Imperial College London
Department of Computing

Socio-Economic Aware Data Forwarding in Mobile Sensing Networks and Systems

Usman Adeel

Submitted in part fulfilment of the requirements for the degree of
Doctor of Philosophy in Computing of Imperial College London and
the Diploma of Imperial College, April 2015

Abstract

The vision for smart sustainable cities is one whereby urban sensing is core to optimising city operation which in turn improves citizen contentment. Wireless Sensor Networks are envisioned to become pervasive form of data collection and analysis for smart cities but deployment of millions of inter-connected sensors in a city can be cost-prohibitive. Given the ubiquity and ever-increasing capabilities of sensor-rich mobile devices, Wireless Sensor Networks with Mobile Phones (WSN-MP) provide a highly flexible and ready-made wireless infrastructure for future smart cities. In a WSN-MP, mobile phones not only generate the sensing data but also relay the data using cellular communication or short range opportunistic communication. The largest challenge here is the efficient transmission of potentially huge volumes of sensor data over sometimes meagre or faulty communications networks in a cost-effective way.

This thesis investigates distributed data forwarding schemes in three types of WSN-MP: WSN with mobile sinks (WSN-MS), WSN with mobile relays (WSN-HR) and Mobile Phone Sensing Systems (MPSS). For these dynamic WSN-MP, realistic models are established and distributed algorithms are developed for efficient network performance including data routing and forwarding, sensing rate control and and pricing. This thesis also considered realistic urban sensing issues such as economic incentivisation and demonstrates how social network and mobility awareness improves data transmission. Through simulations and real testbed experiments, it is shown that proposed algorithms perform better than state-of-the-art schemes.

Declaration

I declare that this thesis was composed by myself, that the work contained herein is my own except where explicitly stated otherwise in the text.

Copyright

Acknowledgements

I would like to express my sincere gratitude to my advisor Prof. Julie A. McCann for the continuous support of research, for her patience, motivation, enthusiasm, and immense knowledge. Her guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisor and mentor for my PhD research.

I would like to express deepest thanks to my friend Shusen Yang, who helped in my personal and academic development. Shusen is a mentor and friend, from whom I have learnt the important skill of disciplined critical thinking.

I would also like to thank Yaad Tahir, and Sokratis Kartakis who, as good friends, were always willing to help and give their best suggestions. I graciously thank my colleagues for providing me valuable feedback on my work.

I would also like to thank my parents. They were always supporting me and encouraging me with their best wishes.

Finally, I am thankful to my lovely wife Sara for her constant support, patience and love. She has been a great source of motivation and inspiration throughout my PhD.

Dedicated to my daughter Inaya. You mean the world to me

Contents

List of Tables

List of Figures

Chapter 1

Introduction

Wireless Sensor Networks (WSNs) [204] have gained an increasing importance over the last number of years due to their suitability for monitoring complex physical world phenomena at levels of detail that was previously impossible. Wireless Sensor Networks (WSNs) are distributed systems composed of low cost sensor devices monitoring an aspect of interest such as temperature, humidity, light, acoustics, vibration, and pressure. Sensor networks are rapidly growing due to their large potential in various application areas including ecosystem management, smart homes and buildings, natural hazard monitoring, intelligent transportation, and human behaviour sensing. Current approaches are mainly application-specific, where numbers of static wirelessly-connected sensor nodes are deployed at a place of interest. The sensed information is sent to sink(s) over wireless links through single-hop or multi-hop communication. Research into WSNs has also motivated emerging research areas for general purpose WSNs, such as the Internet of Things (IoTs) [22], Cyber-Physical Systems (CPSs) [154], and smart sustainable cities [148, 161].

1.1 Motivation

Statistics show that the numbers of people living in urban areas is dramatically increasing (http://world.bymap.org/). Like with any overloaded system, when a city is close to capacity,

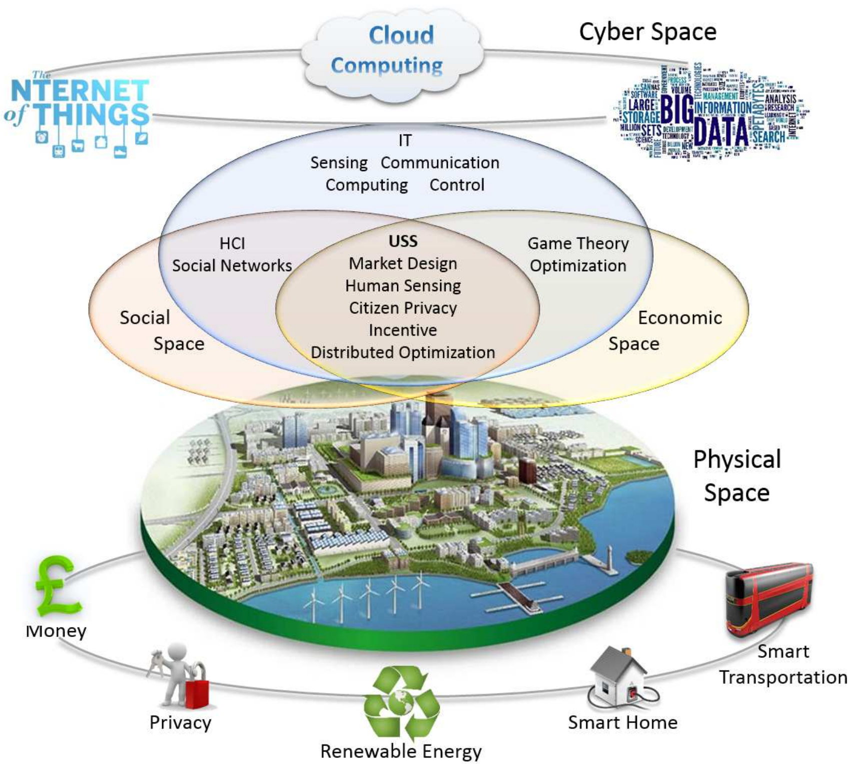


Figure 1.1: Design considerations for WSNs with Mobile Phones (WSN-MP.

services and resources begin to fail. Therefore, to make more efficient use of such resources, it becoming necessary to better understand the state of urban capital and its usage. Key to this knowledge is the ability to instrument the city with mechanisms that monitor it, and react to the data gathered from the physical environment autonomously. Therefore, urban sensing is beginning to become more prolific [148]; from light and temperature sensors in a smart building, water leak detection in the sewer, smarter cars, smart power grids, to noise and air-quality monitoring in the streets.

With the evolution of such smart cities and the IoTs, WSNs will become an increasingly pervasive form of instrumentation for the gathering and analysis of data of all kinds. For these approaches to be effective, autonomic sensor networks require large number of specially designed hardware and a communication infrastructure which makes it expensive. Moreover, fixed infrastructures for such large scale WSN have limitations regarding sensor maintenance, deployment and connectivity.

Therefore, Ubiquitous sensor-rich smartphones are beginning to play an increasingly important role in the evolution of the IoTs, which bridge the digital space to the physical world at a societal scale. Their powerful computing and communication capacities as well as huge market proliferation, and inherent mobility makes WSNs with Mobile Phones (WSN-MP) [63, 109] a much more flexible and cost-effective sensing and communication solution compared with traditional static WSNs.

In a WSN-MP, a mobile phone can act as a sink or relay to collect the data from statically-depolyed sensor nodes; or produce sensor data by sampling its embedded sensors. The main advantages of WSN-MP are.

1. **Powerful sensing and communication platform.** The sensors available on smart phones can monitor a diverse range of human activities and commonly encountered contexts. Apart from that mobile phones can also communicate with each other using widely available interfaces (e.g., wifi-direct, bluetooth, 3G/4G).

2. **Increased coverage.** Due to increasing number of smart phones, WSN-MP can collect data with much higher granularity. Mobility of the humans also enable more penetration and coverage than previously possible.

3. **Cost-effective.** Smart phones enable sensing, collection and communication of data without need of any pre-installed infrastructure. Mobile phones are carried by humans as part of their daily life therefore, mobility of mobile phones can be utilized at no extra cost.

Different to existing sensor networks, WSN-MP bring people *in the loop* because people are no longer just consumers of sensed data; but also are the source of sensed data and responsible for its communication. Furthermore, the mobility of phone users plays an important role in WSN-MP by improving the performance of networks [52] and serving as core requirement of many applications [89]. Due to this human involvement, their social and economic behaviours (e.g. social networks, selfishness etc) have a significant impact on the sensing and communication performance of the WSN-MP. Such behaviors can be exploited in the design of communication schemes due to the close coupling between modern sensor networks and the physical world in which they reside. Evolution of smart sustainable cities and CPS lead to many research issues based on the close relationships of human behaviours to design effective approaches for WSN-MP as shown in Figure 1.1. This thesis investigates how to design, analyse, and evaluate cost-effective sensing and wireless networking schemes for WSN-MP in this big picture.

Figure 1.2: Example of WSN-MP to highlight the challanges.

1.2 Challenges

Here I use a simple example to illustrate challenges of WSN-MP shown in Figure 1.2. Imagine a small-scale static sensor network deployed at the surface of Hyde park lake in London. The sensors monitor water quality and detect if some one falls into the lake (e.g.using sonar sensors). There are many people in the park carrying mobile phones, which sense light conditions at different areas of the park. Free Wifi Access-Points (APs) are located at the corners of Hyde Park. Sensors in the lake cannot send data directly to APs due to their short communication range but they can send the data to other nodes or any mobile phone in their communication range using Wifi-Direct. Mobile phones can also communicate with APs and each other using Wifi-Direct. In addition, mobiles phone can also send data directly through cellular communication (i.e 3G/4G). In this example, we focus on John and William, who are two users carrying a mobile phone. Now we discuss the key issues that can arise in this WSN-MP scenario.

1. Data forwarding to mobile nodes.

A sensor node in the middle of the lake gathered some sensor data. It cannot send the data directly to any mobile phone due to its short communication range. However, it can send to another relay sensor node using multi-hop communication. How can it find a relay node, who has the best chance to forward its data to mobile phone? (e.g, A node near the road)

Current scheme for data collection using mobile sinks can be divided into two categories, those that assume controllable sink mobility and uncontrollable sink mobility [51]. WSN-MP falls into the second category of WSN-MSs where the sink mobility is uncontrollable (it is unlikely that a mobile phone user is willing to move according to a planned trajectory). There are several data collection protocols for WSN-MSs [105, 108, 115, 123, 193]. Lee *et al.* [108] propose a routing protocol based on information potentials [124] and a constructed mobility graph. However, this scheme requires mobility prediction and may suffer from heavy communication overheads when there are a large number of mobile sinks. WARP [117] and the routing protocol developed by Li *et al.* [123] are based on fast and efficient routing structure repairs, but are still limited to single mobile sink settings. Data Stashing [115] supports multiple mobile sinks but requires mobility prediction, network-wide flooding, and linear programming solving on each node, leading to large communication and computational overheads. In chapter 3, we propose a novel routing metric Contact-Aware ETX (CA-ETX) to determine the ability of a static node to forward sensor data to mobile nodes. Using CA-ETX, we developed opportunistic data forwarding scheme to reduce end-to-end delay with out any mobility prediction and communication overhead.

2. Support heterogeneous applications with different delay requirements.

John received critical data from a sensor about falling some one in the lake while has to be sent urgently. He also has periodic light quality data sensed by his phone. Should John send all data using faster and reliable cellular connection?

Most current mobile sensing applications transmit mobile sensor data to the server through cellular networks. With the increasing popularity of mobile sensing applications, this simple solution will suffer from not only significant battery [167, 190] and 3G/4G financial costs [128] to the phone users, but also will produce heavy traffic load on the underlying bandwidth-limited cellular networks, especially for the applications that require continuous sensing with fine granularity (e.g. [155]). Besides expensive cellular communication, current smart phones are being equipped with more and more short-range wireless technologies such as WiFi, WiFi direct, and Bluetooth 4.0, which enable opportunistic phone-to-phone and phone-to-sink communications (e.g. through WiFi routers). Due to their low energy and financial cost, it is promising to exploit the potential of short-range communications, especially for delay-tolerant mobile sensing. In recent years, many research efforts such as [153] and Delay-Tolerant Networks (DTN) [93] aim at exploring communications among mobile devices through short-range communication, where a continuous end-to-end connectivity may not be possible. Current opportunistic data forwarding approaches either suffer from high cost due to the replication of messages or high storage and maintenance cost to store and update history information and routing paths. Back-pressure routing [56] is a promising approach due to low overhead and optimal throughput. The main disadvantage of opportunistic schemes is high delay, which make them unsuitable to forward critical sensor data. To address this problem, there is a need for hybrid architecture to support both real-time and delay-tolerant urban sensing applications via the seamless integration of inexpensive short-range opportunistic transmissions and reliable long-range cellular radios. Forwarding scheme to support heterogeneous applications is developed in chapter 4 where cost-benefit analysis is used to decide to forward data through cellular communication or short-range communication.

3. How to exploit mobility patterns and social behaviours for efficient data transmission.

Lets consider a scenario, where John and William approach a sensor near the lake. Should sensor send the data to John or send it to William? William works at the cafe near the lake while John's friends are currently sitting close to the Hyde Park corner. Can this information help to determine better relay for the data?

Social network theory [162] is studied as a useful tool to model the structure, properties and relationships among people. For example, community is a property of social networks that groups people with similar interests, who meet frequently, spend time together and they are more willing to share information and resources. The knowledge of underlying social structures and ties among the users, understanding the mobility patterns and predicting contact opportunities is the key to design the effective opportunistic network services; such as data forwarding, resource allocation and resource utilization. The networks consisting of mobile phones are human-centric. Understanding social characteristic of humans depends on many factors, such as personal preferences, relationships, interests and popularity. These characteristics define social properties that include community, centrality, similarity and mobility patterns. These social properties can be used to design efficient networking solutions. In this context, social-aware forwarding protocols and dissemination strategies are the most important components. Social network metrics such as centrality and community structure have been used for many opportunistic routing schemes [45, 66, 67] in Delay Tolerant Networks (DTN) [60]. However, all of them focus on packet routing (i.e. unicasting or multicasting a single packet or multiple packets) rather than the flow routing. Flow routing is required by many sensing applications, who send packets to the sink continuously. Chapter 5 proposes a sensing paradigm for WSN-MP, where mobile phones are used to relay the sensor data from sensor nodes to sinks. A data forwarding metric, Sink-Aware (SA) centrality, is proposed to measure the potential sensor data forwarding ability of human with mobile phones. This scheme determines better relays for the data.

4. Incentivize mobile devices to sense and forward sensor data.

Let go back to our example. A sensor wants to send data to John. Why would John carry and forward the data for the network, which will cost him in terms of battery and computation? Similarly, he is not willing to use his mobile phone for sensing light readings. Can network provide him any incentives to encourage user participation?

Multi-hop data forwarding in communication networks depends on the cooperation from the participant nodes (i.e. willingness to help forward messages for other nodes). This cooperation is achievable in infrastructure based networks where all nodes belong to a single user or organization. In WSN-MP, many nodes such as human with mobile phones are not willing to forward messages to conserve their limited resources (e.g. battery power and buffer) and increase their own benefits (e.g. reducing communication cost). The non-cooperative behaviour of nodes can have negative impact on the performance of data forwarding algorithms in a multi-hop network [88]. Moreover, user participation is the most important element in mobile sensing applications to achieve good service quality. Most of the existing mobile sensing applications are based on voluntary participation of users to sense the data. In reality, a user may not be interested in participating in mobile phone sensing due to its potential cost (in terms of battery usage and communication cost) and privacy considerations. Traditional economic models generally model these non-cooperative nodes or users as rationally-selfish actors [135] , who always act for their own interest to maximize its profits. Economic considerations such as incentives are central to the design of any useful data dissemination schemes and, maximizing user participation in mobile sensing applications. This is a challenging problem that has attracted attention from many researchers. Many incentive-aware schemes [192] [168] [33] have been proposed to for data forwarding in mobile networks with *selfish* nodes but they do not consider pricing user's effort to encourage participation in mobile sensing applications. To stimulate the participation of mobile users to sense data and forward data generated by others, we developed fully distributed, joint routing and pricing schemes to incentivize phone owners in chapter 4, 5 and 6.

5. Reward or punish mobile devices based on their effect on the performance of the network.

Imagine John is willing to participate in collection and forwarding of the data in order to earn credits. He is subscribed to a cheap cellular data package. He found out that he can earn more profit if he sends only his own sensed data through cellular connection than forwarding data from others. So he informs the network that he don't have any Wifi-direct to maximize his profits. This reduces the performance of the network but key question is that will he be able to get more profit?

Strategic self-interest users are those individuals who try to maximize their profits through their strategic actions learned from participation in a system. Strategic self-interest users can subvert the network behaviour to increase their own profits. Mechanism design [91, 146, 169] is concerned with how to make a global decision with desirable properties in systems consisting of strategic self-interest individuals with private information where misreporting of private information can effect the functioning of the system. Recent theoretical work [152, 179] on distributed Vickrey-Clarke-Groves (VCG) mechanisms enables the faithful implementation of algorithms producing desired outcomes in a distributed way. VCG discourages manipulation of the system by charging each individual the harm they cause. However, these approaches focus on deterministic rather than stochastic systems, therefore cannot be applied in highly dynamic WSN-MP scenario. We used mechanism design to develop a fully distributed taxation scheme, which can provide subsidy or impose tax on mobile users based on their positive role. Our scheme discourage mobile users, who can subvert the network behaviour to increase their own profits in chapter 6.

1.3 Our Approach

Although few existing work for Mobile WSNs can be applied to WSN with Mobile Phones (WSN-MP), but most of them are either practical work based on heuristic approaches without performance guarantees, or theoretical approaches with unrealistic assumptions and high complexity. This thesis aims to bridge the gap between theory and practice for data forwarding in WSN-MP and to develop fully distributed algorithms and protocols, guided by both quantitative insights gained from mathematical theories and the practical principles of real WSN-MP.

Also, from recent research, it is apparent that the design of a generic adaptive scheme for sensing and data forwarding leveraging the ubiquity of mobile devices can be challenging. Therefore, many individual application oriented or characteristic oriented design solutions exist. With the popularity of mobile phone sensing applications and initiatives like smart cities, cyber-physical systems, it is envisioned that many sensing applications (e.g, environmental monitoring [7, 156], smart transportation [184], social sensing [35]) will use mobile phones to sense or transmit data simultaneously. Therefore, we aim to design comprehensive solutions to support multi-users and multiple applications in a WSN-MP. Another key challenge is to design network algorithms that are adaptive to various WSN-MP dynamics. In this thesis, we focussed on distributed approaches, which are more scalable and adaptive to the network dynamics compared with centralised schemes.

We addressed the challenges and the limitations of the existing work discussed in previous sections to develop fully distributed sensing and networking algorithms for WSN-MP based on the application of coupled communication networks and socio-economic behaviours of people.

1.4 Contributions of This Thesis

The broad aims of this thesis are to develop efficient distributed algorithms for network problems in WSN-MP. Specifically, we addressed three kinds of WSN-MP.

1. WSN-MP with mobile phones to collect data from static WSN in chapter 3 reffered to as WSN with Mobile Sink (WSN-MS).
2. WSN-MP with mobile phones to relay data from static sensors to sinks in chapter 5 reffered to as WSN with Human Relay (WSN-HR).
3. WSN-MP with mobile phones to sense and relay the data to the sinks in chapter 4 and 6 reffered to as Mobile Phone Sensing System (MPSS).

This thesis makes the following contributions to the state-of-the-art:

- Chapter 3 studies ubiquitous sensor data collection in large-scale WSNs with mobile sinks. Based on queuing analysis theory, a novel routing metric, called CA-ETX, is proposed to estimate the packet transmission delay caused by both link unreliability and intermediate connectivity. By integrating CA-ETX into Lyapunov optimisation theory, a throughput-optimal data collection algorithm is then developed. Testbed experiments and extensive simulations show that the proposed algorithm can achieve much better performance than current state-of-the-art approaches, in terms of energy consumption, end-to-end delay, scalability, and sensitivity to sink movement speeds. In addition, CA-ETX can work seamlessly and synchronously with the well-known routing metric, ETX [49], illustrating that existing ETX-based routing protocols, such as the de-facto TinyOS routing standard CTP [72] and IETF IPv6 Routing Protocol RPL [11], can be easily applied to WSN-MSs, using the CA-ETX.

- Chapter 4 develops a network architecture to provide a cost-effective networking service for ubiquitous mobile phone sensing, where mobile phones are used to sense and relay the sensor data to the sink. A joint pricing and routing scheme is proposed to support both real-time and delay-tolerant mobile sensing applications through the seamless integration of cellular and short-range communications of mobile phones. By trading mobile sensor data in a virtual free market, this scheme provides an incentive system for phone owners, while achieving network throughput optimality and minimizing phone users total costs in terms of their 3G budget and battery levels.

- Chapter 5 proposes a sensing paradigm for ubiquitous sensing, where mobile phones are used to relay the sensor data from sensor nodes to sinks. By exploiting underlying social and economic networks in context of human relays, a socio-economic aware data forwarding scheme is designed. A novel data forwarding metric, Sink-Aware (SA) centrality, is proposed to measure the potential sensor data forwarding ability of mobile relays. By combining complex network theory and wireless sensing, a distributed algorithm is developed for joint rate control, opportunistic routing, and resource pricing. This algorithm not only maximises global social profits, but also manages to incentivize selfish phone users to participate.

- Chapter 6 develops a cost-effective data collection solution and faithful market design for MPSS. It considers mobile phones to sense data and also relay sensed data, using hybrid cellular and opportunistic short-range wireless communications. An adaptive and distributed algorithm OptMPSS is developed to maximize phone user financial rewards accounting for their costs across the MPSS. Based on distributed mechanism design theories, BMT scheme is proposed to incentivize phone users to faithfully implement OptMPSS, by imposing taxes or providing subsidies for each phone user. Experiments with Android phones and trace-driven simulations demonstrate that this approach manages to improve the system performance significantly while confirming that our system encourage the faithful implementation of BMT.

1.5 Publications

The work presented in this thesis is supported by the following publications.

1. Usman Adeel, Shusen Yang and Julie A. McCann. **Self-Optimizing Citizen-centric Mobile Urban Sensing Systems.** *International Conference on Autonomic Computing (ICAC), 2014*

2. Shusen Yang, Usman Adeel and Julie A. McCann. **Selfish Mules: Social Profit Maximization in Sparse Sensornets Using Rationally-Selfish Human Relays**, *IEEE Journal on Selected Areas in Communications (JSAC), vol.31, no.6, pp. 1124 - 1134, 2013.*

3. Shusen Yang, Usman Adeel and Julie A. McCann. **Practical Opportunistic Data Collection in Wireless Sensor Networks with Mobile Sinks**, *IEEE Transactions on Mobile Computing(TMC), Revised.*

4. Usman Adeel, Shusen Yang and Julie A. McCann. **A cost effective data forwarding scheme for heterogeneous data in Mobile Sensing System**, *ACM Transactions on Autonomous and Adaptive Systems (TAAS), under review*

5. Shusen Yang, Usman Adeel and Julie A. McCann. **Backpressure Meets Taxes: Faithful Data Collection in Stochastic Mobile Phone Sensing Systems**. *IEEE INFOCOM, 2015 (accepted to appear)*

Chapter 2

Background

Network technologies in the last decade have revolutionized the ways in which persons and large organizations communicate and exchange information among themselves and organize their activities. In the near future, we will observe another revolution that includes surveillance and control of the physical world.

Wireless Sensor Networks (WSNs) are a new class of distributed system, that are an integral part of the physical space they inhabit. Unlike most systems, which work primarily with data created by humans, sensor networks capture the state of the environment around them. This bridge to the physical world has captured the attention and imagination of many researchers, leading to a broad spectrum of ideas, from environmental protection to military applications.

Future wireless sensor networks are envisioned to consist of hundreds or thousands of sensor nodes communicating over a wireless channel, performing distributed sensing and collaborative data processing tasks for a variety of vital applications. Such sensor networks will improve the safety of our buildings and highways, enhance the viability of wildlife habitats, shorten disaster response times, and contribute in many other vital functions. We can imagine ad hoc sensor networks deployed for various kinds of applications, providing continuous and spatially dense observation of biological, environmental and artificial systems.

Traditional WSNs consisted of large number of nodes sprinkled over an area of interest. They were initially also termed as *Smart Dust* [100] due to their small size and high density . The nodes are equipped with limited battery, which makes energy a scarce resource in these networks. All the nodes in WSNs collect data from the surroundings and send it to base-station node typically through multi-hop routing algorithms. High density of the nodes in the network area is necessary for multi-hop forwarding as the transmission range of the low powered nodes is not enough to communicate with the other nodes at long distance.

Urban Sensing Systems come to complement previous efforts to deploy wireless sensor networks to sense our environment, providing a vision for smart sustainable cities. To this end, there have been many initiatives that involve wireless sensing, cyber-physical systems and the Internet of Things (IOTs).

The work in this thesis deviates from this traditional WSNs in the sense that it focuses on sensor networks with mobile nodes. Specifically, we investigate WSNs with Mobile Phones (WSN-MP) in this thesis. A similar direction has been already taken by other researchers. This chapter briefly introduces the necessary background related to the main work presented in this thesis. In this chapter we discuss the general motivation of mobile WSNs in section 2.1, we analyze the characteristics of WSN-MP in section2.3. Opportunistic data forwarding approaches related to our work are presented in section 2.4. We defer the analysis of the state-of-art topics particularly related to our contribution to the corresponding chapters, as well as clear definitions of the terms and symbols used.

2.1 Mobility in WSNs

In many application scenarios of Wireless Sensor Networks, the deployment of some or all mobile nodes is possible and it can greatly enhance the utility and functionality of network. There are many ways mobility can benefit the system.

- The mobility of nodes is the functional requirement of various WSN applications, for instance in exploration or monitoring of moving entities. In many systems, mobile sensors are attached to animals, humans and moving assets. Similarly, autonomous robots could be of great use in search and rescue missions. ZebraNet project [98] is one of the examples, where sensors were mounted on zebras to understand the migration pattern of Zebras. CargoNet [134] is another platform designed to monitor goods and freight activities. Environmental monitoring using unmanned aerial vehicles (UAV's) [14] and vehicles [89] is another representative example of mobile WSN's.

- The mobility of nodes can also help to improve the performance and overcome the shortcomings of static sensor networks. Mobile nodes can be deployed to reduce infrastructure costs such as installation costs and maintenance costs. Mobile nodes can respond to dynamic conditions and requirements by rearranging themselves to optimize the performance goals of the network such as connectivity, coverage etc. [17,70] address the solutions to the coverage problem using mobile nodes.

- The mobility of nodes can also optimize utilization in resource constrained systems. In sparse or disconnected networks, mobile nodes can visit static nodes and collect data using one-hop transmissions. Data mule approach presented in [94] focuses on ensuring the connectivity in the sparse networks by providing another layer of mobile nodes for data retrieval from the disconnected sub networks or different clouds of nodes. For underwater WSNs, [25] determines an optimal collection path for autonomous underwater vehicles (AUVs) to collect maximum information from sparse nodes and deliver it to sink.

- Mobile nodes or sinks can be relocated to balance energy consumption in the network. Relocation also reduces the funnelling effect for the nodes near the sink [16]. Studies have shown that the mobile nodes significantly reduce energy consumption and increase the lifetime of the network [131, 191, 196]. [26] used a distributed approach to maximize the lifetime of WSNs through controlling and coordinating mobility of multiple sinks.

  Furthermore, mobility can be used to increase the overall throughput of a network at the expense of the delivery delay [52].

The mobility of wireless sensor networks not only provide solution to meet the requirements of certain applications but also improves the performance of networks. In our work, we have focused on using mobility to collect and relay data in sensor networks. The deployment of mobile nodes is feasible and useful in many application scenarios in urban environment, ranging from the environmental monitoring and public safety applications, to the industry, healthcare and vehicular applications.

In chapter 3 we study how to improve the delay and throughput performance for delay-tolerant data collection applications in Wireless Sensor Networks with Mobile Sinks (WSN-MSs) whereas chapter 4, 5 and chapter 6 focus on leveraging human mobility for sensing and data forwarding. In the next sections, we focus on WSNs with mobile nodes and Mobile Phones. We also present various challenges brought about as a result of introducing mobile nodes in the network.

2.2 Data Collection in WSNs with Mobile Nodes

In WSNs with mobile nodes, mobile nodes are used to collect data from static sensor nodes. A sensor node can either send data directly to a mobile node as it passes by or it can send data vial local multi-hop routing to other static sensor nodes who currently have contact with, or who have better chance to be in contact with mobile node in the future.

Data Collection in WSNs with mobile nodes can be divided into these sub-problems.

1. Mobile Node Discovery
2. Data Transfer to Mobile Node
3. Local Multi-hop Routing

2.2.1 Mobile Node Discovery

A sensor node is required to detect the presence of mobile node in its communication range in order to transfer data. The discovery protocols aim to detect mobile nodes in less time while consuming less amount of energy.

The existing approaches for detection of mobile nodes in WSNs [94] [160] are based upon periodic listening. In these approaches, mobile node sends periodic beacons to notify its presence. Sensor nodes in the communication range of the mobile node receive the beacon and start transferring the data. If a sensor node do not receive any beacon message, it goes to sleep. Discovery parameters and duty-cycle needs are defined [18] to ensure the discovery of the mobile node is independent of the duty-cycle scheduling of the sensor node. These approaches are not optimal as they do not define the solution for timely detection of the mobile node in duty-cycled networks.

The information about the mobility of the mobile nodes can be taken into the account for timely discovery of the mobile nodes.

Mobile node discovery becomes more challenging, when mobility is unpredictable or network is duty cycled. In [99] a framework for energy-conservation is presented which uses the mobility pattern of the mobile nodes for discovery in opportunistic networks. In [101] simple periodic wake up and sleep scheme is used for discovery and stop-and-wait protocols are used for data collection. Another scheme use controlled mobility in [170] and strict mobility patterns for mobile node discovery.

The architecture presented in [94] considers the Mules, which are not controlled and move randomly in the area. It addressed both discovery and data transfer phase and evaluates different mobility pattens. This approach considers duty-cycle operation of the nodes but it does not present any specific MAC protocol for data transfer.

A window based ARQ transmission scheme is studied in [20] and [19] which shows better results than stop-and-wait protocols but it does not considers the effect of discovery phase on subsequent data-transfer phase. In [114], the model is presented to derive overall energy efficiency considering the combined effect of discovery and data transfer phase which considers simple asynchronous scheme for discovery and ARQ based protocol for data transfer.

Prediction of Mobile Node

The prediction of the time of arrival of the mobile node is relatively an unexplored area. To discover the mobile node immediately when it enters in the communication range can be done by predicting the time of arrival of the Mule. If the prediction is absolutely correct, *data transfer phase* can start immediately after the arrival of mobile node in the communication range.

The prediction of arrival time is simple if mobile node is travelling along a pre-defined trajectory with fixed speed. The nodes can calculate the arrival time from the speed of mobile node and length of the path (distance). The predictable mobility of mobile node is used to reduce energy consumption in [28] and [99], where sensor nodes sleep until next expected arrival of the mobile node. The scheduled wake up of the nodes is possible because the sensor nodes have correct information about the time of arrival of mobile nodes. These approaches assume strict synchronization of the nodes and a strict scheduling of the mobile nodes which is usually difficult to hold unless we assume controlled mobility.

The approach of using multiple radios can be used to find accurate information about the time of arrival of mobile nodes in duty-cycle WSNs. One low-power radio stays awake to detect the mobile node and another high-power radio operates on its duty-cycle to transfer the data. When the low power radio receives the beacon from the mobile node, it wakes up the sensor node and the node transfers the data using high-power radio. Another approach is to use wake up messages sent by mobile nodes which have enough power to generate an interrupt at the sensor node. The interrupt can wake up the sensor node and start sending the data. These approaches require special hardware which is not generally available for commercial solutions for WSNs.

Mobility Models

The analysis of different mobility models plays important role in the study of mobile networks. In opportunistic networks, mobility models are very important, because mobility is an integral part of the network to deliver the messages. *Contact time* (The time spent by mobile node in the communication range of a node) and *inter-connect time* (time in which no mobile node is in the communication range of the node) can be calculated by using mobility models which can be helpful for the efficient design of routing techniques.

The mobility of the mobile node can be controlled or uncontrolled depending upon design of the network. In controlled mobility, the motion of mobile node can be used to design efficient data transfer protocols because the trajectory and speed of the mobile node can be controlled. This reduces the complexity of mobile node discovery because mobile nodes can be scheduled to visit certain nodes at specific times. Furthermore, mobile nodes can stop to collect data until a node empties its buffer. The different approaches for optimal scheduling of the mobile nodes under controlled mobility are studied in [174] and [175].

Uncontrolled mobility can be divided in to deterministic and random mobility. In the deterministic mobility, mobile nodes enter in the communication range of a particular node at specific or periodic times. The Buses or trains acting as mobile nodes usually represent deterministic uncontrolled mobility pattern. Study the traffic pattens of Bus-based networks is presented in [194] and [208] and they used the information about mobility for efficient communication between the nodes mounted on the Buses.

In random mobility pattern, the contacts of mobile nodes with sensor nodes do not take place regularly but with some distribution probability. [12] provides a close approximation for Random Waypoint (RWP) and Random Direction (RD) mobility models under typical opportunistic network settings. They calculated the contact time T_c for the RWP and RD model as

$$T_c = \frac{\pi r}{2v}$$

These results assume that speed of the mobile node v does not change and the radio range of the sensor node $r << a$ where a is the width or diameter of network area. Similarly inter-connect times are approximated using exponential distribution in these models. The study of mobility models can be used for the efficient discovery of the mobile nodes by increasing the duty-cycle of the nodes in the communication range of mobile node to 100 % before expected arrival of a mobile node.

MACs

The optimization of MAC protocols remain a very important area in wireless sensor networks. Researchers have proposed many MAC protocols for wireless sensor networks depending upon the nature of the network and applications. MACs can be broadly categorized into synchronous and asynchronous. Synchronous protocols assume that nodes in the network are synchronized. Basic protocols like S-Mac [201] periodically listen and sleep at its scheduled time. Some improvements are made in [178] [202] [176] . Asynchronous MAC's do not require nodes to be synchronized and more flexible towards traffic patterns and duty cycles. In Preamble based asynchronous MAC's [30] [187], sender sends a preamble message to the intended receiver and starts sending the data after recieving acknowledgement from the receiver. The other type of Asynchronous MAC's is receiver initiated. In receiver initiated MAC's such as RI-MAC [177] and O-MAC [37], the receiver node informs the sender when it is ready to receive data and sender start sending the data.

The existing approaches [18] uses the basic asynchronous protocol to discover the arrival of mobile node. The mobile node sends periodic beacons continuously and sensor nodes periodically wake up to listen to the beacon.

2.2.2 Data Transfer

Once the mobile node is discovered by the sensor node, data transfer phase starts immediately. The amount of data that can be transferred to the mobile node after the discovery phase is a random variable and it depends upon factors such as communication range of the node, speed of the mobile node and data communication rate. Data transfer protocols aim for efficient utilization of the remaining contact time by maximizing the total number of messages transferred and minimizing the energy consumption during data transfer phase.

Efficient utilization of contact time

In the model presented in [94] the amount of data K which can be transferred can be given by

$$K = CT * B$$

where CT is contact time in which mobile node is in the communication range of a node and B is radio transfer rate. Average contact time CT can be computed as

$$CT = (\frac{\pi}{2}\frac{r}{v})$$

where r is the radio range of the sensor node and v is velocity of the Mule.

There requirement of efficient data transfer protocol to transmit data from the nodes to the mobile node has been discussed in many approaches . Opportunistic Aloha MAC [188] is a protocol specifically designed for wireless sensor networks with flying vehicles as mobile data collectors. Simple ARQ-based data transfer protocol is analysed in [20] and Adaptive Data Transfer (ADT) protocol is proposed which adjusts communication parameters based upon previous history.

Multiple Mobile Nodes

The approach of using multiple mobile nodes in WSN is studied in [95]. The use of multiple mobile nodes can be due to the the high density of the nodes in an area where buffer size of the relay node is filled before a mobile node arrives again to collect the data. The total number of mobile nodes required can be given as [95]

Required number of mobile nodes

$$= BufferFillTime/RTT$$

where RTT is round trip time of the mobile node.

Also, the QoS requirements of the application such as latency can lead to deploy more mobile nodes in the network. If there are multiple mobile nodes in the transmission range of a sensor node at the same time, the sensor node can choose the mobile node depending upon their direction, expected contact time or signal strength. Data-delivery to more than one mobile nodes simultaneously is discussed in [21] , where sparsely deployed nodes transfer the blocks of data to multiple mobile nodes. Hybrid Interleaved data transfer protocol (HI) is presented, which uses encoding techniques to improve data transfer by adding redundant information in the source data.

Speed of the Mobile Nodes

The speed of the mobile node plays an important role in the data transfer phase. If the the speed of the mobile node is fast then the contact time will be less and it may not be possible for a sensor node to transfer its data. On the other hand slow moving mobile nodes can cause the overflow of the buffers of the sensor nodes because its round trip time (RTT) increases and RTT must be less than time taken by node to fill the buffer. The objective is to find the optimal speed such that contact time is enough to transfer the data and RTT is minimum. This problem has been studied in [174] and [175] through controlled mobility. They have formulated the data collection in wireless sensor networks as a scheduling problem with location and time constraints.

Overhead of Control messages

The contact time in which a mobile node is in the communication range of the sensor node is not fixed. Therefore, a sensor node has to constantly monitor if the Mule is still in communication range. In most of the approaches like [114] [21] , which use the ARQ based protocol for data transfer, Acks received from the mobile nodes confirm the presence of the mobile node in the communication range of sensor node. Sensor node assumes that the mobile node is no longer in the communication range if some pre-defined number of Acks are not received. Sensor node then switches to normal duty-cycle and waits for the discovery of mobile node.

The RTS/CTS (CSMA/CA)packets are used to avoid collisions in the dense networks. The sparse networks, in which we can assume only one node in the communication range of the mobile node and vice versa at a given time, the RTS/CTS messages can be avoided. In the dense networks, RTS/CTS messages can decrease the efficiency of data collection by mobile node since the size of these control packets is significant. The other protocols e.g., receiver initiated can decrease the control message overhead.

2.2.3 Local Multi-hop Routing

Depending upon the nature of applications and the terrain, it may not be feasible for mobile node to reach in the communication range of each sensor node in the network. The sensor nodes which are deployed near to the trajectory of mobile node can sent the data directly to the mobile node when it passes nearby. These sensor nodes which are in the communication range of mobile node act as relay nodes. The nodes which are far from the trajectory of the mobile node send the data to the relay nodes, which buffer the data collected from the neighbour nodes and forward it to mobile node, when it passes by.

Traditional multi-hop approaches in wireless sensor networks try to find the best path between source and destinations. The best path can be based on simply the number of hops to the destination or combined with other factors which can contribute in energy consumption of the network such as ETX (expected transmission count) [47] and RSSI (Received Signal Strength Indicator) [9]. This best path reduces the consumption of the energy and increases the life time of the network. WSNs with mobile nodes can use these approaches to build local trees in the network with each Relay node acting as local sink for the nearby nodes. Simple routing approaches e.g., Directed diffusion [90] or tree based approaches e.g., CTP [73] can be used to forward data from the sensor nodes to the Relay nodes via best path.

In WSNs with mobile nodes, the shortest path to the mobile node can change with the time because the movement of mobile node changes the destination of the data constantly. When there are more than one mobile nodes in the network, this problem becomes more challenging. If the trajectory of the mobile node can be predicted than We can compute the time of arrival of mobile node at certain point based upon its speed. This can lead to optimal design for multi-hop routing of the data to the mobile node with shortest path routing.

Path discovery

In WSNs with mobile nodes, mobile nodes can either follow a pre-determined route or they can move randomly in the network. Some approaches like [101] [170] assume that mobile node is moving along a predefined path. Similarly, [170] and [95] assumes that the motion of the mobile node can be controlled. When the movement of the mobile node is pre-determined or controlled, the base station can broadcast this information to all the nodes. If another mobile node is added to the network, updated path information can be sent again by the base station. If the base station do not have the path information, mobile node traverses the path in the start-up phase of the network. During this traversal, mobile nodes do not collect any data from the sensor nodes. The sensor nodes, who detect the mobile node, become relay nodes and announce it to the neighbour sensor nodes. The relay nodes act as the mini-sinks for the sensor nodes which are far from the trajectory of the mobile node. Many approaches are present in the literature which address the routing from the nodes to relay nodes [65] [189].

The approach of Data-stashing [115] is proposed for energy-efficient data delivery to mobile nodes through trajectory prediction . Data is stashed along the predicted trajectory of the mobile node instead of routing directly to the mobile node at its current position. Using previous history of trajectory, cluster matching and alignment is applied to predict the future trajectory and optimal relay node is found by linear programming.

Load Balancing

Traditional wireless sensor networks suffer the problem of unbalanced load on the network. The nodes near to the sink die quickly because they all the data from the network passes through them. Many approaches consider the use of multiple sinks to balance the load on the network. Similarly data aggregation is used in few approaches to reduce the data to be sent to the upper levels of the hierarchy.

This problem also exists in WSNs with mobile nodes, which uses local multi-hop routing to send data to the relay nodes. The relay nodes become the hotspot and the consumption of the energy of the relay node is higher than the other nodes.

An approach for the collection of the data is discussed in [95] which uses multiple mobile nodes to balance the load on the relay nodes. The sensor nodes which are part of the tree of two or more relay nodes can decide to choose the path based upon the number of children of each relay node.

Data balancing is also important due to the limitation of the buffer size of relay nodes. If a relay node receives data from a large number of sensor nodes, its buffer will be filled in small time. if the Round Trip Time (RTT) of the mobile node is greater than *BufferFillTime* then the data will be lost. In controlled mobility, RTT is give by

$$\left(\frac{l}{s}\right) + (NumNodes * ServiceTime) + \left(\frac{l}{s}\right)$$

where area of the networks is $l*l$ and s is the speed of the mobile node. *ServiceTime* can be computed by

$$BufferSize/CommunicationDataRate$$

Multiple mobile nodes can be deployed to avoid overflow of the relay buffers. The optimal number of mobile nodes is given by RTT/*BufferFillTime*. Similarly, data can be forwarded to the other relay nodes if the buffer becomes full to avoid data loss.

2.3 WSNs with Mobile Phones (WSN-MP)

Sensing is a crucial component for smart infrastructures, which can monitor themselves autonomously and make intelligent decisions to operate effectively. Traditional sensor networks consisted of ubiquitous placement static nodes to capture and report the state of the environment around them. For these approaches to be effective, autonomic sensor networks require large number of specially designed hardware and a communication infrastructure which makes it expensive and inflexible for future smart cities.

Given the ubiquity and ever-increasing capabilities of sensor-rich mobile devices, such as smart phones and tablets, WSN-MP can be seen as the backbone for future urban sensing in smart cities. As mobile devices can communicate with each other using widely available interfaces, wifi-direct or cellular communication; therefore, millions of smart phones and devices can be leveraged to sense, collect and communicate data without the need to deploy and maintain thousands of static sensors.

The integration of sensing, computing and communication capabilities in mobile devices has turned them into a cost effective computing and sensing platform. These devices can serve as a bridge to other devices or generate information about themselves and the environment. These mobile devices can play different roles in WSN-MP such as

- Sinks to collect data from static networks.
- Data Mules to relay data from sensors to the sinks.
- Sensors to sense data from the environment.

Proliferation of these sensor-rich mobile devices along with collection of ubiquitous sensor nodes are envisioned to constitute a powerful WSN-MP which can be used to understand and analyse many interesting phenomena of the physical world as shown in Figure 2.1.

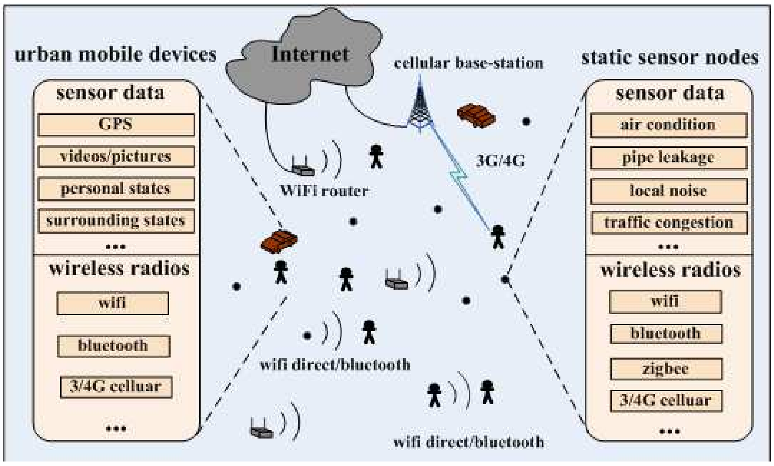


Figure 2.1: Conceptual illustration of the WSNs with Mobile Phones.

2.3.1 WSN-MP Characteristics

Compared with local-area wireless sensor networks (WSNs) [205] such as smart homes or buildings, WSN-MP has the following distinguished characteristics:

- **Large-scale and diverse sensing.** Different from local-area WSNs that consist of tens or hundreds of sensors, billions of low-powered and low-cost sensors could be implemented and connected to the Internet to monitor every interested element in a city. Therefore, in the future smart sustainable cities, the population of smart sensors is expected to be larger than that of citizens.

- **Heterogeneous data types.** Due to the diversity of sustainable city applications and their potential to require large sensing coverage, sensor data will be of many different formats; ranging from simple data structures for physical readings (e.g. temperature) to complex video frames. In turn this data would differ in lifetime, monetary value, QoS requirements and privacy levels, etc. Such heterogeneity results in much more complex data processing, storage and networking.

- **Huge volumes of sensor data.** WSNs are expected to cover small geographic regions such as smart buildings, and typically consist of tens or hundreds of sensors. In contrast, WSN-MPs have the potential to monitor every interesting element in a city, and then relay sensed data to the Internet. As a result, the WSN-MP is predicted to be one of the major sources of *big data*. For instance, the real-time global positioning system data (GPS) for all the taxis in London would amount to several Giga- or Terabytes of data per second. Transmitting large volumes and high frequency data from sensors to the Internet poses great challenges to the current communication infrastructures.

- **Challenging wireless networking.** Current wireless technologies especially WiFi and 3G/4G cellular have successfully provided wireless data traffic. However, will with fundamental information-theoretic limitations (e.g. [77]), it is still questionable whether the future wireless technologies will be able to support cost-effective networking service for transmitting huge volume of sensor data resulting from the proliferation of sensing systems.

- **Social and economic awareness.** In a WSN-MP, sensors can belong to different organizations (e.g. logistics companies, context aware services etc), public bodies (e.g. the meteorological office, city council, utilities companies etc.) or individuals (e.g. mobile users, asthma sufferers concerned about pollutants etc.). There will not be a one size fits all approach to the design and deployment and moreover integration of future sensing systems. Therefore a system is required that adapts, taking account of the different social and economic motivations of customers and service providers. There is a need to better understand what incentives, data and resource pricing schemes are required, as well as comprehending market design issues. This knowledge can be used to maximise the performance of the system and at the same time optimize its quality of service to all its customers.

In this thesis, we focus on how to design networking services for WSN-MP considering the fore mentioned features.

2.3.2 Mobile Phone Sensing in WSN-MP

In this section, we present a brief review of current mobile sensing research. For recent comprehensive surveys, we refer the reader to [64, 104, 110].

Mobile Phone Sensing at Different Scales

According to different sensing scales, mobile phone sensing applications can be categorized into personal sensing, community sensing, and public sensing, where:

- **Personal Sensing** applications focus on monitoring different aspects of the personal life of individuals and collects information about their activities. The sensor data for such application is private, or may be shared only with highly trustworthy entities. For instance, healthcare applications share data with medical professionals to provide necessary support in case of emergency. SPA [165] is an example of personal sensing application, which is smart phone assisted health care self management system. SPA consist of body area sensor network that collects biomedical and environmental data and send it to remote server for analysis and feedback from healthcare professionals. Nike+ (http://nikeplus.nike.com/plus/) is another example of personal sensing which tracks distance, pace, time and calories burned by the user during a run or walk.

- **Community Sensing** makes use of collaborative sharing within a group of phone users who are interested in common goals or a purpose. In community sensing, systems collect and share information between phone users, their friends and social groups, which can promote interaction among people and improve the efficiency of organizing social activities. For instance, CenceMe [136] allows members of social networks to share data such as current activities or location within a group to allow implicit communication between friends.

- **Public Sensing** applications collect rich and heterogeneous data from a large number of people, which can be mined for a variety of purposes to aid decision processes. Phone user's participation is the heart of all public mobile sensing applications where data can be shared, and can return collective benefits. CarTel [89] is such a public sensing application which collects data from users for traffic mitigation, road surface monitoring and hazard detection.

User's Participation in Mobile Phone Sensing

The participation of the general population in data collection and analysis has led to a computational paradigm shift known as *crowdsourcing*. Mobile crowdsourcing can be seen as a distributed approach to solve a complex problem leveraging the participation of a large group of individuals with mobile devices. Defined by awareness of phone users, mobile phone sensing can be classified into two different sensing paradigms: participatory sensing and opportunistic sensing.

- **Participatory Sensing.** requires active participation from the phone users in terms of collecting and sampling the data. (e.g. manual entry of lowest prices or deals for goods or taking a picture). In participatory sensing, the phone user has complete control regarding the contribution of the sensed data which makes it more meaningful and less personal but places a considerable burden on the phone users. MobiShop [164] is a people-centric participatory sensing application, which shares prices of products amongst users. The users use mobile phone camera to capture digital image of store receipts.

- **Opportunistic Sensing.** shifts the burden of tasks from the phone users to the background sensing system and applications can sense and collect data without active participation of the user (e.g. periodic collection of light readings, location information etc), which makes it more suitable for community/public sensing. Along with the benefits of opportunistic sensing, the opportunistic model introduce many challenges. As users do not directly control the data sharing, privacy of the users remains a major concern. Another challenge is design of incentive schemes, which can persuade users to contribute data in sensing system. Nericell [138] is an example of opportunistic sensing, which use GPS and accelerometer sensors of the phones to detect and locate road bumps.

In our work, we consider multiple kind applications at different scales co-existing in WSN-MP. Such applications are envisioned to play a more important role in urban sensing, where one wishes to better understand large-scale phenomena through each citizen's collaboration. These mobile sensing applications will require the communication of a huge amount of data but traffic in the cellular network is already congested and it is predicted to grow at a very fast rate. For example, Cisco predicted that globally, Compound Annual Growth Rate (CAGR) of mobile data traffic will be 66 percent between 2012 and 2017, reaching 11.2 exabytes per month by 2017 [44]. Similarly, Ericsson [71] have predicted that mobile data traffic will grow 15 times by 2017. Off-loading delay-tolerant data to low(or free) cost, short range communication can release the burden on cellular networks. In chapter 4 , we have proposed a cost-effective solution which can meet the demands of heterogeneous mobile sensing applications while reducing the cellular traffic.

2.3.3 Communication Paradigms for WSN-MP

In a WSN-MP, sensors or mobile devices can transmit its sensor data to the Internet by using the following three communication paradigms: one-hop E2E (end-to-end) communications, multi-hop E2E communications, and opportunistic communications [45, 151] . This section discusses the advantages and disadvantages of using above three communication paradigms in WSN-MP, in terms of cost and performance. A brief comparison of the three communication paradigms are summarized in Table 2.1.

Paradigms	**QoS support**	**capacity**	**delay**	**privacy**	**monetary cost**	**energy efficiency**
One-hop E2E	good	low	low	good	high	high
Multi-hop E2E	moderate	moderate	moderate	moderate	high	low
Opportunistic	poor	high	high	poor	low	low

Table 2.1: Comparison of one-hop E2E multi-hop E2E, and Opportunistic communication paradigms.

One-hop E2E Communications

One-hop E2E (End-to-End) communication means that a sensor node can send its sensor data over a one-hop wireless link to a sink. Sinks are generally equipped with wired Internet access capacity, such as the commonly-used WiFi router and cellular base-stations. Due to the one-hop wireless communication and the mature WiFi and cellular technologies, this communication paradigm performs well in QoS guarantees, low E2E delay, and high reliability. However, using one-hop wireless connections for billions of sensors in the future WSN-MP suffers from many critical problems, which can be categorized as follows:

- **Long-range one-hop E2E Communications.**

  Todays cellular networks nearly cover every corner of cities. However, although the wireless spectrum efficiency is continuously improving, the wireless bit rate (or channel capacity) is bounded by the Shannon limit and wireless interference among long-range transmissions [77]. As a result, the wireless bandwidth will still be a scarce and expensive resource in cities consisting of huge number of wireless devices in the future. Emerging technologies (e.g 5G) will further increase the traffic for this limited bandwidth. Due to the explosion of sensor data in the future cities, transmitting sensor data may be more expensive than the worth of sensor data to the WSN-MP. In addition, long-range transmission requires high transmission power, which aggravates the energy scarcity of the low-cost sensor node. For longer deployment lifetime, the more expensive sensor nodes would be required, which have larger-capacity battery or more complex embedded system for energy harvesting [163].

- **Short-range one-hop E2E Communications.**

  To overcome the high power consumption required by long-range communication, short-range E2E communication is used to send data directly to the sinks (WiFi router or IEEE 802.15.4 gateway) through single-hop transmission. Due to the short communication range, a large number of sinks are required to guarantee the large sensing coverage requirement, resulting in high cost of deployment.

Multi-hop E2E Communications

In multi-hop E2E communication, a sensor node can transmit its sensor data to the sink along a wireless path consisting of other sensors or wireless devices. In the past decades, WSNs with tree-based and mesh-based multi-hop topologies have attracted the most attention of both researchers and engineers. Numerous protocols, algorithms, and applications have been developed to improve the sensing efficiency and network lifetime for multi-hop connected networks with IEEE 802.11 and 802.15.4 wireless radios. Today, during the evolution from WSNs to IoT, many efforts are still focusing on supporting this paradigm such as the IETF IPv6 Routing Protocol for Low-power and Lossy Networks (RPL [182]) and applying IPv6 over IEEE 802.15.4 standard (e.g. packet head reduction) by the IETF IPv6 over Low-power WPAN (6LowPAN) working group.

The increasing number of real-world WSN implementations has shown the advantages of the multi-hop E2E communication paradigm in local-area networks. However, for WSNs that requires large-sensing coverage, this paradigm would suffer from the following scalability issues:

- **Poor performance and limited lifetime:**

  Due to the small transmission range of the low-power wireless radios, sensors far away from the sinks have to transmit its sensor data over many faulty and lossy wireless links to the sinks resulting in large delay. In addition, the nodes close to the sink would run out of their batteries fast and become congested as they have to forward the data traffic produced by every sensor node in the network to the base station. As a result, it is very difficult to provide sustainable and reliable networking services to real-time and bandwidth-consuming applications.

- **Flexibility and deployment cost.**

  The deployment topology of an E2E multi-hop sensor network should take both sensing and communication in to account.For example, more important monitoring area should deploy large number of sensors and bottleneck should be avoided for load balancing. For large networks, it would be even prohibiting to make the tradeoff between sensing and communication efficiencies, resulting in additional cost caused by redundant sensors.

The Emerging Opportunistic Communication

Opportunistic communication [45, 151] is an interesting evolution from Mobile Ad hoc Networks (MANETs) and from Delay Tolerant Networks (DTNs). Opportunistic communication enables communication in intermittently connected mobile networks, where the instantaneous E2E(end-to-end) path between a node pair may be absent due to the sparsity of the network or disconnection between connected sub-networks. In city scenarios, WSN-MP can use Opportunistic communication to leverage the prevalence of mobile nodes (e.g. public transport, cars, and individuals) and short-range wireless communication of sensors and mobile devices (e.g. the WiFi Direct radio on smart phones). Sensor data produced by either static sensors or mobile devices can be sent to a static sink (e.g. WiFi router) through other mobile relay nodes in a *carry-and-forward* pattern.

In comparison with E2E (End-to-End)communications, the major advantage of exploiting opportunistic communication paradigm is its flexibility (consider sensing efficiency for deployment topology), low cost (sparse sensor deployment and ubiquitous smart phone sensors), and network throughput [76] (suitable for huge volume of urban sensor data). However, this communication paradigm suffers from large delay as well as poor reliability.

In our work, we have exploited opportunistic communication to leverage the prevalence of mobile nodes (e.g. public transport, cars, and individuals) and short-range wireless communication capabilities of sensors and mobile devices (e.g. the WiFi Direct radio on smart phones). In chapter 5 we study to exploit human mobility in a hybrid sensor and mobile phone network, where we used opportunistic short-range wireless communications between mobile devices to collect and forward data from static sensor nodes to the sink.

2.4 Opportunistic Data Forwarding

Direct Delivery (DD) [76] is the most naive approach for opportunistic forwarding, where a message is delivered only when a source node meets the destination. A message is dropped if its Time To Live (TTL) expires before encounter with destination. Direct delivery is low-cost but it represents the worst case in terms of delivery ratio and delay.

In flooding based algorithms, multiple copies of each message are replicated in the network. In Epidemic forwarding [186], each node sends a copy of the messages in its buffer to every encountered node. Epidemic forwarding guarantees the maximum delivery ratio and finds the delay-optimal path; but it is considered worst in terms of communication costs and storage overheads due to multiple message copies being injected into the network. In Two-Hop-Relay [76], source node sends a message either to the destination or replicates the message to one randomly selected encountered node. Therefore, message forwarding is limited to two hops to reduce message copies on the expense of delay. Spray-and-Wait [171] allows several replicas per message by allowing source and intermediate nodes to replicate (spray) messages to a set of random encountered nodes. In WSN-MP with a huge volume of data, flooding based algorithms suffer from high cost due to the replication of messages.

Knowledge based schemes exchange the information among nodes to make intelligent forwarding decisions. In Seek-and-Focus [171], a node randomly forwards a message to an encountering node in *seek* phase. When a node encounters another node with a more recent encounter time to destination, it shifts to *focus* phase and the message is forwarded to a better candidate node. In PROPHET [125], each node maintains a delivery probability vector based on previous encounters with other nodes . When two nodes encounter each other, they exchange their delivery probability vector. Direct encounter probability vector determines the delivery probabilities of nodes which have never directly encountered each other. MOVE [113] calculates the moving direction based on the Global Positioning System (GPS) to predict the destination of the encountered node. Knowledge based solutions require each node to maintain information about other nodes, which is not feasible in large-scale WSN-MP, which can envisioned to consist of millions of nodes.

Few shortest path routing schemes maintain end to end time varying paths for data forwarding. [50] builds a time varying end-to-end path by estimating the delay between nodes from the information of past contacts. Another approach [127] focuses on hierarchical routing among stationary nodes and other mobile nodes with periodic movement. However, the highly dynamic nature of WSN-MP make it difficult for shortest path approaches to maintain end-to-end paths.

None of above mentioned schemes consider throughput of the network, which is highly important due to huge volume of data in WSN-MP. Backpressure routing [56] is a promising approach because of it achieves optimal throughput and it is highly resilient to the disruption caused by dynamic network. Moreover, backpressure does not require to store any information and it does not compute explicit end-to-end paths. Instead, the routing decision is made based on the difference of queue sizes between two encountered nodes and their link state information.

Due to fully distributed nature of backpressure algorithms, we focus on developing backpressure based forwarding schemes for highly dynamic mobile sensing networks. Some previous works [158,159] use backpressure flow control and routing in disconnected static wireless networks, but they considered fixed gateways with mobile nodes. In contrast, we focus on networks where each sensor node could dynamically serve as a gateway at every opportunistic contact with another node. In addition, Backpressure algorithms suffer from poor delay performance. Several techniques have been developed recently to improve the delay performance of backpressure algorithms [82, 83, 197, 206] but they focus on networks with static topologies only, while our work focus in reducing the delay for mobile networks.

2.4.1 Social-aware Data Forwarding

WSN-MP mainly focuses on utilizing mobile devices for opportunistic data collection in smart cities. Real time data is sent directly to the sink (cellular base station), whereas human mobility is leveraged to forward delay-tolerant data through opportunistic forwarding.

Mobile devices are carried by users, therefore social relationships and behaviours of users have a strong impact on the mobility patterns of the mobile devices. Understanding human mobility is crucial to design efficient schemes for data forwarding for WSN-MP. By exploring social relationships, there are many works that exploited mobility regularities of mobile devices, as well as prediction of contact opportunities for opportunistic data forwarding.

Recent works about social-aware data dissemination in mobile sensor networks can be divided into following categories.

Network structure based Forwarding

Community and centrality are popular properties of social network structures which can be exploited by researchers for opportunistic forwarding in WSN-MP. Human social networks consist of communities (groups) of nodes (people) based on their social relationships. The nodes in a community are connected to each other with strong links and communities are connected to each other via weak links. In addition, some nodes are more central or popular in a community than other nodes. Communities can be seen as the gateway to the destination of relevant data and can assist forwarding approaches to locate destinations. Similarly, a central node can assist the efficient data forwarding inside a community. The movement of the nodes is driven by the strength of the social links between them. A distributed community detection method for opportunistic networking applications is presented in [86]. Another framework [180] is proposed for identification of communities that change with time. In our work, we do not consider the detection of communities but we exploited characteristics of community-based social networks for data forwarding.

Node centrality is used as a metric to forward data in SimBet [48]. [61]] used a semi-Markov analytical model for routing decisions to disseminate data among several communities. Bubble Rap [85] is another social-aware approach which ranks the nodes within their communities and in the network for forwarding decisions. These existing social-aware approaches require high storage and processing capability at a node to handle large amounts of state information by nodes or processing of complex metrics. In chapter 5, we build on these studies and propose a social-aware forwarding scheme for opportunistic mobile social networks.

Mobility Profile based Forwarding

Mobility profile based forwarding avoid the detection of communities and leverage mobility profiles (such as mobility models, mobility-patterns, and mobility-prediction techniques) to decide the best neighbour node for data forwarding.

In earlier work on humans as mobile data collectors, random walk [59] or random waypoint model [96] were used to describe human mobility. Although these mobility models are often used in simulations and analytical models of opportunistic networks for their simplicity but in reality, human mobility is not random and much more complex. Therefore, many recent studies used the human mobility traces and other relevant information (e.g., personal preferences, social information) to understand human mobility and to build realistic models. Periodic patterns of human mobility are identified by many researchers [58, 185]. It was observed that most of the time, humans visit a few well defined geographic locations (such as homes, office, bus-stop) within the network and that the popularity distribution of geographic locations follows a power-law [185]. Many mobility models [207] [106] were derived from these observations to provide more realistic simulations.

Another approach to build mobility profile is by measuring and modelling pair-wise opportunistic contacts between mobile devices [87] [102] [34]. Contact time (contact duration) and inter-contact times (duration between two consecutive contacts) are two important parameters to be considered in order to maximize transfer opportunities using wireless devices carried by humans. Study of pair-wise contacts between wireless devices [87] found heavy-tailed inter-meeting times and contact and inter-contact times followed power-law. Based on this analysis, [38] recommended opportunistic forwarding algorithms between mobile nodes.

In chapter 3, we exploit mobility pattern of mobile sinks using contact and inter-contact times to reduce delays for opportunistic data forwarding.

Trajectory Prediction based Forwarding

Another approach for data forwarding in mobile networks is based on the trajectory prediction of mobile nodes. This prediction relies on historical mobility, where nodes store (and update) information on their identity and meeting history, and use it in a routing metric. Data can be forwarded along the predicted trajectory of the mobile node instead of routing data directly to the mobile node at its current position. [108, 115, 193] address trajectory prediction of mobile nodes but incur considerable overhead in computation, storage and initial training time. Furthermore, prediction schemes suffer from a high error rate.

2.4.2 Economic-aware Data Forwarding

In sensor networks, the routing task is distributed among the participating nodes. Current routing protocols assume that all the nodes are cooperative and each individual node is ready to forward packets for others. However, the nodes (e.g. a phone) in an WSNs with Mobile Phones (WSN-MP) may belong to different users; therefore they may abstain from cooperation in order to save their own resources (such as battery power which is scarce). Furthermore, mobile device users in WSN-MP are reluctant to act as packet relays due to privacy concerns, battery power consumption and potential costs of data communication (3G/4G cost). Such non-cooperative behaviour by selfish nodes results in the sharp degradation of network performance in opportunistic systems. Incentive schemes are necessary to promote cooperation among selfish nodes which can stimulate cooperation, check misbehaviours, and punish selfish nodes.

Recent work on the incentive schemes for WSNs fall into three categories: *barter-based approaches*, *reputation-based*, and *credit-based* categories.

Barter-based Incentive Schemes

Barter-based schemes or Tit-For-Tat (TFT) strategies are based on the realization of mutual benefits, where two encountering nodes exchange the same amount of messages.

Buttyn et al. [32] proposed a mechanism based on the principle of a barter. It allows a node to download a limited volume of messages from another node if it can provide the same volume of messages in return. Two neighbour nodes exchange the list of the messages in their possession and each node decides to trade messages based on its interest. Pair-wise Tit-for-Tat (TFT) is proposed in [168] to maximize the throughput of nodes. However, the requirement for exchanging the same amount of messages is can degrade the routing performance dramatically when there is significant difference between volume of messages of two nodes involved in trading. In such scenarios, the number of messages exchanged is much lower than messages available in the network, which results in low network throughput.

Furthermore, a reliable third party is important to monitor the behaviour of nodes in barter-based approaches. Third party nodes inform the other nodes in the network of their selfish neighbours. However, due to large covered area, availability of such reliable third party nodes at all time is not feasible for opportunistic WSN-MP.

Reputation-based Incentive Schemes

In Reputation-based schemes, nodes collectively detect misbehaving nodes and exclude them from the network by spreading their bad reputation. Each node is assigned a reputation score that reflects its degree of cooperation. By forwarding packets for others, nodes can earn their good reputation scores. High scores help nodes to achieve priorities to deliver their packets across the network.The misbehaviour of a node results in a decrease in its reputation score; and the node is isolated from the network if its reputation score falls below a threshold [120].

Many different approaches for reputation-based schemes are proposed by authors. In [192] , each intermediate node receives a receipt as a proof of cooperation after forwarding a message to another node. The behaviour of intermediary nodes is communicated by the receiver to the network through flooding. In RADON [120], forwarding ability of a node is assessed by integrating the reputation of forwarding data with the possibility of meeting a destination. Similarly, in IRONMAN [27], each node keeps the record of intermediate nodes and forwarding records for each message to detect cooperative and selfish nodes.

Reputation-based incentive schemes work well in WSNs with static and few mobile nodes, where each node has to manage reputation of few neighbours. However, Reputation-based incentive schemes are not easy to implement in infrastructure-less and intermittent connection scenarios due to frequent partitioning and lack of an end-to-end path. It is not feasible for a node to manage reputation of all possible neighbours in large scale WSN-MP.

Credit-based Incentive Schemes

Credit-based incentive schemes introduce some form of credits or virtual currency to discourage selfish behaviour among nodes in multi-user systems. The credit-based scheme was first introduced in [33] where a node earns credits, by forwarding packets for others. These credits can then be used to obtain forwarding service from other nodes in the system. The source node of a packet pays credits to the intermediate nodes which participate in relaying the packet to the destination.

The concept of credit is the motivation behind several forwarding schemes; that are proposed to stimulate cooperation among nodes for packet forwarding. In [33], source node awards credits to intermediate forwarding nodes as an incentive for packet forwarding. However, in mobile opportunistic networks, it is difficult to estimate the number of intermediate nodes that would participate in relaying packet to the destination. Therefore, setting an initial credit for the packet is a challenging problem in these networks. [210] is another credit-based scheme addressed node selfishness through a centralized credit distribution by the server. Close to our work is the message trade model [149], where the receiver pays credits to the sender in exchange of a message in each intermediate transmission. The destination node finally pays for message forwarding when it receives the message.

Regard-less of the performance of these schemes, none of them are designed for social-aware mobile networks and do not provide any performance guarantees. In our work, we have developed a distributed credit-based incentive scheme base on free market with performance guarantees and also exploit social structure and mobility patterns.

2.4.3 Economic behaviour of users in WSN-MP

In a WSN-MP, mobile phones will belong to individuals with different personal preferences. Mobile phone users may not be willing to fulfil a WSN-MP task, due to privacy concerns and the potential costs that would be incurred; impacting battery usage and 3/4G budgets. Therefore, taking account of the social and economic behaviours of phone users, though frequently ignored, is central to the success of WSN-MP.

Social Selfishness

Social Selfishness defines the behaviour of users to cooperate only with the selected trusted users. These trusted users can be friends or peers in a social network. Users may not be willing to cooperate to the strangers due to their privacy concerns. Routing schemes [56, 122] consider the concept of *Social Selfishness*, which describes the willingness of an individual to provide better service to those with strong social ties than those with weaker social ties.

Rational Selfishness

Rational Selfishness considered in our work means that each phone owner is willing to cooperate with other users regardless of social ties as long as it can earn profit as a result. Here phone users relay sensor data as long as he or she can get benefits, which is different form the concept of social selfishness.

Individual Rationality

To incentivise the users to participate in the network, the property of *Individual Rationality* must hold to guarantee profits for each participating user. *Individual Rationality* ensures that each phone user should obtain a non-negative net profit as a reward of its participation in the network.

Server Profitability

Server Profitability defines the feasibility of the deployment of WSN-MP. It ensures that the server should not incur a deficit, which means server always earns a non-negative server profit.

Incentive Compatibility

Due to *Rational Selfishness*, phone user attempt to increase their own profits misreporting their current state. These actions such as wrong reporting of their costs (i.e., 3G cost) or hiding their resources (i.e., no Wifi radio) in attempt to earn more profits are considered as *cheating*. Users may adopt cheating behaviour to avoid relaying data from others or increasing its own data rate, that may result in higher individual profit but it leads to inefficient performance of the network.

Incentive Compatibility ensures that adopting the action suggested by the proposed algorithm should be the best strategy for each phone user, regardless the action of other users. Therefore, a user cannot earn more profit through *cheating* actions, which ensures the optimal performance of the network.

Chapter 3

Mobility-aware Backpressure Collection in WSN with Mobile Sinks (WSN-MS)

In Wireless Sensor Network community, there is currently a movement away from dense and reliable sensing, toward the implementation of larger numbers of low-powered, low-cost, reduced precision sensing technologies. The topologies of such networks are variable and pertain to their application or environment. Some are able to connect to the Internet but this may be prohibitive in some instances due to high communication costs (e.g. 3G cellular costs) or poor connectivity. The mobile wireless devices carried by vehicles or individuals in WSN with Mobile Phones (WSN-MP) provide an attractive alternative and could be used as mobile sinks to collect sensor data in an opportunistic way. We refer such sensor networks to as Wireless Sensor Networks with Mobile Sinks (WSN-MSs) in this chapter.

Due to the requirement of mobility prediction and the lack of focus on delay and throughput performance, state-of-the-art mechanisms for WSN-MS perform poorly in practice. Many state-of-the-art approches use Expected Transmission Count (ETX) as a measure to find high-throughput routes in multi-hop wireless networks but ETX cannot be directly applied to WSN-MSs due to the mobility of the nodes. In this chapter, we propose a novel routing metric, Contact-Aware ETX (CA-ETX), to estimate the packet transmission delays that result from both link unreliability and intermediate connectivity. Using CA-ETX, we develop a throughput-optimal scheme Opportunistic Backpressure Collection (OBC). Both CA-ETX and OBC are lightweight, easy to implement, and require no mobility prediction. Through testbed experiments and extensive simulations, we show that the proposed schemes significantly outperform state-of-the-art approaches. We also show that existing ETX-based routing protocols such as CTP [72] and IETF RPL [11] can be applied to WSN-MSs with minimal modification using CA-ETX.

N^s	The set of all sensor nodes.
N^m	The set of all mobile sinks.
N	The set of all nodes.
L	The set of all wireless links between each pair of nodes in N.
$PRR_{x,y}(t)$	Packet Reception Probability (PRR) of wireless link $(x, y) \in L$ at slot t.
$ETX_{x,y}(t)$	Expected Transmission Count (ETX) of wireless link $(x, y) \in L$ at slot t.
$c_{x,y}(t)$	Channel capacity of link $(x, y) \in L$ at slot t.
$\boldsymbol{c}(t)$	Channel capacity vector for all wireless links at slot t.
$G(N, L, \boldsymbol{c}(t))$	The time-varying weighted graph of the WSN-MSs.
$\boldsymbol{S}$	The discrete state space of all possible channel capacities.
$\pi_{\boldsymbol{c}}$	The stationary distribution probability for channel capacity $\boldsymbol{c}$.
L^s	The set of all wireless links between each pair of sensor nodes.
$G^s(N^s, L^s)$	The graph of the subnetwork consisting of all sensor nodes.
N^o	The set of all sensor node and the virtual sink VS, i.e. $N^o = N^s \cup \{VS\}$.
L^o	The set of all opportunistic contact links, i.e. $L^o = L^s \cup \{(x, VS), x \in N^s\}$.
$G^o(N^o, L^o)$	The opportunistic contact graph.
N^o_x	The opportunistic contact neighbour of sensor node x.
$CA - ETX_x$	The Contact Aware ETX (CA-ETX) value of node $x \in N^o$.
$CA - ETX_{x,y})$	CA-ETX value over opportunistic contact link $(x, y) \in L^o$.
$OSFP(x)$	The shortest path from a sensor node x to the virtual sink VS over $G^o(N^o, L^o)$.
$OP(x)$	The opportunistic parent of sensor node x, i.e. the next-hop node in $OSFP(x)$.
$\mu_{x,y}, \sigma^s_{x,y}$	the mean and variance of service time over link $(x, y) \in L^o$ respectively.
$r_x(t)$	The sensing rate of sensor node $x \in N^s$ at slot t.
$\boldsymbol{r}$	The $\|N^s\|$-dimensional vector of all sensing rates.
$N_x(t)$	The set of node x's instantaneous neighbours at slot t.
$f_{x,y}(t)$	The amount of data transmitted over wireless link $(x, y) \in L$ at slot t.
$Q_x(t)$	The queue backlog of node $x \in N$ at slot t.
$f^{in}_x(t)$	The amount of total incoming data of node x at slot t.
$f^{out}_x(t)$	The amount of total outgoing data of node x at slot t.
$\boldsymbol{\mu}(\boldsymbol{c})$	A contention-free link rate vector for channel state $\boldsymbol{c}$.
$\Gamma(\boldsymbol{c})$	The link rate region for channel state $\boldsymbol{c}$.
φ_x	The gateway quality of a sensor node x.
$w_{x,y}(t)$	The routing weight of wireless link (x, y) at slot t.
$\mathrm{r}^{\max}$, $\mathrm{c}^{\max}$	The finite upper bounds of sensing rate and channel capacity respectively.

Table 3.1: Summary of symbols used in Chapter 3.

WSN-MSs	Wireless Sensor Network with Mobile Sinks
ETX	Expected Transmission Count
OBC	Opportunistic Backpressure Collection
CTP	Collection Tree Protocol [72]
RPL	The IP routing protocol designed for low power and lossy networks [11].
CA-ETX	Contact Aware ETX
DTN	Delay Tolerant Networks
PRR	Packet Reception Ratio
GQ	Gateway Quality
LQF	Longest Queue First
CSMA	Carrier Sense Multiple Access
E2E	End-to-End
BCP	Backpressure Collection Protocol [137]
BP	Backpressure
MG-IP	A routing protocol based on Mobility Graph and Information Potentials [108]

Table 3.2: Summary of abbreviations used in Chapter 3.

3.1 Introduction

Future smart sustainable city environments are predicted to have a huge amount of sensing devices deployed to monitor environmental conditions such as noise level, air pollution, and water pipe leakages [148, 151, 195]. Many urban sensing applications such as traffic monitoring and urban noise data gathering are delay-tolerant, which do not require to deliver sensed data in real-time [115, 151, 193]. In these cases, traditional Wireless Sensor Networks (WSNs) with static sinks may not be a feasible solution for large-scale urban sensing application due to the large sensing coverage requirement and infrastructure costs. Alternatively, wireless devices carried by vehicles or individuals (e.g. smart phones) can act as mobile sinks to collect urban sensor data in an opportunistic way, through short-range wireless communication radios such as Bluetooth, WiFi direct, Zigbee and LTE-direct [6, 36, 119, 133]. With increasing short-range communication capabilities of mobile devices and their huge population, WSNs with mobile sinks (WSN-MSs) have been becoming a more and more realistic and cost-effective solution to delay-tolerant urban sensing applications [108,115,123,193]. Besides urban sensing applications, WSN-MSs can be used in other delay-tolerant sensing applications such as habitat and forest monitoring [4].

In WSN-MSs, a sensor node can either send its data directly to a sink as it passes by, or it can send data via multi-hop routes to other sensor nodes who currently have contact with, or who will be likely in contact with a sink in the future. Therefore, how to choose the best routes to efficiently forward sensor data is a key issue for data collection in WSN-MSs.

Although this topic has received a reasonable amount of research to date, most of the work has limitations that affect their adoption in practice. Some approaches [108, 115, 193] require the prediction of trajectories of the sinks, which incurs considerable overheads and suffers from prediction errors or may not even be possible in large-scale practical scenarios. While other schemes, such as [123], focus on adaptively and smoothly updating routing tree structures as a sink moves in the sensing area. These schemes suffer heavy communications overheads in WSN-MSs with large numbers of fast moving sinks; or where there are intermittently-connected WSN-MSs where the sensor nodes are disconnected from any routes to the sink for reasonable periods of time (e.g. off-peak time in urban roads in [195]). Furthermore, due to the opportunistic availability of mobile sinks and heavy data traffic potentially produced by ubiquitous sensors, throughput performance is an important issue for data collection schemes. However, this has received very little attention in current WSN-MS research.

To overcome the limitations of current work, this chapter aims to develop high-throughput and low-delay opportunistic data collection approaches for practical WSNs-MSs with arbitrary numbers of mobile sinks, arbitrary sink movement speeds, taking account of routes being connected or intermittently-connected. The contributions of this chapter are summarized as follows:

1. Based on queuing analysis theory, we propose a novel routing metric Contact-Aware ETX (CA-ETX), to effectively estimate the packet transmission delay over opportunistic wireless links between sensors and sinks, caused by both wireless link unreliability (i.e. data packet retransmissions) and intermittent connectivity. Beside its efficiency, a major advantage of CA-ETX is that it can simultaneously work with ETX (the most popular link quality metric used by various WSN routing protocols such as the defacto TinyOS routing standard CTP [72] and the IETF IPv6 routing protocol RPL [11]). By implementing CA-ETX in two well-known WSN operating systems TinyOS [1] and Contiki [2], we

demonstrate that current ETX-based routing standards CTP [72] and RPL [11] can be extended easily to support opportunistic data collection WSN-MSs by using CA-ETX.

2. we propose a throughput-optimal approach, Opportunistic Backpressure Collection (OBC), a joint dynamic multi-path routing and scheduling protocol for WSN-MSs by integrating CA-ETX. In contrast to current data collection schemes in WSN-MSs, OBC is lightweight, easy to implement, requires no mobility prediction, and can support a large number of fast moving sinks. To our knowledge, OBC is the first scheme that combines the backpressure approach [137, 181] and mobility awareness for WSN-MSs.

3. We construct real-world experiments and extensive simulations to evaluate the performance of CA-ETX and OBC. The results show that the delay performance of both CTP and RPL can be significantly improved by simply adopting CA-ETX over opportunistic sensor-sink links. In addition, evaluation results demonstrate that OBC can achieve significant performance improvements in terms of end-to-end delay, storage overheads, energy consumption, and scalability compared with state-of-the-art approaches.

The remainder of this chapter is organised as follows: The next section discusses related work. We present our system model in Section 3.3. Section 3.4 proposes the CA-ETX metric. Section 3.5 provides detailed descriptions of the OBC algorithm. Simulation and testbed experiment results discussed in Section 3.6. Finally, we conclude the chapter in Section 3.7. All proofs of theorems in this chapter and related lemmas are placed in Appendix A.

3.2 Related Work

3.2.1 Wireless Routing Metrics

In wireless networks, routing protocols use various link metrics to select the best end-to-end forwarding path from sources to destinations. In static wireless networks such as WSNs, routing metrics like the expected transmission count (ETX) [49] estimate the packet transmission delays caused by link unreliability (i.e. retransmissions). ETX has been used in many routing standards such as CTP [72] and RPL [11], and its efficiency has been validated in numerous real-world experiments. However, it is not a surprise that metrics like ETX cannot be directly applied to WSN-MSs. This is because that data transmission delays not only depend on link unreliability, but also on the intermittent connectivity between the static sensor nodes and mobile sinks. In mobile networks such as Delay Tolerant Networks (DTNs), routing metrics such as inter-contact time [39],or contact probabilities [68] are widely used. However, all these metrics ignore the quality of the temporary wireless link that connects nodes during their moment of contact. In contrast to existing routing metrics in both mobile and static wireless networks, CA-ETX is specifically designed for the WSN-MSs and can efficiently estimate packet waiting times in data buffers, which is the major delay of per-hop packet transmission.

3.2.2 Backpressure Algorithms

Backpressure algorithms [69, 111, 137] are well-known for their optimal throughput but poor delay performance. Several techniques have been developed recently to improve the delay performance of backpressure algorithms [82, 83, 197, 206]. However, this is primarily theoretical work, and focuses on networks with static topologies only rather than time-varying topologies; therefore they cannot be applied to practical WSN-MSs. There are a few backpressure schemes applied to mobile multi-hop networks [57, 158, 159]. Recent interesting work, BWAR [15], develops an adaptive redundancy technique for backpressure routing in DTNs. However, the idea of BWAR cannot be applied to WSN-MSs, in which packet replication is not used due to the limited bandwidth resource.

[158, 159] study backpressure flow control and routing in disconnected static wireless networks with mobile relays and fixed gateways. In contrast, our work focuses on WSN-MSs where each sensor node could dynamically serve as a gateway at every opportunistic contact with a mobile sink.

3.3 System Model

We consider a WSN-MS consisting of static sensor nodes and mobile sinks to collect sensor data using short-range communication radios (e.g. Bluetooth, Zigbee, or WiFi direct). If a sensor node is in contact with a mobile sink, it forwards its sensor data to a mobile sink directly; otherwise, it stores the data and waits for its next connection to a mobile sink or forwards its sensor data through other sensor nodes to any mobile sink in a multi-hop fashion.

Let the sets of all sensor nodes and mobile sinks be N^s and N^m respectively, and denote $N = N^s \cup N^m$. The network operates in discrete time slots (e.g. seconds) $t \in \{0,\ 1,\ ...\}$. We define the packet reception probability (PRR) over a wireless link (x, y), $0 \leq PRR_{x,y}(t) \leq 1$, as the probability of successfully transmitting a data packet, with acknowledgement from node x to y, at slot t. $PRR_{x,y}(t)$ is assumed to be constant within the duration of a slot, but can vary from slot to slot and across different wireless links, due to the time-varying wireless channel quality and intermittent connectivity between static sensor nodes and mobile sinks. According to its definition [49], the ETX value over a link (x, y) at slot t, $ETX_{x,y}(t)$, can be computed as

$$ETX_{x,y}(t) = \frac{1}{PRR_{x,y}(t)} \geq 1 \tag{3.1}$$

We define

$$c_{x,y}(t) = \mathrm{c}^{\max} PRR_{x,y}(t) \geq 0 \tag{3.2}$$

as the logical link-layer capacity of a wireless link from node $x \in N$ to node $y \in N$ at time t, i.e. the maximum (integer) number of sensor data packets (or bits) with acknowledgements that can be successfully transmitted from x to y during slot t, where $\mathrm{c}^{\max}$ is the maximal possible $c_{x,y}(t), \forall(t)$, which is bounded by the data rate of the wireless radio. For instance, experimental studies show that a commonly used IEEE 802.15.4 transceiver, CC2420 (e.g. [8]), can achieve a data rate of approximate 160 40-bytes packets per second [173] in practice. If $c_{x,y}(t) > 0$, we say nodes x and y are ***in contact*** at slot t; otherwise, they are not in contact at slot t (i.e. $ETX_{x,y}(t) = \infty$).

The states of WSN-MS at a given slot $t \geq 0$ can be represented as a directed, complete, and time-varying weighted graph $G(N, L, \boldsymbol{c}(t))$, where L represents the set of all possible wireless links between each pair of nodes in N, and the $|L|$-dimensional vector $\boldsymbol{c}(t)$ represents the vector of channel capacities over all wireless links at slot t.

It can be seen that $\boldsymbol{c}(t)$ can characterize the time-varying channel capacities of the WSN-MS caused by both slow fading between the static sensor nodes and fast fading between sensor nodes and the mobile sinks. Therefore, $\boldsymbol{c}(t)$ also implies the stochastic process of sink mobility. We assume that $\boldsymbol{c}(t)$ is an ergodic Markov chain that takes values on a finite (but which can be arbitrary large) discrete state space $\boldsymbol{S}$, and is has stationary distribution probability $\pi_{\boldsymbol{c}}$ for each channel capacity state $\boldsymbol{c}$. The Markov assumption is realistic and general for both mobility (e.g. [112]) and channel states (e.g. [140]). It is shown that many real mobility traces exhibit a high degree of spatial regularity [74,209]. In the context of WSN-MS, this means that mobile sinks appear in some specific locations with a higher probability than others, resulting in heterogeneous opportunities of sensor nodes to meet mobile sinks (e.g. sensor nodes in shopping centers should have more opportunities to meet mobile sinks than those in park).

In addition, the network consists of all statically-deployed sensor nodes which can be represented as a directed graph $G^s(N^s, L^s)$, where L^s represents all wireless links between sensor nodes. Topologically, the $G^s(N^s, L^s)$ could be either a connected graph, or disconnected graph consisting of several connected subgraphs.

3.4 Contact Aware ETX

3.4.1 Shortest Path Routing based on CA-ETX Gradient in Opportunistic Contact Graphs

This chapter considers anycast routing i.e. the destination of each sensor data packet can be any mobile sink. It is straightforward to extend our work to multi-commodity traffic models, by defining a virtual sink for each commodity. By using a virtual sink VS to represent all the mobile sinks in N^m, we define the opportunistic contact graph as $G^o(N^o, L^o)$, where $N^o = N^s \cup \{VS\}$ represents the set of all sensor nodes and the virtual sink, and L^o represents the set of all sensor-to-sensor and sensor-VS links, i.e. $L^o = L^s \cup \{(x, VS) : x \in N^s\}$. Fig. 3.1 illustrates an example of opportunistic contact graph. For each opportunistic link (x, y) in L^o, we define a

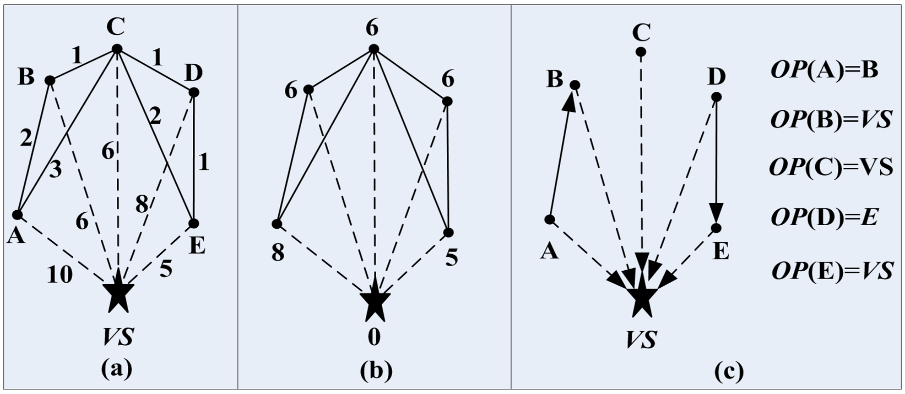


Figure 3.1: An example for CA-ETX gradient and opportunistic shortest path routing in WSN-MS. (a) An example opportunistic contact graph with link CA-ETX values, where solid and dashed lines represent sensor-to-sensor links and sensor-VS links respectively; (b) node CA-ETX values; and (c) opportunistic shortest path routing by using CA-ETX.

metric $CA\text{-}ETX_{x,y}$, to estimate the packet transmission delay over this link. The computation of link CA-ETX values will be discussed in detail in Subsection 3.4.2. Based on the *link* CA-ETX values, each sensor node x can compute its *node* CA-ETX value, $CA\text{-}ETX_x$ in a fully distributed way:

$$CA\text{-}ETX_x = \min_{y \in N_x^o}(CA\text{-}ETX_x, CA\text{-}ETX_y + CA\text{-}ETX_{x,y}) \tag{3.3}$$

where

$$N_x^o = \{VS\} \cup \{y : (x, y) \in L^s\}$$

represents the opportunistic contact neighbor set of x.

Specifically, the CA-ETX value of the virtual sink $CA\text{-}ETX_{VS} = 0$. For instance, Fig.3.1 (a) shows an example of an opportunistic contact graph and the link CA-ETX value for each link, and Fig. 3.1 (b) shows corresponding node CA-ETX values computed by using (3.3). For each sensor node x, let $OSP(x, VS)$ be the shortest path from x to VS in $G^o(N^o, L^o)$, i.e. $OSP(x, VS)$ is the path with the minimal total link CA-ETX values from x to VS. It is easy

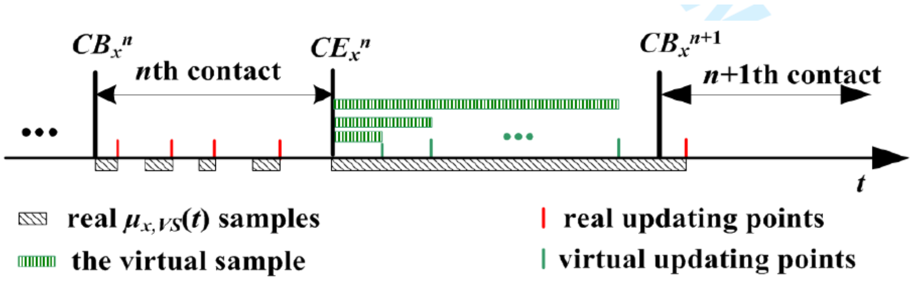


Figure 3.2: An example to show how to compute and update CA-ETX value for sensor-VS links.

to verify that

$$CA\text{-}ETX_x = \sum_{(i,j)\in OSP(x,VS)} CA\text{-}ETX_{i,j}$$

i.e. the node CA-ETX value of each sensor node x represents the total link CA-ETX values of all links in $OSP(x, VS)$.

For each sensor node x, define its opportunistic parent

$$OP(x) = \arg\min_{y\in N_x^o}(CA\text{-}ETX_x + CA\text{-}ETX_{x,y})$$

For instance, Fig. 3.1 (c) shows the opportunistic parents of all sensor nodes based on their node CA-ETX values.

With the CA-ETX gradient, the opportunistic shortest path routing can be easily performed in a fully distributed way: If a sensor node x is in contact with any mobile sink $m \in N^m$, x, it forwards data directly to m. Otherwise, it forwards data to $OP(x)$, if $OP(x)$ is a sensor node; it waits for a sink, if $OP(x)$ is the virtual sink. For instance, sensor node A in Fig. 3.1 (c) will forward data to its opportunistic parent B, if it is not in contact with any mobile sink; otherwise, it will forward data to the mobile sink directly.

3.4.2 Link CA-ETX Calculation

We consider each link (x, y) in L^o as a queue with time-varying packet service times $\mu_{x,y}(t)$, $t \geq 0$, which is the service time duration required for a successful packet transmission over link (x, y) at slot t. If (x, y) is a sensor-sensor link, $\mu_{x,y}(t)$ can be easily computed as

$$\mu_{x,y}(t) = 1/c_{x,y}(t), \forall t \geq 0 \tag{3.4}$$

From (3.2) and (3.4), we can see the the classic time-average ETX value over link (x, y) is

$$ETX_{x,y} = \mathbb{E}[\mu_{x,y}(t)]\mathrm{c}^{\max} = \mu_{x,y}\mathrm{c}^{\max} \tag{3.5}$$

where $\mathbb{E}[\cdot]$ is the expectation operator and $\mu_{x,y}$ is the long-term mean of $\mu_{x,y}(t)$.

For a sensor-VS link (x, VS), we have

$$\mu_{x,VS}(t) = \begin{cases} 1/c_{x,m}(t), \ CB_x^n \leq t \leq CE_x^n \\ CB_x^{n+1} - t + 1/c_{x,m}(CB_x^{n+1}), \ \text{otherwise} \end{cases} \tag{3.6}$$

where CB_x^n and $CE_x^n, n \geq 1$ are the first and last time slots of the nth contact between x and any mobile sink m respectively, shown in Fig. 3.2. Therefore, packet service times for sensor-VS links depend on the dynamic contact durations, the inter-contact time (i.e. durations between each contact), as well as the link quality during each contact, which cannot be reflected in existing link metrics, such as ETX, contact probability, or inter-contact time.

Due to the complex dynamics of the system, both the arrival process and service times over each link $(x, y) \in L^o$ should follow general distributions rather than specific ones. Therefore, we model each link in L^o as a G/G/1 queue. From queueing theory, the average packet waiting time $wd_{x,y}$ in the link queue (x, y) can be approximately represented as:

$$wd_{x,y} = \frac{\sigma^a_{x,y} + \sigma^s_{x,y}}{2(\chi_{x,y} - \mu_{x,y})} \tag{3.7}$$

where $\mu_{x,y}$ and $\sigma^s_{x,y}$ are the standard mean and variance of service time over link (x, y) respectively, and $\chi_{x,y}$ and $\sigma^a_{x,y}$ are the mean and variance of packet arrival intervals respectively. From (3.7), we can see that $wd_{x,y}$ is an increasing function of both $\mu_{x,y}$ and $\sigma^s_{x,y}$. Therefore, we define the CA-ETX value for each link $(x, y) \in N^o$ as

$$CA\text{-}ETX_{x,y} = (\sigma^s_{x,y}/\widetilde{\sigma}^s_{x,y})\mathrm{c}^{\max}\mu_{x,y} \tag{3.8}$$

where $\widetilde{\sigma}^s_{x,y}$ is the variance of the service times during each contact between x and y, i.e. if (x, y) is a sensor-sensor link, $\widetilde{\sigma}^s_{x,y} = \sigma^s_{x,y}$ is the variance of all service times samples; if (x, y) is a sensor-VS link, $\widetilde{\sigma}^s_{x,y}$ is the variance of all service times samples during $CB^n_x \leq t \leq CE^n_x,\ n \geq 1$. In (3.8), we normalize $\sigma^s_{x,y}$ by $\widetilde{\sigma}^s_{x,y}$, in order to follow the concept of classic ETX, and to facilitate its use in current ETX-based routing protocols such as CTP and RPL with the minimal modification. As a result, we have:

$$CA\text{-}ETX_{x,y} = \begin{cases} \mathrm{c}^{\max}\mu_{x,y} = ETX_{x,y} & y \in N^o_x - \{VS\} \\ (\sigma^s_{x,y}/\widetilde{\sigma}^s_{x,y})\mathrm{c}^{\max}\mu_{x,y} & y = VS \end{cases} \tag{3.9}$$

From (3.9), we can see that CA-ETX for sensor-sensor links are identical to the classic ETX. Therefore, we can directly use ETX-based routing schemes in WSN-MSs, by simply using CA-ETX estimations for sensor-VS links.

3.4.3 Updating CA-ETX for Sensor-VS links

For a sensor-VS link (x, VS), service time samples can be easily estimated when x transmits each packet to any mobile sink at run time. For each new service time sample, the values of

$\mu_{x,y}$, $\sigma^s_{x,y}$, and $\widetilde{\sigma}^s_{x,y}$ can be updated based on following efficient online algorithm.

Consider a sequence of samples $X^{(1)}, X^{(2)}, ...,$. When the n^{th} sample $X^{(n)}$ is obtained, the mean $\mu^{(n)}$ and variance $\sigma^{(n)}$ of these n samples can be updated as

$$\mu^{(n)} = \mu^{(n-1)} + \frac{1}{n}(X^{(n)} - \mu^{(n-1)}) \tag{3.10}$$

$$\sigma^{(n)} = \sigma^{(n-1)} + (n-1)(X^{(n)} - \mu^{(n)})(\frac{X^{(n)} - \mu^{(n)}}{n}) \tag{3.11}$$

with $\mu^{(1)} = X_1$ and $\sigma^{(1)} = 0$ [40].

However, when x is not in contact with any sink, packet service time of a sensor-VS link (x, VS) could be very large (e.g. several minutes or hours), as shown in Fig. 3.2. As a result, simply updating $CA\text{-}ETX_{x,VS}$ after each service time sample (e.g. the red time point in Fig. 3.2) would result in the CA-ETX gradient being non-agile to the network dynamics. We solve this problem by using a virtual sample of service times shown in Fig. 3.2 before a large real service time sample is obtained. This virtual sample (also $CA\text{-}ETX_{x,VS}$) is updated at time points (e.g. green time points in Fig. 3.2) with a small interval (e.g. current mean service time $\mu_{x,VS}$). The virtual sample is abandoned when the real large service time sample is obtained.

3.4.4 Discussion

For a sensor node x, it is easy to see that the mean service time $\mu_{x,VS}$ depends on its contact probability with any mobile sink and link quality during each contact, while the variance $\sigma^s_{x,y}$ mainly depends on the inter-contact time between x and any mobile sink. It is also not difficult to verify that the former depends on a spatial distribution of mobile sinks and deployments of sensor nodes, while the latter is greatly affected by the movement speeds of the mobile sinks. Therefore, CA-ETX is very useful in practice, due to regular spatial behavior [74, 209] and heterogeneous movement speeds (e.g. walking people, bikes, and vehicles) of mobile sinks. For instance, by using CA-ETX, packets are relayed via sensor nodes close to a fast moving highway rather than via nodes close to a pedestrian path even though they both have similar traffic rates, resulting better delay performance.

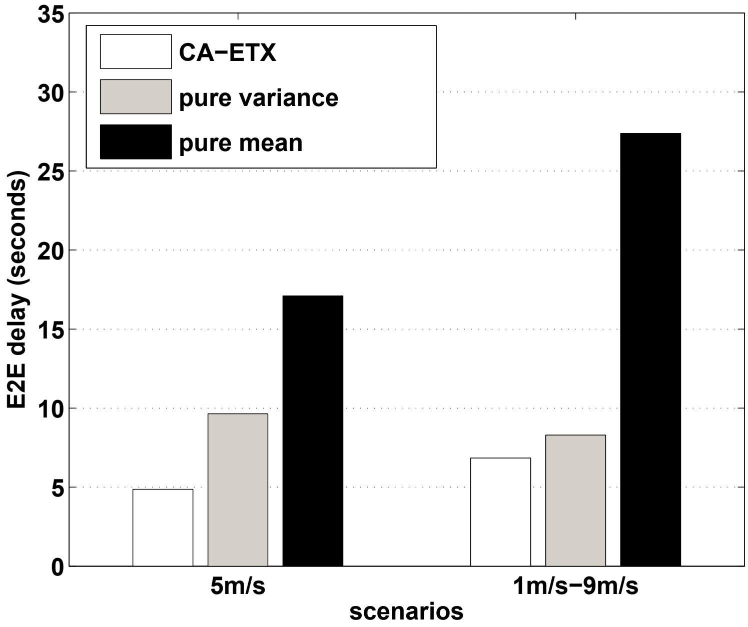


Figure 3.3: Delay performance of opportunistic shortest path routing with different metrics in a WSN-MS with homogeneous and heterogeneous sink moving speeds.

As an example, Fig.3.3 illustrates the average end-to-end delays of opportunistic shortest path routing that uses the following three metrics to measure the delay of all sensor-VS links (x, VS), $x \in N^s$: (1) $CA\text{-}ETX_{x,VS}$, (2) $\mu_{x,VS}$ (**pure mean**), and (3) $\sigma^s_{x,VS}$ (**pure variance**).

The WSN-MS used in the simulation consists of 400 sensor nodes and 50 mobile sinks in a 1000 m$\times$1000 m simulation area. In the first and second simulations, the sink movement speeds were fixed to each 5m/s and uniformly distributed between 1m/s and 9m/s respectively. Other simulation settings can be found in Subsection 3.6.2.

Due to the same mean of sink movement speeds (i.e. 5m/s), the average service time are similar are same for both simulations. However, the variances of packet service times $\sigma^s_{x,VS}$ are different, which results in significant delay performance difference shown in Fig.3.3. Since CA-ETX manages to provide a fine measurement of the packet waiting time in each sensor-VS link, it achieves the best delay performance in both simulations, demonstrated in Fig.3.3.

3.5 Opportunistic Backpressure Collection

By forwarding sensor data through the minimal-cost routes, shortest path routing has good delay performance especially in WSN-MSs with light-weight sensor data traffics. However, the poor throughput performance of shortest-path routing limits its practical applications in WSN-MSs with potential large volume of sensor data traffic. Therefore, this section develops a novel throughput-optimal algorithm, OBC, by integrating CA-ETX into the backpressure algorithm for WSN-MSs. Before presenting OBC, we first model the WSN-MS as a dynamic networked queuing system.

3.5.1 Queueing Dynamics

Every sensor node $x \in N^s$ collects data at a sensing rate of r_x . We assume the constant sensing rate for analytical brevity. However, it is straightforward to extend our results to general ergodic sensing rates. Let $0 \leq f_{x,y}(t) \leq c_{x,y}(t)$ represent the actual amount of sensor data transmitted over the wireless link (x, y) at slot t. Define $N_x(t) \subset N$ as the set of nodes that are in contact with node x at slot t, i.e the set of node x' all current neighbors. Each sensor node maintains a queue (i.e. data buffer) to store the sensor data packets received from other sensor nodes and produced by itself. Let $Q_x(t) \geq 0$ be the queue backlog (or queue length) of $x \in N^s$ at slot $t \geq 0$. From slot t to $t + 1$, queue backlog updates as follows:

$$Q_x(t+1) = |Q_x(t) - f_x^{out}(t)|_+ + r_x + f_x^{in}(t), \forall x \in N^s \tag{3.12}$$

where $f_x^{in}(t)$ and $f_x^{out}(t)$ are the amount of total incoming and outgoing data of node x at slot t respectively, i.e.

$$f_x^{in}(t) = \sum_{y \in N_x(t)} f_{y,x}(t), \quad f_x^{out}(t) = \sum_{y \in N_x(t)} f_{x,y}(t)$$

and for any real number a, the operator $|a|_+ = a$ if $a > 0$; $|a|_+ = 0$ otherwise. It is worth noting that the queue backlog $Q_m(t) = 0$, for all $m \in N^s$, $t \geq 0$.

3.5.2 Link Rate Region

We say a set of wireless links in L are contention-free if they can be active (i.e.transmitting) simultaneously, which depends on the interference relations between them. For a channel state $\boldsymbol{c}$, we define a L-dimensional contention-free link rate vector $\boldsymbol{\mu}(\boldsymbol{c})$, where each entry l is the capacity c_l of the link l if link l is scheduled to transmit; otherwise, entry l is zero. The wireless links associated with the non-zero entries in $\boldsymbol{\mu}(\boldsymbol{c})$ are contention free. We further define the link rate region $\Gamma(\boldsymbol{c}(t))$ associated with channel state $\boldsymbol{c}(t)$ as the convex hull of all possible contention-free link rate vectors.

3.5.3 Network Capacity Region

We define a network capacity region Λ by saying that any given $\boldsymbol{r} \in \Lambda$ if there exists a joint routing and scheduling algorithm that controls $f_{x,y}(t),\ (x,y) \in L$ at every slot $t \geq 0$ such that

$$\overline{f_x^{out}} - r_x - \overline{f_x^{in}} = 0,\ \forall x \in N^s \tag{3.13}$$

$$\boldsymbol{f}(t) \in \Gamma(\boldsymbol{c}(t)), \forall t \tag{3.14}$$

where $\overline{f_x^{out}}$ and $\overline{f_x^{in}}$ are the long-term averages of $f_x^{in}(t)$ and $f_x^{out}(t)$ respectively, and $\boldsymbol{f}(t)$ is the vector of all $f_{x,y}(t),\ (x,y) \in L$. Constraints (3.13) and (3.14) state the flow conservation law and the link rate region constraint respectively.

3.5.4 OBC Algorithm

At each slot $t \geq 0$, the OBC algorithm operates as follows:

1. **Weight Calculation**. Each sensor node $x \in N^s$ computes the weight $w_{x,y}(t)$ for each of its current neighbors $y \in N_x(t)$,

$$w_{x,y}(t) = (Q_x(t)/\varphi_x - Q_y(t)/\varphi_y)c_{x,y}(t) \tag{3.15}$$

where

$$\varphi_x = \frac{1}{CA\text{-}ETX_{x,VS}}$$

is called the Gateway Quality (GQ) of sensor node x. To guarantee the stability of OBC in theory, we set deterministic lower and upper bounds for all sensor nodes $x \in N^s$, i.e. $0 < \varphi^{\min} \le \varphi_x \le \varphi^{\max} < \infty$. In addition, we set φ_m for each mobile sink m as any fixed non-zero value between $\varphi^{\min}$ and $\varphi^{\max}$, which has no impact on OBC algorithm.

2. **Scheduling.** The set of scheduled links $F^*(t)$ is chosen as

$$F^*(t) = \arg \max_{F(t)\in\Gamma(\boldsymbol{c}(t))} \sum_{(x,y)\in F(t)} w_{x,y}(t) \tag{3.16}$$

It is clear that $F^*(t)$ is the set of contention-free links with the maximum aggregated weights at slot t.

3. **3. Routing and Forwarding.** Based on $F^*(t)$, each sensor node $x \in N^s$ transmits a sensor data packet to the next one-hop node by setting $f_{x,y}(t), y \in N_x(t)$ as follows:

$$f_{x,y}(t) = \begin{cases} c_{x,y}(t) & (x,y) \in F^*(t) \wedge w_{x,y}(t) > 0 \\ 0 & \text{otherwise} \end{cases}$$

Hence, each node x chooses the next hop node y such that link (x, y) is scheduled and $w_{x,y}(t) > 0$ (routing), then transmits $c_{x,y}(t)$ amount of data packets to y (forwarding).

4. **Queue Update.** Each sensor node x updates its queue backlog using (6.3).

According to the OBC algorithm, it can be seen that sensor nodes with higher GQs have a higher opportunity to receive more packets than those with lower GQs, which naturally combines mobility awareness (i.e. CA-ETX) and queue gradient of backpressure algorithm. Since the packet waiting time over sensor-VS link can be precisely estimated by CA-ETX , the packet transmission delay can be significantly reduced. Theorem to demonstrate throughput optamality OBC can be found in Appendix A.

It is worth noting that OBC *does not* require any future knowledge of the network, and makes routing and scheduling decisions based on only the network information at the current slot.

3.5.5 Distributed Implementation and Practical Issues

Now we discuss distributed implementation of OBC in practical WSN-MSs.

1. **Distributed Scheduling.** The optimal solution to the scheduling problem (3.16) is centralized and NP-hard for practical wireless networks with general interference relations (e.g. [166]), which is therefore intractable in practical WSN-MSs. To solve this problem, we implemented a fully distributed suboptimal scheduler, the greedy **Longest Queue First (LQF)**, which can achieve a near-optimal performance in practical wireless networks [29, 42, 97]. The details of distributed LQF implementation can be found in [42]. The algorithm works as follows.

   Each node i carries out the following steps:

   1) Calculate weight w for each link to neighbour nodes.

   2) Find a neighbour node j with maximum link weight

   – If it received a matching request from j, then link i, j is a matched link. Node i sends a matched reply to j and a drop message to all other neighbours. – Otherwise, node i sends a matching request to node j.

   3) Upon receiving a matching request from neighbour j:

   – If j is neighbour node of node i with maximum link weight, then link i, j is a matched link. Node i sends a matched reply to node j and a drop message to all other neighbours. – Otherwise, node i just stores the received message.

   4) Upon receiving a matched reply from neighbour$j$,link i, j is selected as a matched link, and send a drop message to all other neighbours.

   5) Upon receiving a drop message from neighbour j, node i excludes j from its neighbours set.

6) If node i is in a matched link or has no free neighbours, no further action is taken. Otherwise, it will repeat steps 2 – 5.

7) Only Matched links are allowed to transmit.

2. **WSN-MSs with CSMA radio.** Consider the discrete time slot modeling in our system, the *distributed OBC with LQF scheduler* can be directly used in synchronized TDMA networks. However, Since most current wireless devices adopt CSMA-based radios, our evaluation used an efficient technique proposed in [13,172] to implement OBC with LQF: If a link (x, y) is scheduled to transmit (decision made by OBC with LQF), node x will reduce its CSMA backoff window size to aggressively access the channel; otherwise, x accesses the channel with normal backoff window size..

3. **Mobile Sink Discovery and neighbour management.** In our evaluation, each mobile sink declares its presence to its current nearby sensor nodes by periodically broadcasting one-hop beacons. We set the beacon broadcasting interval of mobile sink $\mathrm{T}_{\mathrm{sink}}$ as 1 second and 250 milliseconds for testbed experiments and simulations respectively. Each sensor node also broadcasts a beacon per second to inform its neighbours its local information (e.g. queue backlog). In addition, each node updates its neighbour table every $\mathrm{T}_{\mathrm{neighbor}} =$ 50 milliseconds. To reduce the control packets, we implemented the overhearing (or snooping) mechanism (e.g. [137]). Here, each data packet or acknowledgement message includes the sender's local information in a packet header filed. A node (sensor node or mobile sink) does not need to broadcast a beacon in a broadcasting interval, if it has already send a data packet or replied an acknowledgement in the same broadcasting interval. It can be seen that our implementation achieves both precise and lightweight sink discovery and neighbour table updating.

3.6 Evaluation

In this section, we constructed extensive real-world experiments and simulations to evaluate the proposed CA-ETX metric and OBC algorithm. Our evaluation are based on two popular

Table 3.3: Implementation details of CA-ETX and OBC.

Protocol	CA-ETX-CTP	CA-ETX-RPL	OBC
Platform	TinyOS	Contiki	TinyOS/Castalia
Code Changed	35 lines	68 lines	-
RAM (kB)	3.1	9.8	2.7 (TinyOS)
ROM (kB)	30.5	42.5	26.9 (TinyOS)

Table 3.4: Summary of Evaluation Parameter Settings

Evaluation	**CA-ETX Evaluation**		**OBC Evaluation**			
Method	**Testbed**	**Simulation**	**Testbed**	**Simulation**		
Protocols	CTP(CA-ETX *vs* ETX)	RPL(CA-ETX *vs* ETX)	OBC&BCP	OBC	BP	MG-IP
Platform	TinyOS/MicaZ	Contiki/Cooja	TinyOS/MicaZ	Castalia	Castalia	Castalia
MAC Layer	IEEE 802.15.4	IEEE 802.15.4	IEEE 802.15.4	CSMA+LQF	CSMA+LQF	CSMA
TX power	-25 dbm	0 dbm	-25 dbm	0 dbm	0 dbm	0 dbm
Mobility	Real Mobility	HHW Model	Real Mobility	HHW Model	HHW Model	HHW Model
Prediction	no	no	no	no	no	more than 95% accuracy
Packet Size	34 Bytes	40 Bytes	34 Bytes	34 Bytes		
Retransmission	10 times	10 times	10 times	10 times		
Scale	20 sensors+2 sinks	200 sensors+10 sinks	20 sensors+2 sinks	up to 1000 sensors+20 sinks		
$\mathbf{T_{sink}}$	1 second	250 milliseconds	1 second	250 milliseconds		
$\mathbf{T_{neighbor}}$	50 milliseconds	50 milliseconds	50 milliseconds	50 milliseconds		
Buffer size	20 packets	40 packets	20 packets	300 packets		

WSN and IoT operating systems, TinyOS [1] and Contiki OS [2]; and a realistic WSN simulator Castalia [3].

We applied the CA-ETX metric to the defacto TinyOS routing standard CTP [72] and the IETF IPv6 routing protocol RPL (in Contiki OS) [11], and implemented the OBC algorithm in both TinyOS and the Castalia simulator, as shown in Table 3.3. In addition, we adopted the following implementation approaches in our evaluations.

Throughout the evaluation, we collected the follow four metrics to measure the performance of all protocols.

- *End-to-End Delay.* The time taken for every packet from source to destination.
- *Queue Backlog.* The number of data packets in each node's data buffer, which indicates the storage efficiency.
- *Communication Overhead.* The number of transmitted and received packets (including all data and control packets) per node per second. This performance metric can measure the efficiency of routing algorithm in terms of energy and bandwidth consumptions.
- *Packet Loss Rate.* The percentage of lost data packets, indicating the reliability perfor-

Figure 3.4: Sensor Deployments for the ETX and CA-ETX (with CTP) experiments.

mance. Here, the remaining packets at the end of each evaluation was not considered as packet losses.

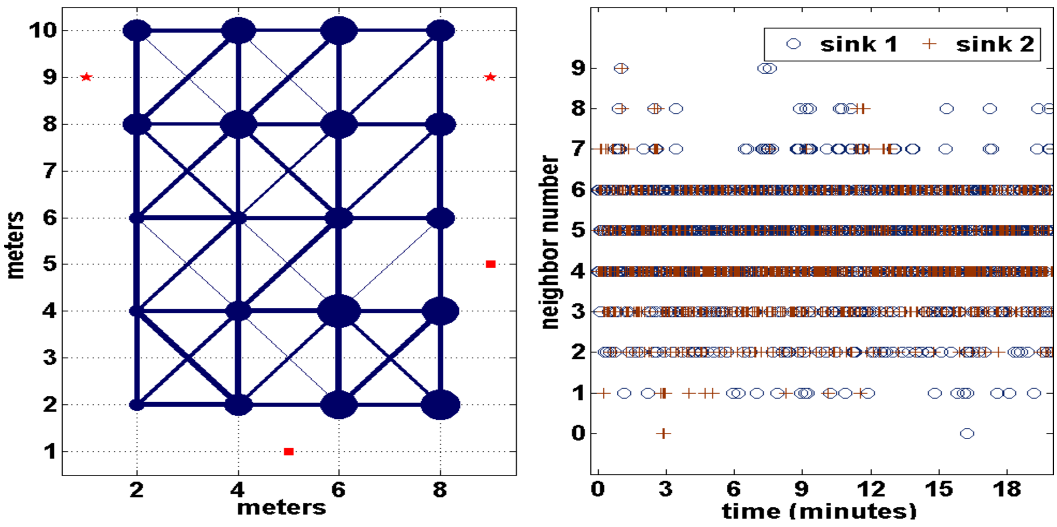
Table 3.4 summarizes the parameter settings of our evaluation, which will be discussed in detail in specific subsections below.

3.6.1 Evaluation of CA-ETX

In this subsection, we demonstrate how to use the CA-ETX metric to extend current WSN routing protocols to WSN-MSs. Specifically, we first applied CA-ETX to CTP [72], and evaluate the performance improvement through a testbed experiment using TinyOS based MicaZ motes. Then, we constructed a simulation using Cooja, the simulator of Contiki OS [2], to demonstrate the performance gain of RPL [11] by using CA-ETX.

CA-ETX with CTP

We constructed experiments using MicaZ motes to evaluate the practical performance of applying CA-ETX in CTP [72]. As shown in Table 3.3, only 35 lines of nesC code is used to



(a) connectivity visualization (b) sink mobility

Figure 3.5: Visualization of the ETX and CA-ETX (with CTP) testbed experiment settings: (a) Sensor deployment topology and mobility hot spots, (b) the sequences of neighbor numbers of the two sinks.

implement CA-ETX in CTP. We compared the performance of CTP with ETX and CTP with CA-ETX by using the following methods:

As shown in Fig.3.4 and Fig.3.5, two WSN-MSs were concurrently deployed for the ETX and CA-ETX experiments respectively in London Hyde Park. Each WSN-MSs consisted of 20 sensor nodes (i.e. the blue cycles in Fig. 3.5(a)), which were deployed in a grid topology with a cell size of two meters. The mobile sinks were held by two researchers respectively (each researcher carries two sinks for ETX and CA-ETX respectively). The experiment lasted for 20 minutes during which the two researchers (mobile sinks) roamed around the deployment area, simulating both high probabilities of visiting some hot points (i.e. red circles and stars in 3.5(a)) and other low probability locations. Therefore, the two WSN-MSs had the exactly same deployments and sink mobility. In order to avoid inter-interference between the two WSN-MSs, they were operated on two orthogonal channels of the CC2420 radio (channels 13 and 26) respectively. In addition, neither channel experienced interference by other external 2.4 GHz wireless transmissions during the experiment, as no WiFi signal was detectable around the deployment area.

Fig.3.5 illustrates above experiment scenarios, where the diameter of each blue cycle in Fig.3.5(a) is linearly proportional to the percentage of time that the corresponding sensor node was in

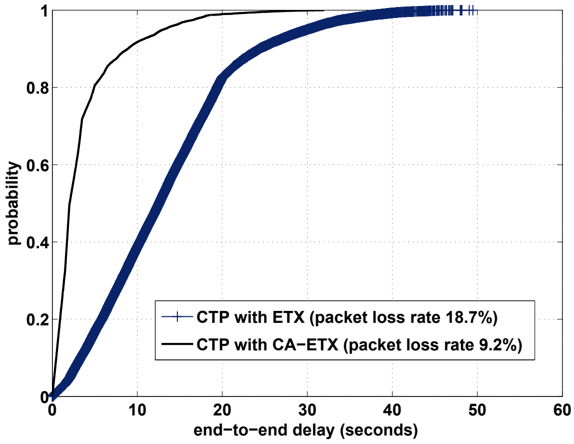


Figure 3.6: Packet loss rate and CDF of end-to-end delay: CTP with ETX and CA-ETX.

contact with a mobile sink; the width of each blue line is linearly proportional to the percentage of time that corresponding pair of sensors were connected as neighbors; and Fig.3.5(b) illustrates the time sequences of neighbor numbers of the two mobile sinks. Since the collected illustration results are almost same for the two WSN-MSs with orthogonal channels, we plot the CA-ETX experiment results in Fig.3.5 for brevity.

In this experiment, the packet size, node transmission power, and sensing rate, were set as 34 bytes, -25 dBm (results in around 2-3 meter transmission range), and one packet per five seconds for each sensor node respectively. Fig. 3.6 shows the Cumulative distribution function (CDF) of the end-to-end packet transmission delays. It can be seen that more than 90% of packets were transmitted within 10 seconds by using CTP with CA-ETX, while only 40% packets in the CTP with ETX experiment achieved this performance. The average delay of the CA-ETX experiment is 3.6 seconds, which is 73% less than that of ETX experiment (i.e. 13.4 seconds).

Table 3.5: End-to-end delay (in seconds) of opportunistic shortest path routing with different link metrics.

	queue backlog	Tx/Rx Rate
CTP with ETX	10.32 packets	5.63 packets/node/second
CTP with CA-ETX	7.83 packets	5.77 packets/node/second

As shown in Fig.3.6 and Tab.3.5, CTP with CA-ETX also achieves smaller packet loss rate, lighter average storage overhead (queue length) and similar communication overhead (Tx/Rx rate), compared with CTP with ETX. This experiment shows that although the real-time routing protocol CTP is originally designed for static WSNs, it has a great potential to be

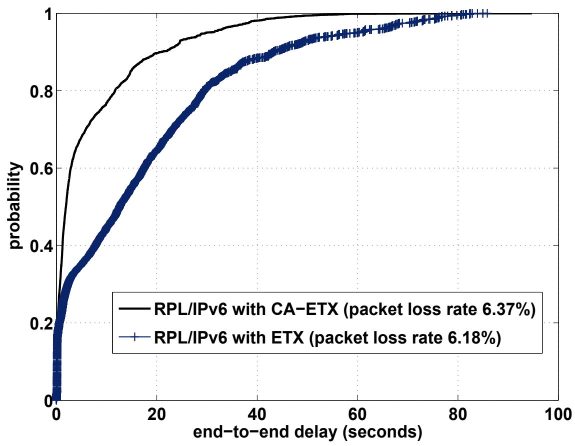


Figure 3.7: Packet loss rate and CDF of end-to-end delay: Modified RPL/IPv6 with ETX and CA-ETX.

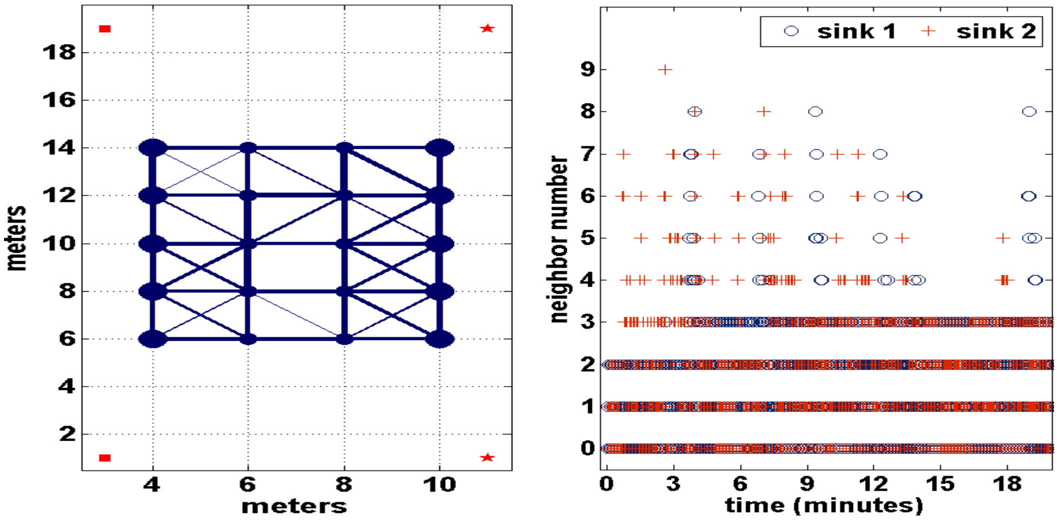
extended to support delay-tolerant data traffic in the intermittently-connected WSN-MSs, by using CA-ETX.

CA-ETX with RPL

In this subsection, we evaluated the performance of RPL with CA-ETX by using Cooja, the simulator of Contiki OS. As shown in Table 3.3, we modified 68 lines of code to implement CA-ETX and an enhanced loop detection scheme to RPL. We also changed some codes in Contiki's IPv6 stack to support node mobility and delay-tolerant applications, including implementing a queue at the network layer, faster neighbour table updates, and broadcasting neighbour solicitation messages more frequently.

In this set of simulations, we consider a multi-hop WSN-MS consisting of 200 randomly-deployed sensor nodes and 10 mobile sinks in a 500 m$\times$500 m simulation area. Each sensor node generated one UDP packet (40 bytes) per 40 seconds. We use a realistic mobility model, the Heterogeneous Human Walk (HHK) model [199] to simulate sink mobility. The sink movement speeds were randomly set between 1 m/s and 9 m/s respectively. Each simulation lasted for 2000 seconds. Fig. 3.7 shows the CDF of end-to-end delay for the modified RPL/IPv6 with ETX and CA-ETX respectively.

It can be seen that using ETX and CA-ETX result in similar packet loss rates of RPL, and the



(a) connectivity visualization (b) sink mobility

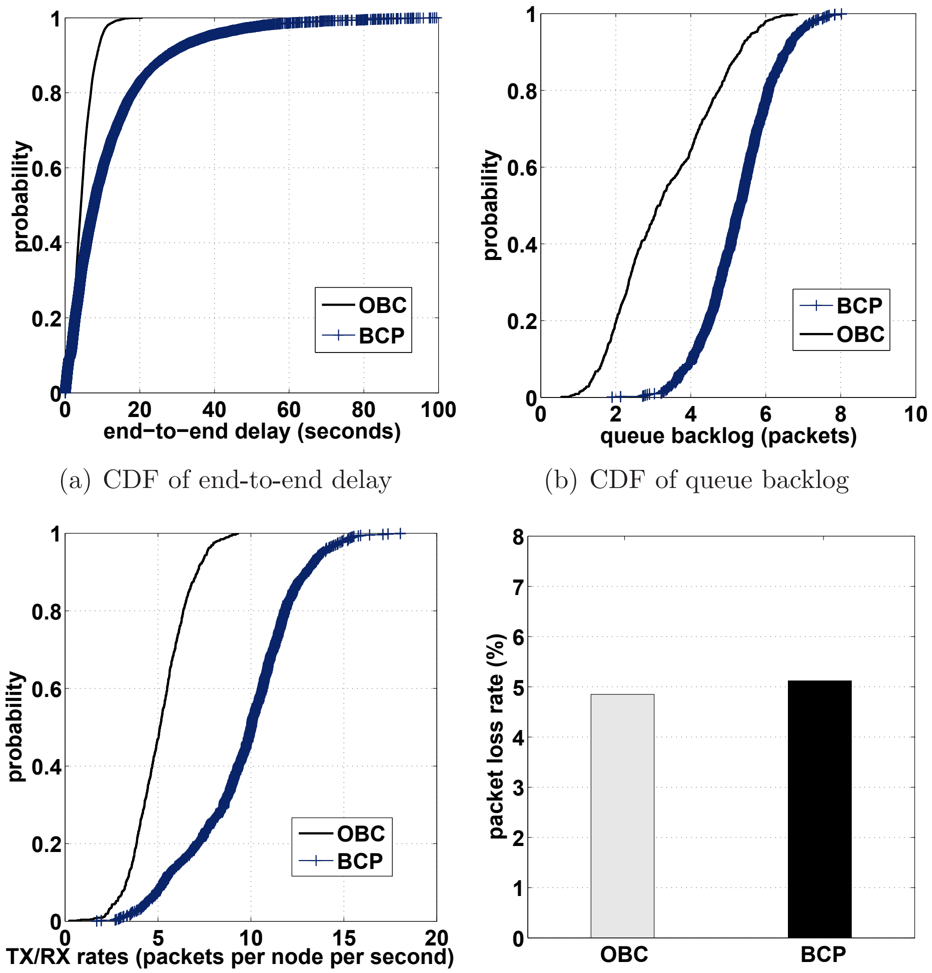
Figure 3.8: Visualization of the BCP and OBC testbed experiment settings: (a) Sensor deployment topology and mobility hot spots, (b) the number of neighbors for the two sinks.

transmission delay of around 20% data packets are similar (less than 1 seconds) when using the two metrics. These 20% data packets could be either sent from sensor nodes directly to the sink within one-hop, or through temporally existing multi-hop paths. In comparison to ETX, however, CA-ETX significantly reduces the delay of data packets (around 80%) that were transmitted through opportunistic multi-hop paths. As a result, the overall average of end-to-end delay of RPL with ETX (19.92 seconds in average) is approximately three times larger than that of CA-ETX (6.63 seconds in average).

3.6.2 Evaluation of OBC

Testbed Experiments

In this subsection, we compared OBC with a practical backpressure-based WSN routing protocol BCP [137], using real-world experiment with TinyOS-based MicaZ motes. Since BCP is a pure backpressure routing protocol without scheduling, we only compared the routing part of OBC with BCP for fairness. The method of this experiment was the same as that of the CA-ETX experiment in Subsection 3.6.1 but sensor layouts and sink mobility were different, which are illustrated in Fig. 3.8. In addition, sensing rates of OBC and BCP were set as



(a) CDF of end-to-end delay (b) CDF of queue backlog

(c) CDF of power consumption (d) average packet loss rate

Figure 3.9: OBC experiment results.

one packet per two seconds, while all other parameter settings were the same as that in the CA-ETX experiment.

As shown in Fig. 3.9, OBC significantly outperforms BCP in end-to-end delay, storage overhead, and communication overhead, while achieving similar packet loss rates as BCP. When a sensor node x transmits data packets to a mobile sink in the WSN-MS running BCP, its queue length reduces, resulting in a queuing gradient towards x. However, the mobile sink may disconnect from x before the convergence of such a gradient, resulting in severe routing loops. Such time-varying queue gradients caused by sink mobility aggravate the hop count performance of pure backpressure routing which is known to perform poorly already in static networks (e.g. [206]). In OBC, however, data packets are continuously forwarded to sensor nodes with low node CA-ETX values, which significantly reduces convergent time therefore mitigates routing loops (i.e. CA-ETX gradient helps the convergence of queue gradient). This results in much less unnecessary data transmissions (46% less average TX/RX rates), as shown in Fig. 3.9(c).

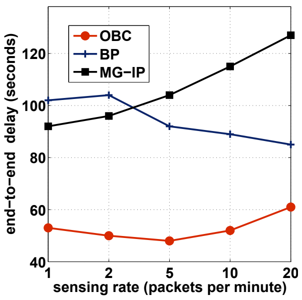
Furthermore, the network capacity resource (i.e. opportunistic contacts between sensors and mobile sinks) is better utilized by OBC, compared with BCP. This is because that a sensor node cannot transmit data to any nearby mobile sink during the slots when its queue is empty, while OBC minimizes the number of such slots by ensuring that better gateways (sensors with lower CA-ETX value) have high probability to receive more data. However, BCP treats all sensor nodes homogeneously and inherently tries to balance all queues in the network. Therefore, the probability of empty queues in good gateways is much higher, resulting in a waste of sensor-sink contact opportunities. Due to the the better network capacity resource usage and less data forwarding hops (less routing loops), OBC achieves much less network congestion and therefore much smaller queue backlog (59.1% in average) and less end-to-end delay (38.2% in average) than BCP, shown in Fig. 3.9 (b) and (a) respectively.

Simulations

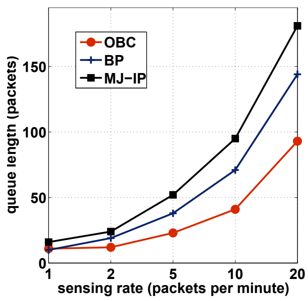
We constructed extensive simulations to further evaluate the performance of OBC, in terms of throughput, adaptability to sink movement, and scalability.

Simulation Settings. We compared OBC with a state-of-the-art protocol in WSN-MS that is based on mobility graph and information potentials (MG-IP) [108], and the classic backpressure routing/scheduling algorithm (BP). As shown in Table. 3.4, all the three protocols were implemented on the top of Castalia CSMA link layer. We implemented LQF and the back-off window adjustment technique [13] for the distributed scheduling of OBC and BP.

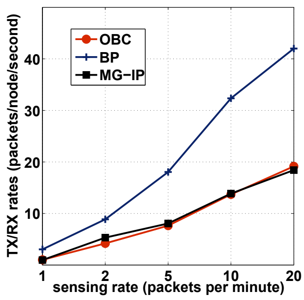
It is worth noting that OBC and BP *do not* require any knowledge of the dynamic system states, but MG-IP requires to forecast future sink mobility. In our simulations, we set the mobility prediction accuracy of a mobile sink m, PAC_m as a decreasing function of m's speed



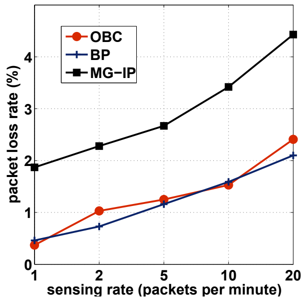
(a) end-to-end delay



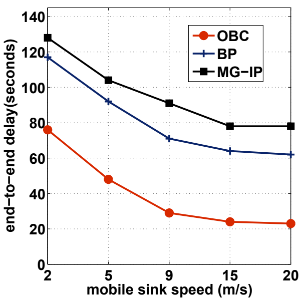
(b) storage overhead



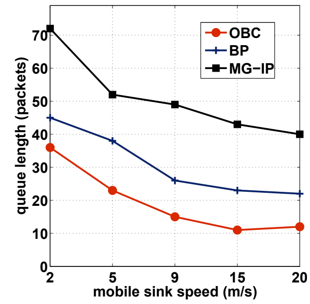
(c) communication overhead



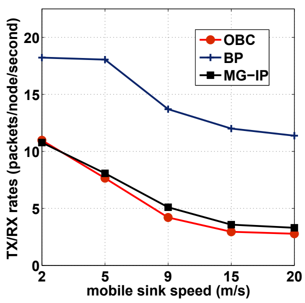
(d) packet loss rate



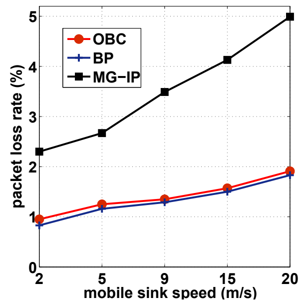
(e) end-to-end delay



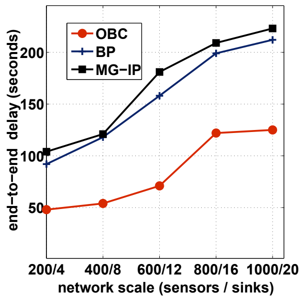
(f) storage overhead



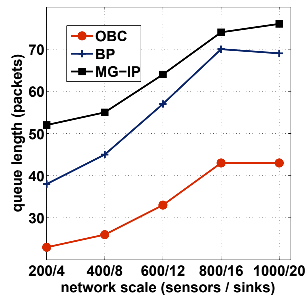
(g) communication overhead



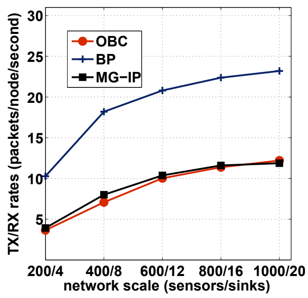
(h) packet loss rate



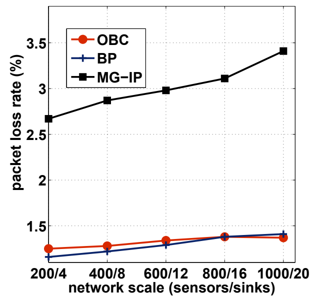
(i) end-to-end delay



(j) storage overhead



(k) communication overhead



(l) packet loss rate

Figure 3.10: Simulation results of OBC, BP, and MG-IP:(a)–(c) with different sensing rates (average sink speed is 5 m/s); (d)–(f) with different sink speed (sensing rate is 2 packets per minute); (g) –(i) with different scale WSN-MSs (average sink speed and sensing rate are 5 m/s and 2 packets per minute respectively, the average node density of different scale networks remained the same, by adjusting the size of corresponding simulation areas.

spd_m for MG-IP

$$PAC_m = \begin{cases} (100 - spd_m/4)\%, & \text{if } \mathrm{spd_m} \leq 20 \text{ (m/s)} \\ 95\% & \text{otherwise} \end{cases} \tag{3.17}$$

which is much higher than the accuracy of mobility-graph based prediction reported in [108]. The aim of this setting is not only to show the limitations of prediction-based WSN-MS protocols, but also to demonstrate that OBC can significantly outperform MG-IP even when its prediction accuracy is nearly perfect. Every simulation was run five times to obtain the average result.

Impact of Traffic Loads. Fig. 3.10(a)-(d) show the performance of the three algorithms with different sensing rate in a WSN-MS consisting of 200 sensor nodes and 4 mobile sinks, in a 600m $\times$ 600m area. The average sink mobility speed is 5m/s. As shown in Fig. 3.10(a), MG-IP and BP show opposite trends when the network data traffic changes. In general, end-to-end delay mainly depends on two factors: queue backlog (i.e. Little's Theorem) and routing path length. Since the routing decision of the mobility-aware MG-IP is independent of network traffic, it suffers from larger delay in simulations with heavier network traffics, caused by purely increased queue length shown in Fig. 3.10(b). However, BP makes routing decision based on queue backlogs. As network traffic load increases, its delay decreases. This is because the reduced routing loop is a more dominant factor than the increased queue backlog. Fig. 3.10(b) and (c) respectively show that purely queue-aware BP achieves better queue length but worse communication overhead than the purely mobility-awareness MG-IP. By combining mobility and queue awareness, OBC achieves the best delay performance in almost all these simulations. Finally, the packet drops of MG-IP are caused by both wireless channel contention and imperfect mobility predictions, which result in worse reliability performance than BP and OBC (no prediction requirement), as shown in Fig. 3.10(d).

Impact of Sink Movement Speed. Fig. 3.10(e)-(h) show the performance of the three algorithms with different sink movement speeds. In this set of simulations, the WSN-MS also consists of 200 sensor nodes and 4 mobile sinks in a 600m $\times$ 600m area. As the sinks moved faster, the variance of packet service times over sensor-VS links decreases, as we discussed in Subsection **??**.

This results in faster packet transmission, lighter network congestion, and less routing loops, for all three routing protocols, shown in Fig. 3.10(e)-(f) respectively. In addition, MG-IP exhibits a large the packet loss rate in simulations with high-speed mobile sinks, caused by non-ignorable mobility prediction errors, while BP and OBC are relatively insensitive to sink movement speed in terms of transmission reliability. Furthermore, that reliability performance of MG-IP would be significantly degraded in practical WSN-MSs, where high prediction accuracy (e.g. more than 95% in our simulations) is impossible to achieve.

Scalability Study. Finally, we studied the scalability of the three protocols. Since all three protocols only require local information (i.e. queue backlogs, CA-ETX, and information potentials) to make routing and data forwarding decisions, their control complexities are O(1) with respective to the network size, which demonstrates their potential to scale to large WSNs-MSs. However, the simulation results shown in Fig. 3.10 (i)- (l) show that the performance of all three protocols generally decreases as the network scale increases, but the performance degradation speed of OBC is the slowest. It can also be seen that OBC outperforms the other two protocols in both small-scale and large-scale WSN-MSs.

Besides network size, routing protocols in WSN-MSs are also expected to scale with respect to the number of sinks. Both OBC and BP only require to maintain one data queue for anycast data traffic, and adopt lightweight mobility scheme (i.e. one-hop beacon and CA-ETX) without maintaining any information of a specific moving sink. Therefore, they are relatively insensitive to sink population. In contrast, MG-IP needs to store n information potential values for n mobile sinks respectively. Furthermore, MG-IP also suffers from high complexity due to the maintenance of a mobility graph for sink mobility prediction, which restricts its application in large-scale WSNs with a large number of sinks.

3.7 Summary

In this chapter, we study how to improve the delay and throughput performance for delay-tolerant data collection applications in Wireless Sensor Networks with Mobile Sinks (WSN-

MSs). We propose a novel routing metric, CA-ETX, based on queueing analysis theory to estimate the packet transmission delay caused by both link unreliability and intermittent connectivity. By implementing CA-ETX in CTP and RPL, we demonstrate that current ETX-based routing protocols for WSN with static sinks can be easily applied to WNS-MSs by using CA-ETX. We also introduce a throughput-optimal data collection scheme, OBC, by integrating CA-ETX into the Lyapunov optimization framework. In contrast to current routing schemes for WSN-MSs, OBC does not require any mobility prediction and is suitable for large-scale sensor networks with a large number of fast moving sinks. Test-bed experiments and extensive simulations demonstrate the significant performance improvement achieved by OBC, compared with state-of-the-art approaches. Interesting future directions lie in the extension of CA-ETX and OBC to duty-cycled WSN-MSs and to support a hybrid of delay tolerant and real-time sensing applications.

Several problems remain for ubiquitous sensing using WSN-MSs.

(1) What should the mobile sink do after collecting sensor data? Could the data be sent via the Internet through expensive long-range communication (e.g. 3G cellular), or through low-cost short-range communications (e.g. WiFi) when the mobile sink passes a free access point?

(2) Since wireless communications result in energy and bandwidth costs, and even monetary bills, how can mobile device owners be incentivised to collect and transmit data?

To address these problems, the next chapter will consider a more realistic and low-cost sensor data collection paradigm for WSN-MP in which mobile devices such as smart phones are used to sense and forward collected data to statically-deployed sinks, for example, free WiFi access points and cellular base stations can be utilised as sinks.

Chapter 4

Citizen-centric Network Architecture for Mobile Phone Sensing System

The mobility in Wireless Sensor Networks (WSNs) improves the performance of networks in terms of throughput, coverage and cost. In chapter 3 we studied data collection from static sensor nodes for delay-tolerant applications in Wireless Sensor Networks using mobile devices. In last few years, the focus of wireless sensor networking research has evolved from static networks of sensor nodes to Mobile Phone Sensing Systems (MPSS) relying on the smart devices and mobility of people. In a MPSS, ubiquitous sensor-rich mobile phones contribute to the understanding of future cities using its embedded sensors. Moreover, mobile devices can transmit sensor data using short-range or long-range communication.

This chapter addresses following key challenges in MPSS.

- How to communicate huge volumes of sensor data from mobile phones to sinks (i.e cellular base station or Wifi Access point) in a cost-effective way.
- How to incentivize mobile phone users to participate in the sensing system.

In this chapter, we develop a novel citizen-centric networking scheme to support both real-time and delay-tolerant urban sensing applications via the seamless integration of inexpensive short-range opportunistic transmissions and reliable long-range cellular radios. Core to this is trading of mobile sensor data in a virtual free market where we demonstrate that our scheme provides a strong incentive system for phone owners, while achieving network throughput optimality and minimizing phone users total costs in terms of their 3G budget and battery levels. The proposed scheme considers realistic urban sensing issues such as user privacy and demonstrates how social network awareness improves data transmission. Our scheme is fully distributed and it self-configures and self-adapts to the environmental changes regarding mobility, topology, and channel conditions. We evaluated our approach using a real testbed and extensive simulations.

WSN with Mobile Sinks (WSN-MSs) studied in Chapter 3 considers data forwarding from sensor nodes to and mobile devices. In this chapter we focus on data forwarding from mobile devices to the Wifi Access points or cellular base station.

TTL	Time-to-Live value of each packet.
$NEI(x,t)$	One-hop neighbour table of x at slot t.
$rate_{x,y}(t)$	Transmission capacity between a phone x and its neighbours y.
$scost_x(t)$	Monetary cost of sending a packet at slot t.
$rcost_x(t)$	Monetary cost of receiving a packet at slot t.
α	A positive system parameter set by the server t.
$sell_x(t)$	Monetary value of the data packets of a phone x.
$profits_{x,y}(t)$	Potential individual profits a phone x at slot t.
$max-lifetime$	Simulation variable to set maximum life time of packets.
$ST(x,y)$	trust metric between two phones x and y.
$max-trust$	Simulation variable to set maximum trust between two phones.

Table 4.1: Summary of symbols used in Chapter 4.

MPSS	Mobile Phone Sensing System
WSNs	Wireless Sensor Networks
GPS	Global Positioning System
D2D	Device-to-Device communication.
E2E	End-to-End
QoS	Quality of Service.
OP	Opportunistic communication
PAYG	Pay as you go

Table 4.2: Summary of abbreviations used in Chapter 4.

4.1 Introduction

Sensor-rich smart phones are predicted to play an increasingly important role in urban sensing, so much so that a number of next generation phones will be augmented with even more environmental sensors such as GPS, accelerometer, microphone, camera, gyroscope, digital compass and barometer. Given the ubiquity and ever-increasing capabilities of sensor-rich mobile devices, Mobile Phone Sensing System (MPSS) provide a highly flexible and ready-made wireless infrastructure for future smart cities than traditional static Wireless Sensor Networks (WSNs) [110] [104] [64]. At the same time, inherent mobility of phone users enables increased sensing coverage both spatially and over time, providing opportunities to collect data at a higher granularity and with more penetration. Furthermore, increasing short range communication capabilities of smart phones (LTE, bluetooth, wifi-direct) has enabled mobile data off-loading using low cost(free), short range communications. This releases the burden on traditional communications technologies, which will reach upper physical bounds if all future MPSS systems use them [78]. Mobile sensing can also exploit the social structures of the physical world to improve the performance of cyber world and in doing so provides better services to the users in the physical world by optimising the organization of the available resources in cyber world. This paves the way towards large-scale citizen-centric urban sensing applications for smart cities [148].

Fig. 4.1 illustrates a typical Mobile Phone Sensing System (MPSS). According to the demands of specific sensing applications, mobile phones produce sensing data such as available parking places, traffic congestion, noise levels, air pollution, and smart meter readings. The sensor data can be sent to the MPSS server through either cellular communication or short-range radios such as WiFi direct [36], Bluetooth, and LTE direct [119].

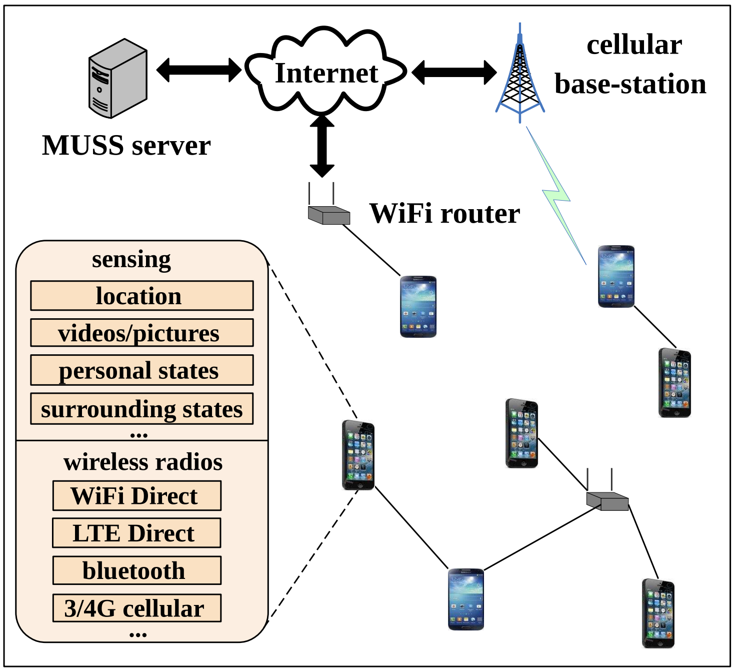


Figure 4.1: Conceptual illustration of the Mobile Phone Sensing System.

MPSS is predicted to be one of the major sources of *big data* due to prolification of mobile devices and their potential for many urban sensing applications. Furthermore, sensor data would differ in lifetime, monetary value, QoS requirements and privacy levels, etc. due to the diversity of sustainable city applications.

Efficient transmission of potentially huge volumes of sensor data with different QoS requirements over expensive cellular networks with limited bandwidth is the key challenge in MPSS [78]. However, recent forays into cheaper and robust opportunistic alternatives, leveraging urban mobility (e.g. human and vehicle) and short-range communications of mobile phones, have suffered from large end-to-end delays.

In this chapter, we show how to provide autonomic and cost-effective networking services for MPSS by combining short-range and cellular communication. We firstly examine the major communication techniques designed to support MPSS. We then propose a hybrid architecture for MPSS by combining both cellular communication [128] and opportunistic networking [45] for a large scale city network. To demonstrate the feasibility of our architecture, we develop a ***joint*** pricing and data routing algorithm aiming to support both real-time and delay-tolerant MPSS applications.

The MPSS is underpinned by an economic market model in which phone users produce sensor data according to the demands of the MPSS applications, trade sensor data with each other to obtain profits, and finally sell data to WiFi routers or cellular base-stations for returns. By doing so, the proposed scheme can achieve not only effective incentivization for the phone users, but also throughput optimality for big sensor data transmission and minimal costs of phone users (e.g. 3/4G budgets and battery energy). Our scheme is distributive and scalable which makes it robust to dynamic environment and variable user preferences and behaviours. Through simulations, we show the impact of users preference and privacy on the performance of MPSS. We also show that MPSS is able to adjust itself under dynamic conditions such as disasters, variable pricing and joining and leaving of users. Finally, we developed an android application (OppCom) to show the feasibility of our approach in real world scenario.

4.2 Communication Supports for Mobile Phone Sensing

4.2.1 Cellular Communication

Currently, the majority of mobile sensing applications send sensing data directly to the server through single-hop 3/4G *cellular radio communications*. This is very suitable for real-time applications, since the use of this expensive cellular service is justified for important and delay-sensitive data. However, due to limitations such as 3/4G costs to the phone users [128] and cellular system's capacity bounds [54], using cellular communication solely would not be a feasible solution for the potential huge volume of urban sensing data.

4.2.2 Opportunistic Networking

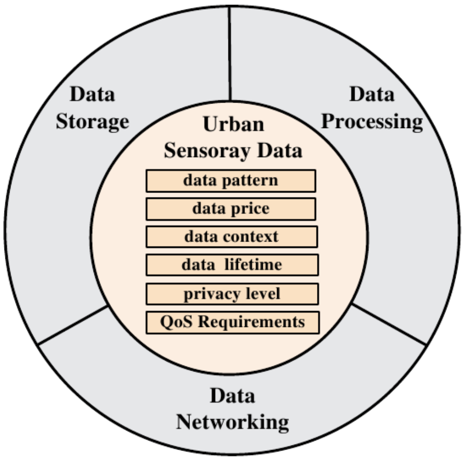
Opportunistic networking [45] enables data communication in intermittently-connected delay-tolerant mobile networks, where for a given instance an end-to-end communication path between sensor and ultimate destination may be absent. By leveraging inherent human mobility and low-cost short-range communication, sensor data can be sent to base-stations (e.g. WiFi routers) in a "carry-and-forward" fashion by relaying the data in short hops via different mobile phones.With the increase in the short-range communication capabilities of smart phones, such as in WiFi Direct for Android OS 4.0+, efficient neighbour discovery [23], and the development of smart Device-to-Device (D2D) communications [54]; it becomes more and more promising to use *opportunistic networking* for delay-tolerant MPSS applications [23], [151], [200].

This opportunistic networking can significantly reduce energy and telephony costs for phone users and at the same time mitigate sensor data traffic load over cellular communication channels [78]. However, for such schemes to support sensing at the scale of a city, a distributed approach is required; in terms of initial configuration and maintenance over time, and be agile enough to overcome failure. Added to this is the privacy and security concerns, these could discourage phone users from relaying data from unfamiliar parties [75].

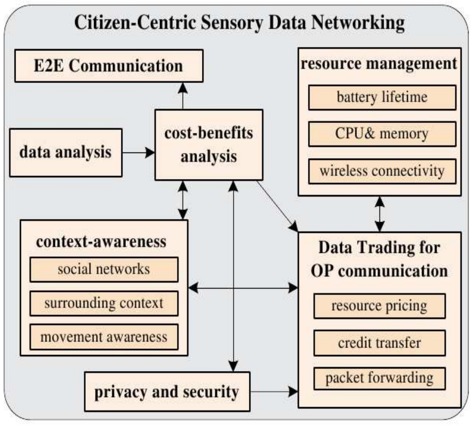
4.3 MPSS Network Architecture

As shown in Fig. 4.2(a), networking is an important component in a MPSS. Considering the characteristics of MPSS data and the communication paradigms we have discussed in earlier sections, we propose a network architecture to provide cost-effective sensing applications for the future intelligent cities, illustrated in Fig. 4.2(b). The proposed architecture consists of the following three components:

1. **Data Analysis**. This component analyzes networking-related properties of every sensor data packet or data stream, such as the packet lifetime metric time-to-live (TTL); monetary data value (i.e. the importance of the data; the value of a fire alarm message should



(a) MPSS Components



(b) Network architecture

Figure 4.2: MPSS Architecture.

be much more expensive than periodic raw temperature data for daily environment monitoring); QoS requirements (e.g. jitter tolerance for video streams); and a value attributed to the notion of privacy. The MPSS can decide what kind of services should be provided for a given data packet, or data stream, by using the information above.

2. **Cost-benefit analysis**. This component quantitatively analyzes the benefits and costs that a sensor data packet, or stream, can bring to the networking service providers such as; phone users, E2E network owners, and the Internet access provider (e.g. WiFi router). This may be based on the networking service providers context, computing and communication resources, as well as privacy concerns. For instance, where there is no 3G cellular coverage, a phone may only forward a packet to other nodes that are in the nearby region, or if this is not possible, the packet is dropped. This component optimizes the networking efficiency, and more importantly, provides the notion of incentivisation to drive the policies of MPSS. For instance, a phone user may not be willing to forward low valued sensor data to through the 3G/4G network where he or she has a limited 3G/4G usage contract (e.g. 1G data per month), or the phone currently has less than 10% battery energy left. In addition, personal preferences, as well as security and privacy concerns, have significant impacts on the network provider's decision making.

3. **Decision making and data transmission**. After cost-benefit analysis, the network

service provider can not only decide whether to forward the sensor packets or not, but also decide on the type and volume of data that should be transmitted (in terms of computation, communications and the human effort that would be required, and the rewards received for carrying out the transmission) .

In many societies, individuals work to improve their social profit and obtain rewards according to their performance. In this sense the proposed network architecture treats the MPSS as a digital society, where every citizen can be a network provider as long as he or she is willing, and able, to improve the profits of MPSS. Sensor data in future cities will be a commodity like any other commodity in current economic systems. Here a large volume of easily attainable goods are cheap, yet rarer goods are expensive. So too in this network architecture. Here the citizen is encouraged to transmit larger volumes of low-value (and presumably delay tolerant) sensor data using OP communications, which has a great network capacity.Then more expensive Internet access can be used for the small amount of valuable (potentially real-time) data through the expensive, higher quality E2E communications mechanisms.

4.4 A Citizen-centric Networking Scheme for MPSS

In a MPSS, mobile phone will belong to individuals with different personal preferences. Mobile phone users may not be willing to fulfil a MPSS task, due to privacy concerns and the potential costs that would be incurred; impacting battery usage and 3/4G budgets. Therefore, taking account of the social and economic behaviours of phone users, though frequently ignored, is central to the success of MPSS.

To demonstrate the feasibility and efficiency of the proposed scheme in terms of social and economic behaviours of both citizens and the system, we present a simple distributed data trading and networking algorithm in to support both real-time and delay-tolerant MPSS applications in a cost-effective way, through the combination of cellular communication short-range communication.

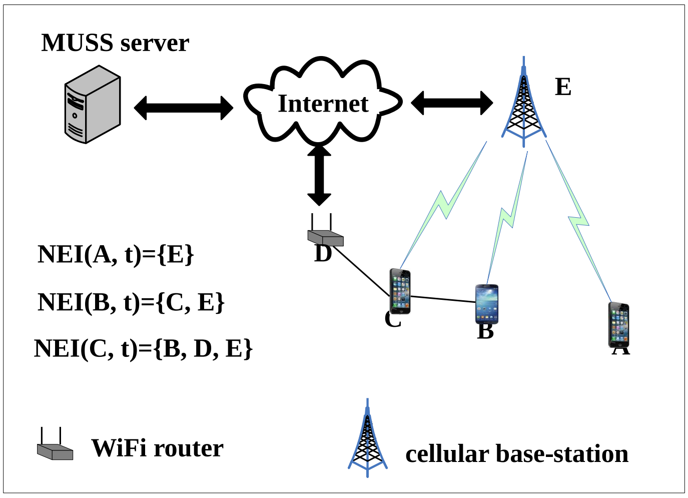


Figure 4.3: Example of MPSS network to describe the proposed scheme.

4.4.1 Network Model Scenario

We consider a MPSS network that consists of three types of nodes: smart phones, static WiFi routers, and a cellular base-station as shown in Fig. 4.3 Each phone can report sensed data to the mobile sensing server through 3G cellular radio directly, or through a WiFi router nearby. In addition, two nearby phones have the opportunity to communicate directly to each other through WiFi Direct during their contact duration, such as phones B and C shown in Fig. 4.3. In our model, each data packet produced by a smart phone has a monetary value (e.g. which can be represented in terms of a national currency or tokens to be traded in other ways such as to purchase mobile phone apps). Further, each packet is has a lifetime *e.g.*, 10 minutes, and its duration is tightly coupled to the worth specific applications attribute to the packet. The MPSS operates in discrete time with a unit time slot t =1, 2, ... Every phone x maintains a data buffer that stores the sensor data packets generated by its own sensors, and the data received from other phones. Initially the queue backlogs of the base-station in the MPSS is assumed to be zero.

4.4.2 Algorithm Description

At every time slot t =1, 2, ..., our scheme operates as follows:

Sensor Data Sampling

1. According to the requirements of the MPSS application (e.g. the demands of external MPSS users), each phone x generates sensor data packet(s), and then assigns its monetary value and initial Time-To-Live (TTL) value to each packet based on data analysis discussed in earlier section. Then, x inserts the sensor data packets into its phone data buffer.

Instantaneous Neighbour Discovery

2. Each node x builds a one-hop neighbour table $NEI(x,t)$, consisting of the cellular base-station, and all phones and WiFi routers that can connect to x through WiFi radios at current slot t. Neighbour discovery schemes such as [23] can be used to populate the one-hop neighbour table. Take Fig. 4.3 for instance, the instantaneous one-hop neighbour table of phone C, $NEI(C,t)$, consists of three nodes: B, D, and E.

Transmission Quality Estimation

3. Each node x estimates its transmission capacity, $rate_{x,y}(t)$, between itself and each of its instantaneous neighbours y in $NEI(x,t)$, i.e. the maximum number of packets that x can transmit to y, based on the data rates of their wireless radios and WiFi duty-cycle settings of x and y [23].

4. Each node x estimates the monetary costs of sending and receiving a packet, denoted as $scost_x(t)$ and $rcost_x(t)$ respectively, based on its remaining energy, system resource usage, and 3G bills costs. It worth noting that if x is a WiFi router or the cellular base-station, its receiving cost, $rcost_x(t)$, is equal to zero.

Pricing

5. Each phone x sets its current data selling price, $sell_x(t)$, as the total monetary value of the data packets in its data buffer multiplied by a positive system parameter α, set by the server. For instance, if $\alpha = 0.1$ and x's data packets are worth 10 cents, therefore $sell_x(t) = 1$ cent per packet. Then phone x communicates the selling price $sell_x(t)$ to all nodes in $NEI(x,t)$. Recall, the selling prices of any cellular base-station and WiFi router are set as zero for every slot.

Profit computation

6. Each phone x computes the potential individual profits, $profits_{x,y}(t)$, it could obtain by selling data to each of its neighbours y in $NEI(x,t)$. $profits_{x,y}(t)$ is computed as a function of the cost (that would incur in this potential data trading) and selling price differences between x and y

$$profits_{x,y}(t) = (sell_x(t) - sell_y(t) - scost_x(t) - rcost_y(t))rate_{x,y}(t) \tag{4.1}$$

Data Trading

7. Denote y^* as the neighbour that can currently give phone x maximum profit, if it sells on its data. If $profits^*_{x,y}(t) > 0$, then x sells $rate^*_{x,y}$ (t) number of packets to y^*. Note that the number of packets are a function of the communications rate so as to not overload that link. A data packet with a smaller TTL will be forwarded with a higher priority. Packets which have reached a 0 TTL value will be dropped as they are deemed no longer useful to the application.

8. Upon receiving data packets from the seller x, the buyer y^* pays $(sell_x(t) - sell_y(t) - rcost_y(t))rate_{x,y}(t)$ total amount of money to x, which means that the cost incurred in this trade is paid by the seller x.

It can be seen that the proposed scheme is fully distributed and scalable, because it requires only the local information of each mobile phone and its current one-hop neighbours. In addition, this scheme is very lightweight, as it implements simple arithmetic calculations and does not require any historic information to be maintained, nor does it require future knowledge of mobile phones and their trajectory to be speculated.

4.4.3 Economic Interpretation of the System

By using the proposed networking scheme the MPSS acts as an economic system, shown in Fig. 4.4. Smart phones produce and trade sensor data with others in a virtual market, according

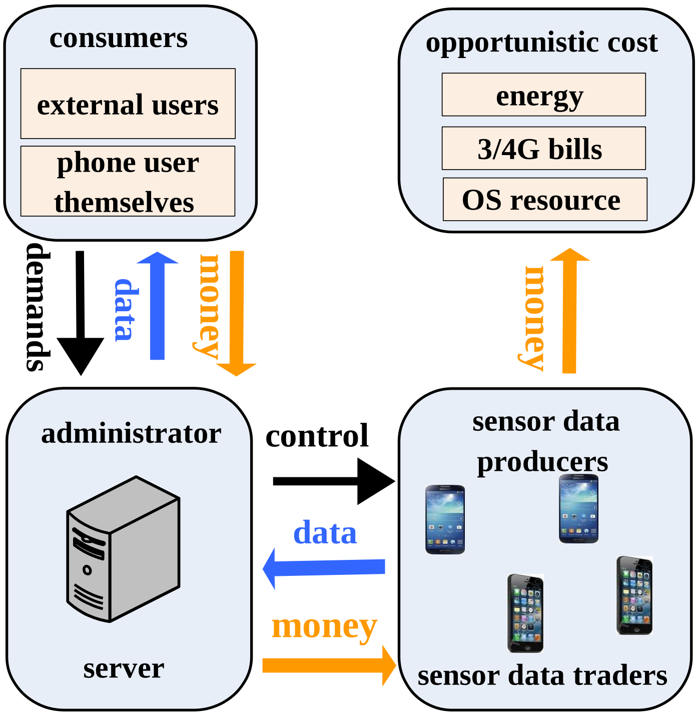


Figure 4.4: Economic Interpretation of the MPSS.

Now we discuss how the proposed scheme incentivizes phone users to participate. On one hand, by selling sensor data to others, a phone user can directly reap profits computed by Equation (4.1). On the other hand, a phone user would also be willing to *buy* sensor data from others because of the following two reasons: firstly a buyer can obtain sensor data at no cost; according to the data trading rule of the proposed scheme. Secondly, the selling price of a buyer is increased due to the grown data value, which will in turn increase its overall revenue. That is, the buyer will obtain more benefit by selling on sensor data to others (e.g. WiFi routers and the cellular based-station) at a higher price in the future; similar to real-world trading systems. The proposed scheme is suitable for a large-scale MPSS, where the trading system can be considered as a pure competitive market. Here, each phone user is described as a *price-taker*. This means that they believe that the pricing and trading rules are given as constants by the system and that these cannot be manipulated by actions of individuals. In small-scale MPSSs, where there are a small number of phone users in the system; phone users can strategically subvert their actions to manipulate the market and thus their return. In this case, algorithmic economic schemes such as distributed mechanism design [179] can be used for faithful implementation of the proposed scheme but this is beyond the scope of this chapter.

4.4.4 Network-Theoretic Interpretation of The System

Since the total value of the data carried by each phone in its buffer is proportional to its queue backlog, it can be verified that the proposed scheme implicitly solves a stochastic optimization problem (i.e. we minimize the total transmit and receive costs for all phones) in a fully distributed way, by using the Lyapunov "drift-plus-penalty" method [139]. According to Lyapunov optimization theory, optimal throughput and long-term minimization of global system costs can be achieved, by controlling the weight between queue backlogs and communications costs [139]. In our scheme, this weight is controlled by the price scaling parameter α. Based on the Lyapunov "drift-plus-penalty" method, it is not difficult to verify that as α decreases, the global system costs (total cost of all phone users) also decrease, but the average queue backlogs increase resulting an increase in end-to-end transmission delays. Therefore, by controlling the pricing parameter α , one can prove that the proposed scheme can not only achieve throughput optimality, which is highly desirable when transmitting large volumes urban sensing data; but it can also minimize the total cost incurred by the phone users [200], [139].

Since the instantaneous neighbor table on each phone can include a cellular base-station, and all WiFi routers and other phones nearby, the phones can automatically switch data transmission between WiFi radio and cellular radio, according to selling prices and transmission costs. In addition, heterogeneous data packet types are addressed by sending smaller-TTL packets with higher priorities. This buffer management operation does not affect the global queue backlog gradients of the system, and therefore does not affect the Lyapunov optimization guarantees.

4.4.5 Self-*

In MPSS, mobile phone users can join and leave the network very frequently. Similarly disasters and the dynamic nature of urban environments can cause network failure. Therefore, MPSS should be able to add nodes to the existing network without any manual configuration. Similarly, MPSS should be robust towards node and link failures and it should adapt autonomously to current network states such as channel conditions and evolving logical network topologies.

Besides achieving throughput optimality, our scheme exhibits the following autonomic behaviours verified by simulations in the next section.

1. *Self-optimization* Since the neighbour table on each phone can include a cellular base-station, and all WiFi routers and other phones nearby, the phones can optimize their profit by automatically switching data transmission between WiFi radio and cellular radio depending upon selling prices and transmission costs.

2. *Self-organization* This scheme is fully distributed, because it requires only the local information of each mobile phone and its current one-hop neighbours. This enables MPSS to self-organize based on current network state and topology. Moreover it is flexible enough to cope with partial failure of communication infrastructure *e.g.*, by natural disasters and can scale across urban space.

3. *Self-adaptation* Our scheme provides a tailored service based upon the available resources and requirements of the application and nature of the data. It reacts timely to the continuously changing properties of data according to performance objectives such as delay and cost.

4. *Self-Healing* Our scheme is self-healing under various permutations, such as node(s) leave or join the network. The system reconfigures it-self to choose the best available path to route the data to the server and maintains stability.

4.5 Evaluation

4.5.1 Simulation Settings

To evaluate the performance of our scheme, we constructed extensive simulations using the realistic simulator Castalia (http://castalia.npc.nicta.com.au/). We randomly deployed a 151-node MPSS in a $800m \times 800m$ geographic area, consisting of 10 WiFi routers, 140 mobile phones, and one cellular base-station. We set the duration of a slot to 1 second and each simulation lasts for 10^6 slots (around 12 days). The transmission ranges of the WiFi direct radio was set to 50 meters i.e. the typical WiFi direct transmission range in practice (http://www.wi-fi.org). The time-varying transmission capacities of all cellular and WiFi radios were randomly set between 1 and 50 packets per second. We used a realistic human mobility model, Heterogeneous Human Walk (HHW) [199], to simulate the mobility of smart phones. The movement speed of each phone was randomly distributed between 1 and 10m/s (i.e. representing walking speeds and typical urban vehicular speeds).

Each sensor and mobile phone produces sensor packets with a random monetary value of 10 credits at a rate of one packet per slot. For every mobile phone, the receiving and transmitting costs of WiFi radios were randomly set between 0.1 and 1 credits per packet, while that of the cellular communications were set between 1 and 10 credits per packet.

4.5.2 Impact of Packet Lifetime and Pricing Parameter α on the MPSS Performance

In this set of simulations, we study the impact of different packet lifetimes and the pricing parameter α on the global system cost and global social profits. The lifetime (i.e. the initial TTL value) of each generated packet was randomly set between 5 seconds and the $max - lifetime$ minute, this latter parameter is a simulation variable ranging from 10 to 50 minutes. The randomness of the packet lifetime assignment can reflect the heterogeneity of mobile sensing data. The simulation results are shown in Fig. 4.5, 4.6(a) and Fig. 4.6(b). In all simulations,

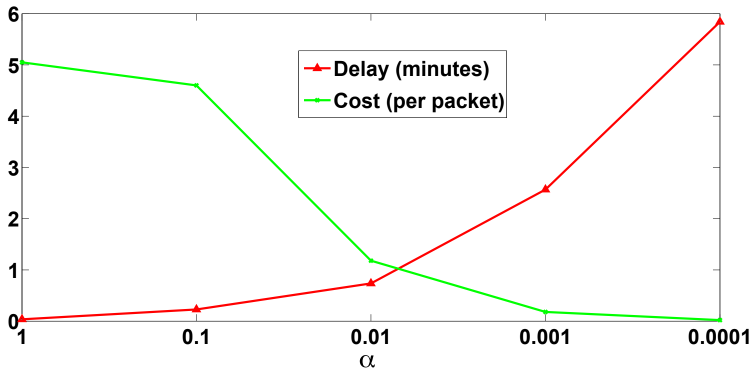
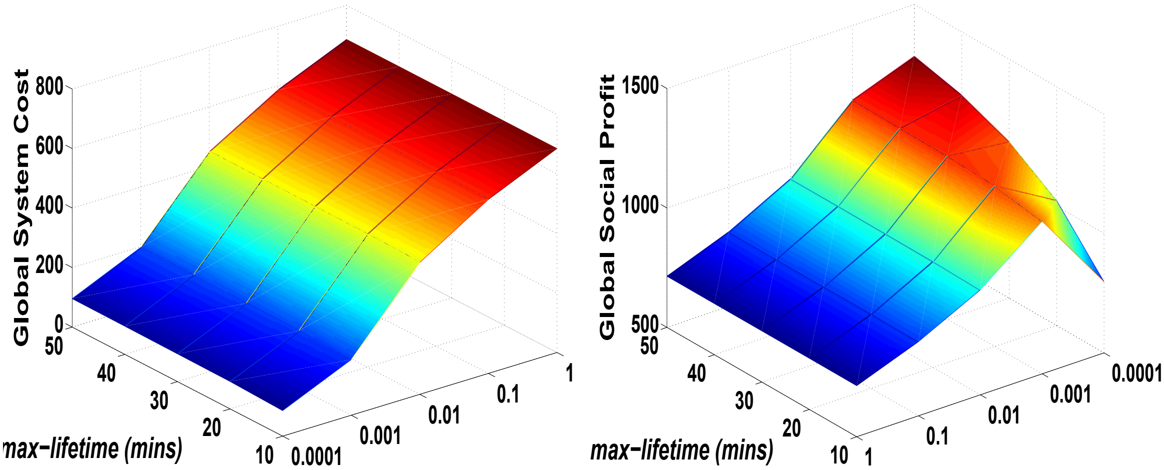


Figure 4.5: Impact of parameter α on packet cost and delay.



(a) Global Cost

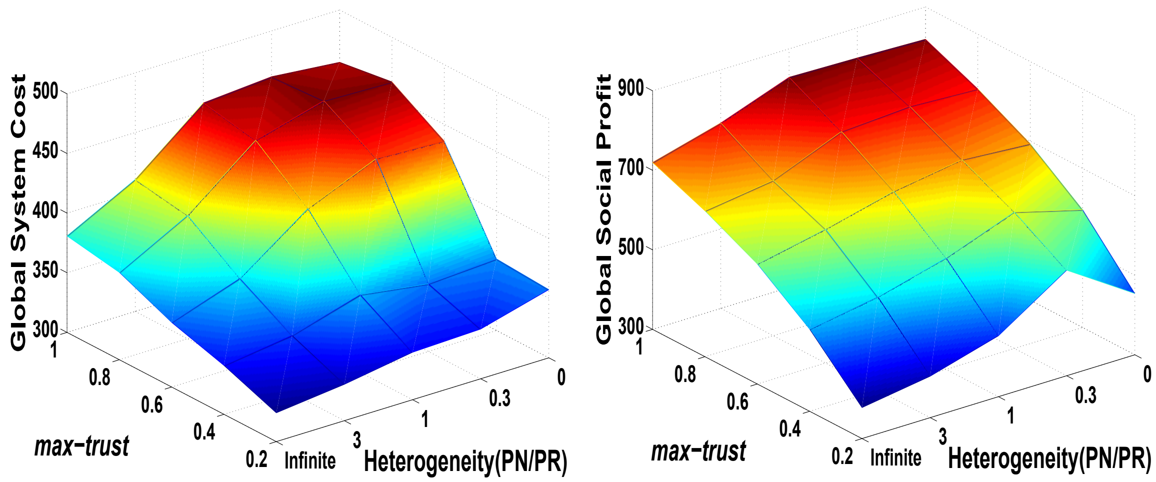
(b) Global Profits

Figure 4.6: Impact of sensor data lifetime and parameter α .

around 10%-65% of the sensor data traffic is sent through cellular radios, and the rest is sent over WiFi direct radios.

As illustrated in Fig. 4.5, the packet cost shows a monotonically decreasing trend as the pricing parameter α decreases; this verifies our optimal throughput discussion in Subsection 3.2. End to End delay increases with the decrease in packet cost. This is due to more packets are sent through Wifi direct radios to reduce the cost and less packets are sent through cellular radios with low delay.

We use time-average global system costs and global social profits (both in credits per second) to measure the performance of our scheme. Here the global system cost is measured as the sum of both the transmission and reception costs of all phones, and global social profits is computed as the total value of all the successfully received packets (by the MPSS server) minus the total



(a) Global Cost (b) Global Profits

Figure 4.7: Impact of user trust and social network topology.

system cost.

As shown in Fig. 4.6(a), by setting a sufficiently small α, the global system cost can be arbitrarily close to the minimal, according to Lyapunov optimization theory. However, the end-to-end delay becomes large as α decreases, resulting a higher risk of a packet being dropped, taking TTL into account. This is reflected in Fig. 4.6(b), where the global social profit shows a concave curve as α decrease when the packet life time is small. This is caused by the joint effects of decreased system cost and increase in dropped packets.

When max-lifetime is sufficiently large, global social profits exhibit a monotonically increasing function of α. This is because the impact of packet loss caused by expired TTL on the global social profits can be ignored. It worth noting that every phone obtained positive profit in all simulations. This means that our scheme has the potential to incentivize phone users to participate in the MPSS because they receive a fair reward.

4.5.3 Impact of User Preference and Privacy on System Performance

People have different privacy concerns and preferences. For instance, a phone user may restrict transactions to buy (or sell) sensor data from his or her friends than from strangers. To see how this would affect the MPSS, we constructed a set of simulations to study the impact of such social behaviours. A social network was established to simulate the social relationships between phone users, by using a realistic social network model proposed by Jackson et al [92]. We define a trust metric $0 < ST(x, y) < 1$ between two phones as the percentage of sensor data packets, which x is willing to sell to y (and in turn y is willing buy from x). For instance, if the trust between$ST(x, y) = 10\%$, and the link capacity, $rate_{x,y}(t) = 50$, then x can send at most 5 packets to y at time slot t. If two phone users share a social link, then their trust value will be randomly assigned between 0 and a positive value $max - trust$, the latter being a simulation variable ranging from 0.2 to 1; otherwise, their trust value was assigned as zero. The randomness of the trust assignment reflects different personal preferences. Fig. 4.7(a) and Fig. 4.7(b) show how the strength of trust and the underlying social network topology affect the system performance. We can see that the high level of trust between the phone users across the network results in high global social profits for the network as a whole. As one would imagine, low levels of trust reduce the utilization of WiFi direct communications between mobile phones and force phone users to send more packets directly through cellular communication, which results in higher costs and low global social profits. In addition, we can see that both global system cost and social profits changes considerably as the heterogeneity of the social network [92] changes, which demonstrates that the underlying social network topology also has a significant impact on the system performance.

4.5.4 Self-Configuration in Dynamic Environment

To study the ability of our proposed scheme to adapt to changing scenarios, we constructed experiments by dividing the total simulation time into five periods of equal duration of 2×10^5 secs. In first period, MPSS operates normally with operational cellular and Wifi communication. In the second period, we disabled the cellular communication of all nodes so that data packets can only be transmitted and received directly through Wifi direct radios (simulating cellular failure similar to what has occured in disaster situations). MPSS returns back to normal (mixed) state in the third period. In fourth period we disabled Wifi direct communication

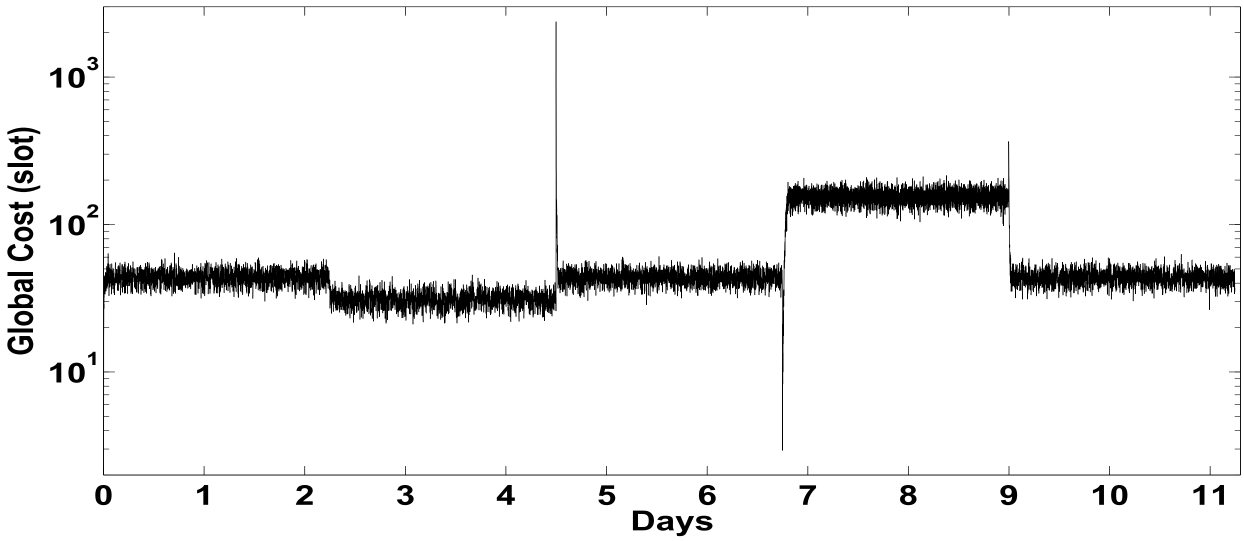


Figure 4.8: Impact of Dynamic environment on System Cost.

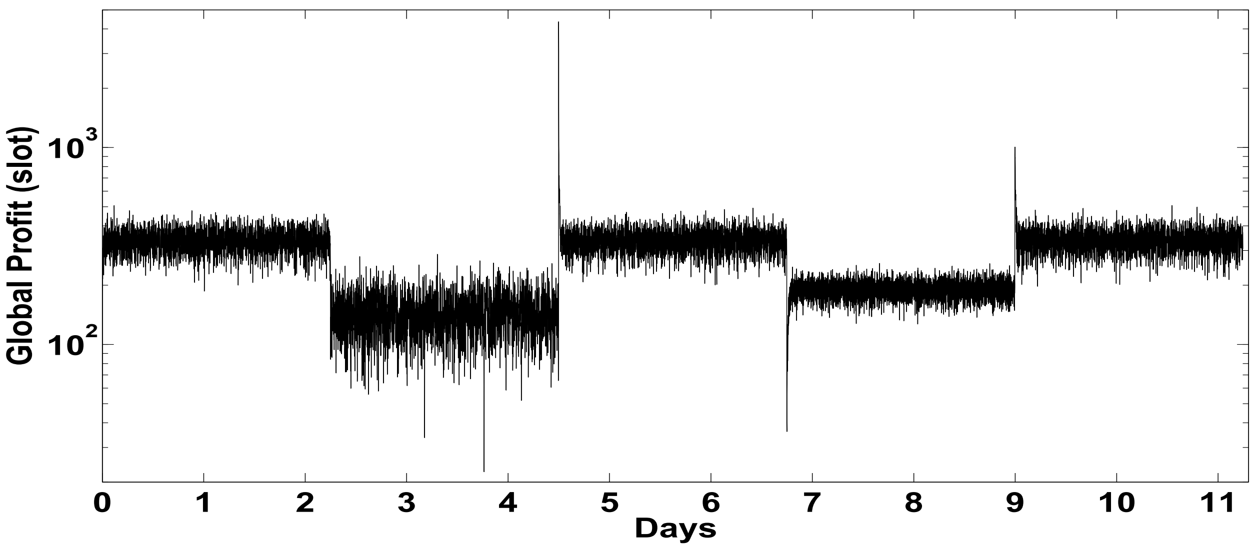


Figure 4.9: Impact of Dynamic environment on System Profits.

between all the nodes in the network, so that nodes can only transmit data packets through cellular communication. Finally, the network returns again to its normal state in the fifth period.

Here we used global system costs and global social profits in every slot (both in credits) to measure the effect of changing topology over time. Here the global system cost is measured as the sum of both the transmission and reception costs of all phones in a slot. Global social profits is computed as the total value of all the successfully received packets (by the MPSS server) minus the total system cost in a slot.

In Fig. 4.8, we can see that in the second period, the global system cost decreases when cellular communication is disabled. This is due to all the transmissions being relayed through Wifi direct only, which is cheaper than cellular communication. However, Fig. 4.9 show that global phone user profits also decrease in spite of the decrease in cost. This is caused by the large delay in multi-hop transmission which results in increased number of dropped packets with smaller TTL. In the fourth period, system costs increase significantly when Wifi direct communication is disabled due to the high cost of cellular communication. This is also reflected by the decrease in global profits of the MPSS. We can also see that the network self-adjusts very quickly to the changing conditions of the network. When the network returns to normal operation in third and fifth period, the data buffer of the phones contain large numbers of data packets that are sent instantly after availably of alternate option. This is the reason behind sudden spikes in system cost and global profits at the start of these periods. Once the backlog reduces, the system becomes stable.

4.5.5 Self-Adaptation to dynamic pricing models and user behaviours

To evaluate the impact of customer's actual mobile phone usage on system; we categorised users according to their usage behaviours against the available tariffs in the market. The pricing models for phone users can be divided in to following three categories with respect to data.

1. **Pay as you go(PAYG)/ Prepaid.**

   PAYG is the simplest pricing model in which users pay fixed cost per unit for data usage. PAYG subscribers are usually charged at higher rate for data usage than users who are subscribed to monthly data packages/allowances.

2. **Monthly package with limited data allowance.**

   Majority of the pricing plans for mobile networks include monthly data allowance (*e.g.* 2GB per month) for additional price at the start of each month. Per unit cost of the data allowances is less than cost of data usage for PAYG users. If a user consumes data allowance before the completion of monthly cycle, data usage is charged at a flat cost per unit for the rest of the monthly cycle. This flat cost is considerably higher and is similar to that of PAYG users.

3. **Monthly package with Unlimited data allowance.**

   Few pricing plans include unlimited data allowance for the users at a higher monthly price than the plans with limited data allowance. The users can consume data through out the monthly cycle at no extra cost.

According to the analysis of on-line bills of users to find out exactly how they use their phone [10], the number of mobile phone users paying for mobile data usage outside of allowance has declined at an accelerated rate. In contrast with usage of tariff-inclusive data allownces; users appear highly aware of and averse to using chargeable services out of allowance. Based upon the user's behaviour of mobile data usage, we divided users in to two groups.

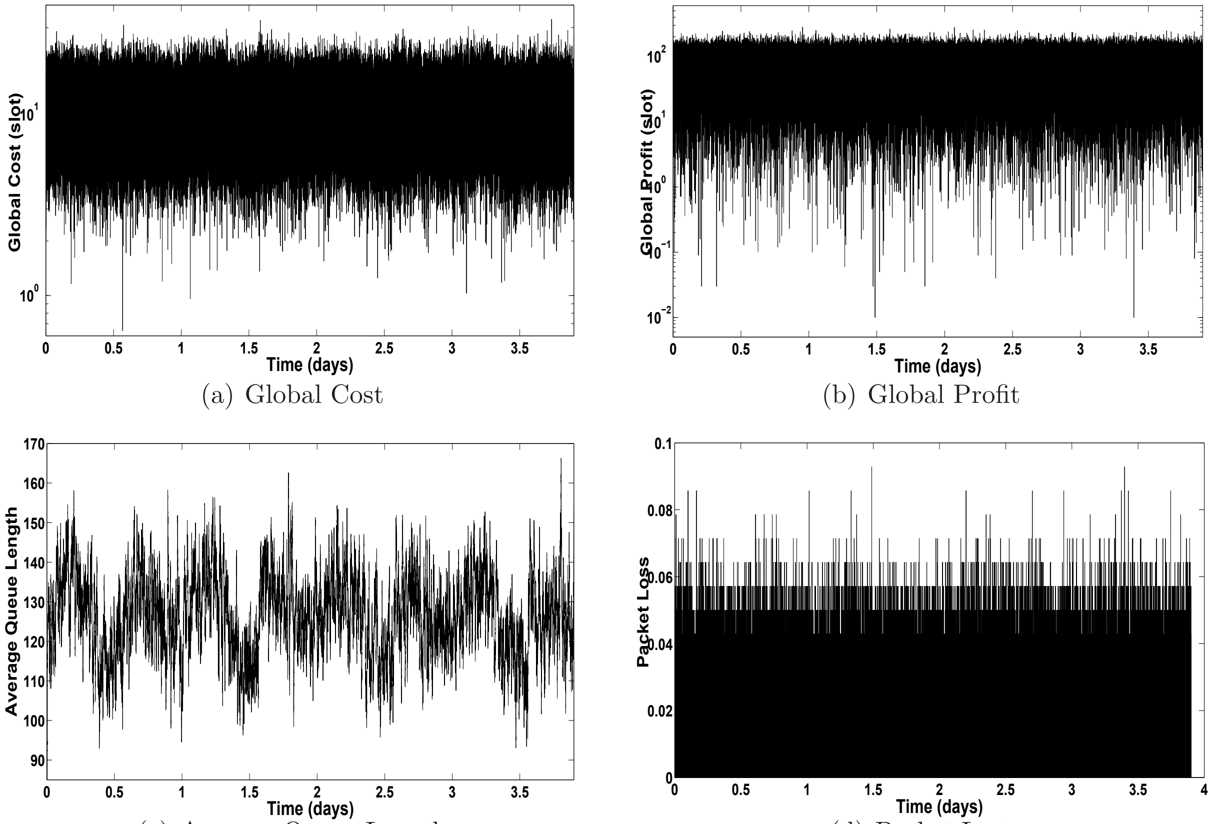
1. **Opportunistic users.** The users who are willing to allow MPSS applications to use data at any cost as long as they will earn profits in return.

2. **Conservative users.** The users who do not want to consume data out of their data allowance. Their usage of data decreases with the decrease in remaining data allowance. Therefore they decrease the data allocation for the MPSS applications with the decrease in their monthly data allowance.

We investigated the effect of user behaviour on the system by incorporating different user's preferences. For PAYG users, cellular communication cost is set between 5 - 10 credits per packet and it remains fixed over time. For users with the limited and unlimited data allowance, cellular cost was set between 1 - 5 credits per packet because these plans are available at less data cost than PAYG plans. For users with a limited data allowance, we allocated 500 mb - 5GB at the start of each month according to typical data plans available in the market. Remaining data allowance is decreased randomly each day from 10mb - 200mb. If a user run out of allowances, the cellular data cost increases to 5-10 credits per packet and it stays constant until the end of the month, similar to PAYG plans. For conservative users, we decreased allocation of data for MPSS applications linearly with respect to remaining data allowance. On the other hand, opportunistic users allow MPSS applications to use the data only based on cost benefit analysis.

We evaluated the capability of our approach to self-adapt the network as a consequence of dynamic change in cost. We constructed a scenario where 100 mobile users are subscribed to monthly package and 40 are PAYG users. Each cycle is of 30 days for monthly package subscribers. We set slot size to 10 seconds and each simulation lasts for 10^6 slots (110 days).

The monthly cycle of limited or unlimited data package subscribers does not start at the same day of the month. Network service providers assign different starting dates to each mobile phone user as per their policy. To observe the effect of dynamic pricing on global network performance, we constructed a scenario where each monthly data subscriber has been assigned random day of the month as starting day on their monthly cycle. Although dynamic pricing has a significant impact on single individual but we can see in Fig. 4.10(a) that the system adapts itself and remains unaffected regardless of change in the data price of monthly subscribers.Mobile users choose to forward data to users who have low cellular cost i.e they are in the start of their monthly cycle. Due to different starting time of cycles, change in the the cost of individual mobile phone user does not affect the performance of the global profit as shown in Fig. 4.10(b). If the cellular cost of a user increases as the result of montly data allowance consumption, cost benifit analysis will deter it to send MPSS data packets through cellular communication. This will increase its backlog and other user will not choose it as as a relay for MPSS data. The



(a) Global Cost (b) Global Profit

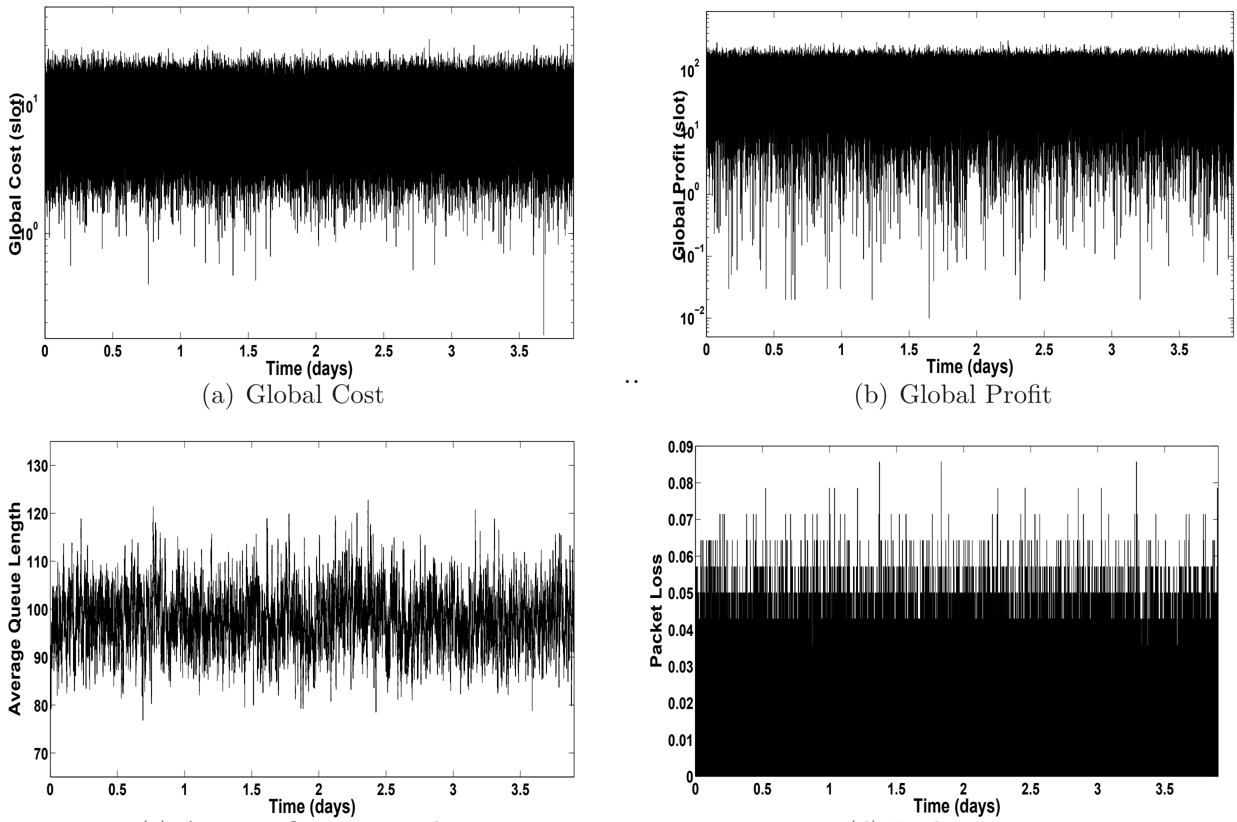
(c) Average Queue Length (d) Packet Lost

Figure 4.10: Global Impact of pricing models and user behaviour

stability of the system can also be verified by queue lengths and packet loss shown in Fig. 4.10(c) and 4.10(d).

4.5.6 Self-Healing

In last set of simulations, we study the self-healing capability of our scheme. In particular, we analysed the critical situation when a substantial number of the available mobile phone users leave the network. We constructed a scenario where 40 phone users leave the network during second month at random time and join back during third month. Fig. 4.11. shows the effect on the network performance due to nodes leaving and joining back in the system. In Fig. 4.11(a) and Fig. 4.11(b). we can observe that this does not affect global cost and global profit of the network. The network self-heals itself when nodes leave the network as the total amount of data also decreases and mobile phones wait until they find suitable cost-effective way to forward the data. This is also validated through stability of average queue length and low packet loss as



(a) Global Cost

(b) Global Profit

(c) Average Queue Length

(d) Packet Lost

Figure 4.11: Impact of mobile phone users leaving and joining the network

shown by the Fig. 4.11(c) and Fig. 4.11(d). Furthermore the network is also able to reconfigure itself when the users join back during the third month and remains stable.

4.5.7 OppCom App

We developed an application (OppCom App) for android to evaluate our approach in real world. We also deployed a server to collect data from the phones running OppCom App shown in Fig. 4.12. OppCom uses Wifi-Direct for short-range communication whereas it can also send data through cellular network (3G) to the MPSS server.

For OppCom App, slot interval was set to 10 seconds to accomodate discovery delay in Wifi-Direct and it senses location data every 30 seconds i.e. 1 packet in every 3 slots. The communication cost using Wifi Direct was set as 1 credit per packet whereas 3G transmission cost was kept random between 1 to 10. We set WiFi Direct Channel capacity to 100 packets/slot

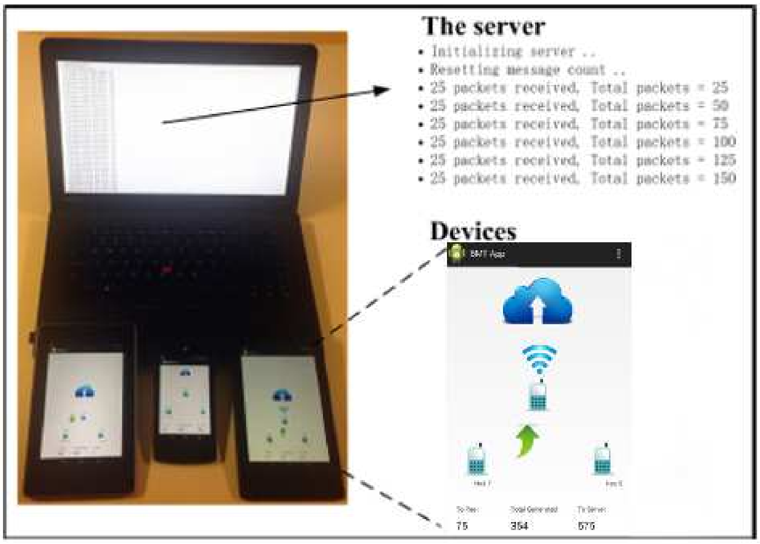


Figure 4.12: OppCom App.

and 3G/WiFi channel capacity to 150 packets/slot. We installed OppCom App on 10 mobile devices including Nexus 5, Nexus 4, Samsung Galaxy S5, HTC M8 and Nexus 7. Two of the devices have no 3G connection and they can only use Wifi Direct to communicate with other devices. The devices were distributed to students at Imperial College, London. Steps were taken to make sure OppCom App was deployed to a device that would be carried around as opposed to being kept stationary. Also, users were chosen to ensure that a device would come into contact with at least another device at a point during the day. However, the deployment was kept sparse enough so that devices were not always in each other's transmission range. We conducted the experiment for approximately 3 days.

Table 4.3 and 4.4 contains the statistics that we gathered from the clients and the server. Table 4.3 provides various statistics from user's devices.

Individual Profits

We can see that every device obtained positive profit in the experiment, which shows that our scheme can incentivize phone users to participate in the MPSS by providing them fair reward. The variation of individual profit depends mainly on the extent of their participation in data relaying for other devices. Individual profits also depend on 3G cost and access to Wifi routers of the users. The devices E, G, H and J have used Wifi Direct communication for relaying of

Device	Total Sensed	From Peer	To Peer	To Server	Local Cost	Profit
	(packets)				(credits)	
A	866	0	602	0	1468	144.32
B	2399	10	266	0	2665	343.85
C	8282	49	10	6867	56361	836.53
D	6844	48	131	6199	35770	695.8
E	6343	2562	1435	7124	38514	766.45
F	5124	0	0	3618	26712	515.99
G	9141	2636	6663	4919	40399	1781.28
H	9215	8512	1920	14720	90517	209.13
I	42	57	80	0	122	5.62
J	6086	666	3433	1703	17431	1800.57
Total	54342	14540	14540	45150	309959	7099.54

Table 4.3: Statistics gathered from mobile devices

Number of Packets	45150
Average Delay	1988.92 seconds $\approx 30mins$
Average Hops	1.308
3G Packet Loss	0.0 (0%)
Wifi Direct Packet Loss	0.0 (0%)

Table 4.4: Statistics gathered from MPSS server

data and earned larger profits than others. It is interesting to note that although device H sent more packets using Wifi Direct than device E but device H earned lower profits than device E. This is due to tha fact device H more packets than device E and only sent a small portion through Wifi Direct. Therefore, the local cost of device H is high and its profit is low. The devices C and F also earned high profits although they relied majorly on 3G communication. This is due to their lower 3G cost than other devices.

Packet Received by Server

From the table 4.3 we can see that 83% of the packets sensed are delivered to the server. The remaining packets are either in the message buffer of the devices or were lost when the devices re-started the application. We can also see that a couple of devices participated much less that other devices. A limitation on WiFi Direct in android was that it does not automatically accept WiFi Direct connection requests. When two new devices try to connect for the first time, android pops-up a message on screen prompting for user verification to allow the devices to connect. This deterred some users from actively participating in the MPSS but this did not affect the performance of the system, which verified our discussion of self-* properties of the approach.

Delay

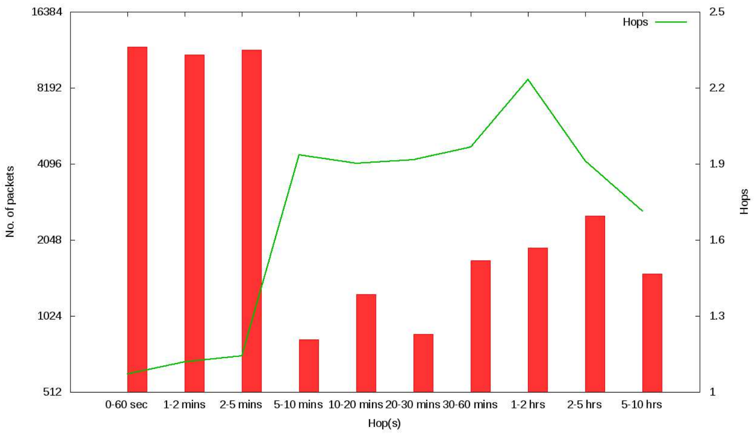


Figure 4.13: Packet delay.

In table 4.4 we can see that the average delay reported by the server was around 30 minutes. This delay is due to sparsity of the network and devices were isolated during certain times of the day (i.e. night time) Also, because two of the devices cannot communicate directly with the server, they wait to find other peer to send their packets.

Fig. 4.13 shows the package delay from all the packets received by the server. We have distributed the packets according to different time intervals. We observed that more than half of the packets were received within 5 minutes. We also observe that the average hops is close to 1 for the packets received within five minutes. This is due to the devices, which can directly connect to internet, therefore they were able to upload packets as soon as they are generated. There were few packets with smaller delay that were received by the server with more than one hop count. This is due to the devices which are in contact with other devices with internet access at a given time, a device with no internet access can forward its packets to other devices and these packets are sent to the server with small delay. We can see that with the increase the delay, the average hop increases significantly. These packets with large delays are from the devices with no or high cost internet access and no peers with internet access in their communication range. These devices send packets to each other, and finally send the packets to a peer with internet access upon encounter.

Packet Loss

During the experiment, there were no packet loss on 3G and WiFi Direct communication network. This shows that the communication technology is consistent and reliable in terms of transmission quality and speed.

4.6 Conclusion

In this chapter, we developed a citizen-centric network architecture to provide a cost-effective networking service for real-time and delay tolerant applications in Mobile Phone Sensing Systems (MPSS). Based on the guiding principles of the proposed network architecture, we developed a joint pricing and routing scheme to support both real-time and delay-tolerant MPSS applications through seamless integration of cellular and short-range communications of mobile phones. The proposed scheme is not only lightweight and fully distributed, but can also achieve optimal throughput, which is highly suitable to deliver large amount of mobile sensing data. Moreover, it is autonomous and scalable in highly dynamic environment of large cities. Through simulations, we demonstrate that our scheme can minimize global system costs, as well as effectively incentivize phone users to participate in the MPSS. We also show that our scheme is adaptive to user's preferences, privacy and self configures, self adapts and self heals in dynamic network conditions. Finally we developed an android appliaction (OppCom App), which shows the applicability of our scheme in real world deployments.

Although our network architecture addresses key design principles of MPSS networking services but it does not consider the mobility characteristics of the network. Unlike traditional wireless sensor networks, mobile sensing brings people in the loop. Therefore, the question remains; How to exploit the underlying social networks of human relays to utilize their mobility patterns for an efficient networking system.

In the next chapter, we address this problem, where we exploit underlying social networks structure of people to design joint rate control, routing, and resource pricing scheme for a hybrid sensor and mobile phone network consisting of static sensors, mobile phones and static sinks.

Chapter 5

Socio-Economic aware Opportunistic Data forwarding in WSN with Human Relays

Fixed infrastructures of wireless sensing techniques have limitations in terms of sensor maintenance, placement and connectivity in future smart cities. One approach to overcome some of these problems is to utilise the ubiquity of mobile phones as discussed in chapter 4. Unlike traditional wireless sensor networks, mobile sensing brings people in the loop, not only as the owners of sensing devices but also as the sources and consumer of the network. WSNs with Mobile Phones (WSN-MP) can utilize the interactions between static sensors, mobile sensors (people with mobile phones) and sinks with internet access to enable cost effective data collection for delay tolerant networks. This has encouraged researchers to study the social characteristics of people in order to design efficient data forwarding approaches in these systems.

In the chapter we build on economic-aware scheme for MPSSs described in chapter 4; and exploit underlying social networks of human relays to design a socio-economic aware data forwarding scheme. We propose a novel data forwarding metric, Sink-Aware (SA) centrality, to measure the potential sensor data forwarding ability of mobile relays. Our work combines network science principles and Lyapunov optimisation techniques to maximise global social profit across hybrid sensor and mobile phone networks. Sensor data packets are produced and traded (transmitted) over a virtual economic network using a lightweight, social-economic-aware backpressure algorithm which combines rate control, routing, and resource pricing. Phone owners can receive benefits by relaying sensor data. Our algorithm is fully distributed and makes no probabilistic/stochastic assumptions regarding mobility, topology, or channel conditions, nor does it require prediction. Simulation results further demonstrate that the proposed algorithm outperforms pure backpressure and social-aware schemes, highlighting the advantage of building systems that combine communication with other types of networks.

Different than mobility aware scheme for WSN with mobile sink (WSN-MS) in chapter 3, where we considered mobility of nodes to collect data from static WSN; this chapter considers an WSN-MP, where mobile devices (i.e. smart phones) are used to collect and forward data from static sensor nodes to the sink (Wifi Access Point or cellular base station) using opportunistic short-range wireless communications. We refer such WSN-MP to as WSN with Human Relays (WSN-HR) in this chapter. Furthermore, main focus of this chapter is on how the social and economic behaviours of the phone owners can be utilised for efficient data transmission.

$\mathcal{S}$	The set of all sensor nodes.
$\mathcal{R}$	The set of all human relays.
$\mathcal{D}$	The set of all sinks.
$\mathcal{N}$	The set of all nodes $\mathcal{N} = \mathcal{S} \cup \mathcal{R} \cup \mathcal{D}$.
$\mathcal{L}$	The set of all directed links between each pair of nodes in $\mathcal{N}$.
$c_{x,y}(t)$	The capacity of wireless link $(x, y) \in \mathcal{L}$ at slot t.
$\boldsymbol{c}(t)$	Channel capacity vector for all wireless links at slot t.
$G(\mathcal{N}, \mathcal{L}, \boldsymbol{c}(t))$	The time-varying weighted graph of the WSN-HR.
$\mathcal{L}^{wsn}$	The set of all directed links between each pair of sensor nodes in $\mathcal{S}$.
$G(\mathcal{S} \cup \mathcal{D}, \mathcal{L}^{wsn})$	The graph of the static-deployed WSN.
$\mathcal{N}_x(t)$	The set of node x's instantaneous neighbours at slot t.
$r_x(t)$	The sensing rate of sensor node $x \in \mathcal{S}$ at slot t.
$f_{x,y}(t)$	Actual amount of data transmitted over wireless link $(x, y) \in \mathcal{L}$ at slot t.
$Q_x(t)$	The queue backlog of node $x \in \mathcal{N}$ at slot t.
$ICT_{x,y}$	The inter contact time between nodes x and y.
$\mathcal{L}^{social}$	The set of social ties between human relays in $\mathcal{R}$.
$G(\mathcal{R}, \mathcal{L}^{social})$	Social graph of human relays.
$G(\mathcal{N}_x^{social})$	Social neighbour table of human relay $x \in \mathcal{R}$.
$\mathbb{C}_x$	The set of community(ies) that mobile relay $x \in \mathcal{R}$ belongs to.
$\boldsymbol{h}_x$	Local centrality vector for mobile relay $x \in \mathcal{R}$.
$\mathbb{C}_x^{static}$	The set of community(ies), whose location static node x located in.
$\mathbb{A}$	The set of all static clusters.
$\mathbb{A}^{sink}$	The set of all static clusters containing sinks.
H_x^{sink}	The sink-aware centrality of a mobile relay $x \in \mathcal{R}$.
$\lambda_x(t)$	The selling price if node $x \in \mathcal{S} \cup \mathcal{R}$ in slot t.
$I_x(r_x(t))$	Utility function of sensor node x.
$\gamma_x^{relay}(t)$	The instantaneous profit of a mobile relay $x \in \mathcal{R}$.
$\Gamma^{WSN}(t)$	The sum of the instantaneous profits of all nodes in the static WSN.
$\Gamma(t)$	The sum of the instantaneous profits of all nodes in WSN-HR.
V	The parameter to tradeoff global social profits and queue backlogs.
$\mathrm{r}^{\max}$, $\mathrm{c}^{\max}$	The finite upper bounds of sensing rate and channel capacity respectively.
Q_x^{max}	The finite queue buffer size of node $x \in \mathcal{N}$.
t_{end}	The number of slots of the finite horizon.
$\mathrm{H}_{\mathrm{sink}}^{\max}$	The maximum SA centrality.
λ_{scale}	The price-scaling parameter.
α	The weighting parameter for social awareness.
$w_{x,y}(t)$	The routing weight of wireless link (x, y) at slot t.
$\eta_x(t)$	The maximal possible amount of data can be received by node x at slot t.
$\eta^{\max}$	The finite upper bound of $\eta_x(t)$ for all $x \in \mathcal{N}$ at slot $1 \le t \le t_{end}$.

Table 5.1: Summary of symbols used in Chapter 5.

WSN-HR	Wireless Sensor Network with Human Relays
OBSEA	Opportunistic Backpressure with Social/Economic Awareness
SA Centrality	Sink Aware Centrality
ICT	Information and Communication Technology

Table 5.2: Summary of abbreviations used in Chapter 5.

5.1 Introduction

In this chapter, we apply network science principles to build a resilient architecture consisting of a hybrid of mobile phones and WSNs. Here static sensors are deployed to instrument a space and report sensed readings. However, we deviate from the traditional WSN architecture by not only using static base-stations connected to the Internet to relay data about the space, but also utilising human relays via their mobile devices. This resilient architecture is motivated by sensing applications in sustainable smart cities [46, 148, 151]. For brevity we call our architecture WSN-HR (Wireless sensor networks with Human Relays). For a WSN-HR to be a cost-effective communication solution for smart sustainable cities, the following two key issues must be addressed:

- Since the mobility patterns of human relays are governed by their underlying social networks [121, 199], *how can we exploit social network features, such as centrality and community, for efficient sensor data forwarding?*
- Since using the mobile phone as a relay has costs, in terms of local resources (e.g. memory and energy) and telecommunications, *how can we incentivise individuals to participate?*

5.1.1 Our Approach

To address the above issues, we develop a novel scheme: an Opportunistic Backpressure approach with Social/Economic Awareness (termed OBSEA) for joint rate control, flow routing, and resource pricing. This approach uniquely combines network science principles and Lyapunov stochastic optimisation theory [43,69,203]. Specifically, the contributions of this chapter are summarised as follows:

1. By exploiting mobility patterns and the underlying social networks of human relays, we propose a novel data forwarding metric, Sink-Aware (SA) centrality, to measure the potential sensor data forwarding ability of mobile relays.

2. To incentivise people to serve as data relays using their phone, we establish a virtual economic network for sensor data producing and trading. Here, the static WSN consists of static sensors and sinks. It makes a profit by producing sensor data to maximise the network utility. Each mobile relay acquires profit by dynamically adjusting the selling price of its maintained sensor data and then trading (transmitting and receiving) data with other nodes opportunistically at each moment of contact.

3. We formalise a finite-horizon optimisation problem to maximise the global social profits of the all nodes in WSN-HR. Our formalisation *does not* make any probabilistic/stochastic assumptions (e.g. specific probability distributions or ergodic stochastic network processes) about the network conditions (e.g. mobility, topology, and wireless channel) and thus is suitable for the arbitrary dynamic evolution process of WSN-HR. The lightweight OBSEA solves the problem using only current and local information. This means that OBSEA is *fully distributed* and *does not require* any prediction capacity, therefore maximising the practical application of the work.

4. We evaluate the performance of OBSEA using the Castalia [3] simulator and a realistic mobility model, [199] which exhibits features observed from real social networks and human mobility traces. Simulation results demonstrate that OBSEA is adaptive to different network settings. The algorithm is also shown to outperform pure backpressure routing and pure social-aware forwarding schemes in terms of global social profit, data buffer efficiencies, and end-to-end delay. In addition, the results show that a 'win-win situation' (positive outcome) can be achieved by both the static WSN and all mobile phone owners.

5.1.2 Related Work

Data Muling and Relaying for WSNs. Many data muling schemes have been proposed to improve energy efficiency and coverage in sparse WSNs. Specifically, [31, 55, 81, 151, 183] recognise the potential of human mobility in data mules. As far as we are aware, however, none of them consider exploiting the underlying social networks of human relays or utilising opportunistic multi-hop human contacts.

Intermittently Connected Networks and Social Awareness. Social network metrics, centrality and community structure, have been used for many opportunistic routing schemes [45,66,67,84,121] in Delay Tolerant Networks (DTN) [60]. However, all of them focus on packet routing (i.e. unicasting or multicasting a single packet or multiple packets) rather than the flow routing in our OBSEA.

For general intermittently connected networks [41, 56, 159], backpressure-type flow routing and control schemes are studied in [158, 159]. However, their scheme is based on the assumptions of predetermined gateways and ergodic network conditions (i.e. mobility and channel states). In contrast, our work uses a much more general and realistic network model in which any sensor could serve as a gateway at each opportunistic contact, and no probabilistic/stochastic assumptions (e.g. Markov process of the mobility) are made for arbitrary network conditions. Furthermore, none of them considers social or economic awareness.

Network Optimisation. Cross-layer network optimisation and control is an active networking research area [43, 69, 203]. Most backpressure scheduling/routing schemes [111, 145, 206] are developed for multi-hop wireless networks to achieve infinite-horizon stability, and to maximise long-term network utilities or minimises costs. However, these schemes are limited to ergodic network models, which may not hold true in our highly dynamic WSN-HR. The recent universal scheduling framework developed by Neely [142] optimises finite-horizon general network utility with arbitrary dynamic network processes, and has been used in P2P networks [143] and smart electricity markets [144]. OBSEA is the first approach to combine the universal scheduling framework [142] and network science principles for urban WSNs using mobile data relays.

Pricing and Incentive Schemes. In [103, 141, 143], Lagrange multipliers or queue backlogs are used as prices to solve static convex network problems or dynamic stochastic problems. In contrast, OBSEA uses both queue backlog and social-aware metrics for pricing. [132, 147, 157] study game-theoretic incentive and pricing approaches. Incentive-aware routing schemes [129, 168] are proposed for data forwarding in DTNs. DTN routing schemes [56, 122] consider the concept of *social selfishness*, which describes the willingness of an individual to provide better service to those with strong social ties than those with weaker social ties. However, *rational selfishness* considered by our OBSEA means that each phone owner is willing to relay sensor data as long as he or she can get benefits, which is different form the concept of *social selfishness*. In addition, none of above schemes focus on data muling for WSNs.

5.1.3 Chapter Organisation

In the next section, the network model is presented. Section 5.3 presents the OBSEA algorithm. Simulation and results are presented in Section 5.4. Finally, we summarise this chapter in Section 5.5. All proofs of theorems in this chapter and related lemmas regarding theoretical analysis can be found in Appendix B.

5.2 Network Model

In an intermittently connected WSN-HR, every sensor node collects environmental data (e.g. temperature and humidity) and sends the sensor data to any of the sink(s) through other static sensor nodes and mobile relays if necessary, in an opportunistic multi-hop manner.

5.2.1 Topology and Communication model

Let the sets of sensor nodes, human relays, and sinks be $\mathcal{S}$, $\mathcal{R}$, and $\mathcal{D}$ respectively. Denote $\mathcal{N} = \mathcal{S} \cup \mathcal{R} \cup \mathcal{D}$ as the set of all nodes in the WSN-HRs. The network operates in discrete time with a unit time slot $t \in \{1,\ 2,\ ...\}$. Let $c_{x,y}(t) \geq 0$ be the current capacity of wireless link from node $x \in \mathcal{N}$ to node $y \in \mathcal{N}$ at time t, i.e. the maximum (integer) number of sensor data packets that can be successfully transmitted from x to y during slot t. $c_{x,y}(t)$ is assumed to be constant within the duration of a slot, but can vary from slot to slot and across different wireless links. Specifically, if $c_{x,y}(t) > 0$, we say nodes x and y are ***in contact*** at slot t; otherwise, they are not in contact at slot t.

We model the whole WSN-HR as a directed, complete, and time-varying weighted graph $G(\mathcal{N}, \mathcal{L}, \boldsymbol{c}(t))$, where $\mathcal{L} = \{(x, y)|\ x, y \in \mathcal{N}\}$ represents the set of all possible wireless links between each pair of nodes in $\mathcal{N}$, and the $|\mathcal{L}|-$dimensional vector $\boldsymbol{c}(t)$ represents the vector of channel capacities over all wireless links at slot t. Due to the sparsity and intermittent connectivity of the network, most entries of instantaneous $\boldsymbol{c}(t)$ are zero at a given t. Figure 5.1(b) illustrates an example of instantaneous $G(\mathcal{N}, \mathcal{L}, \boldsymbol{c}(t))$ at a slot.

We *do not* make any probabilistic/stochastic assumption on $\boldsymbol{c}(t)$, such as specific probability distribution, i.i.d., or even ergodicity. This is because the stochastic process $\boldsymbol{c}(t)$ could be affected by many random time-varying events such as unexpected external interference, channel fading, and human mobility, governed by various complex physical rules. It is easy to see that the definition of $G(\mathcal{N}, \mathcal{L}, \boldsymbol{c}(t))$ is very general and can characterise arbitrary stochastic channel states and topology processes (e.g. mobility) of the $|\mathcal{N}|$-node WSN-HR.

Due to the sparse density of the WSN-HR, we assume that the wireless interference among concurrent transmissions over links in $\mathcal{L}$ can be ignored, in order to focus on routing and rate control[1].

For notation brevity, we also define $G^{wsn}(\mathcal{S} \cup \mathcal{D}, \mathcal{L}^{wsn})$ to represent the static WSN, where $\mathcal{L}^{wsn} = \{(x, y)|c_{x,y}(t) > 0\ \forall t \geq 1,\ \forall x, y \in \mathcal{S} \cup D\}$ is the set of all wireless links with non-zero capacities between static nodes.

Definition 5.1 (Static Cluster)**.** *A static connected cluster $\mathcal{A} \subseteq \mathcal{S} \cup \mathcal{D}$ is defined as the set of all static nodes in a connected component[2] of the static WSN $G^{wsn}(\mathcal{S} \cup \mathcal{D},\ \mathcal{L}^{wsn})$.*

For instance, there are four static clusters in the WSN shown in Figure 5.1 (a).

[1] It is easy to add greedy/approximate scheduling functionalities such as [42] into our OBSEA algorithm for efficient distributed implementations.

[2] In graph theory, a connected component of a graph G is a subgraph in which any two vertices are connected to each other by paths, and which is connected to no additional vertices in G.

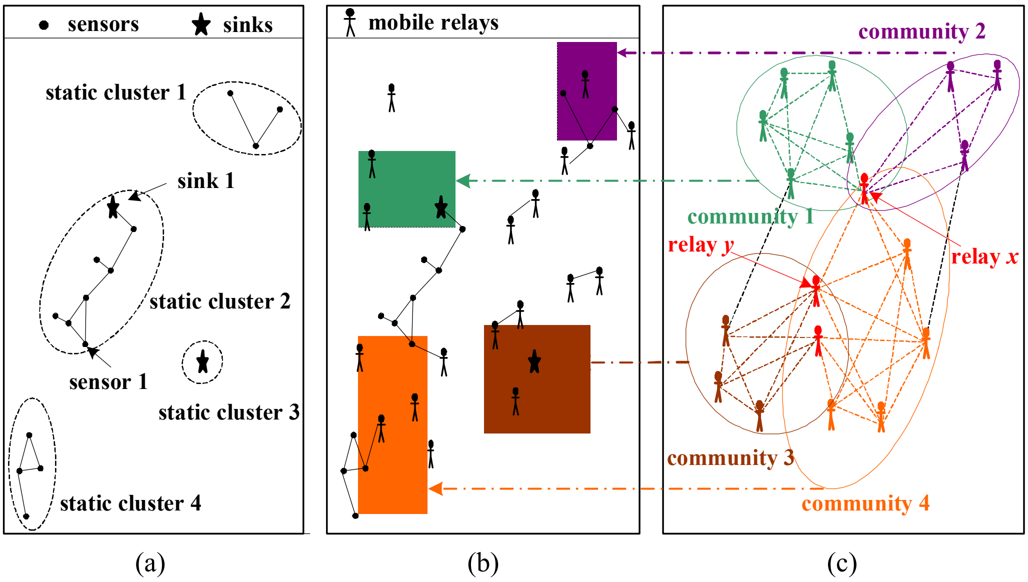


Figure 5.1: Conceptual illustration of a WSN-HR. (a) Four static clusters of a static disconnected WSN $G^{wsn}(\mathcal{S} \cup \mathcal{D}, \mathcal{L}^{wsn})$. (b) An example of instantaneous graph $G(\mathcal{N}, \mathcal{L}, \boldsymbol{c}(t))$ at a slot, and community-location mapping. Solid lines between nodes represent the wireless links with non-zero capacities, and the wireless links with zero capacity are not plotted. The colourful rectangles indicate the geographic areas associated with the communities. (c) The underlying social network consisting of mobile relays, and an example of 4-clique overlapping community structure over the social network.

5.2.2 Queueing Dynamics

Each node $x \in \mathcal{N}$ maintains a queue (i.e. data buffer) for the sensor data, which stores the data packets generated by itself (if x is a sensor node), and by other sensor nodes. Let $Q_x(t) \geq 0$ be the queue backlog (or queue length) of $x \in \mathcal{N}$ at slot $t \geq 1$. Let $\mathcal{N}_x(t) \subset \mathcal{N}$ be the set of nodes that are in contact with node x at slot t (i.e the set of x' instantaneous neighbours),

$$\mathcal{N}_x(t) = \{\ y \mid c_{x,y}(t) > 0,\ c_{y,x}(t) > 0,\ y \in \mathcal{N} - \{x\}\ \}$$

From each node $x \in \mathcal{N}$, its queue backlog updates from slot t to $t+1$ as follows:

$$Q_x(t+1) = \begin{cases} 0 & x \in \mathcal{D} \\ |Q_x(t) - \sum_{y \in \mathcal{N}_x(t)} f_{x,y}(t) + \sum_{y \in \mathcal{N}_x(t)} f_{y,x}(t)|_+ & x \in \mathcal{R} \\ |Q_x(t) - \sum_{y \in \mathcal{N}_x(t)} f_{x,y}(t) + r_x(t) + \sum_{y \in \mathcal{N}_x(t)} f_{y,x}(t)|_+ & x \in \mathcal{S} \end{cases} \tag{5.1}$$

where $r_x(t) \geq 0$ is the sensing rate at which a sensor node $x \in \mathcal{S}$ collects environmental data at slot t; $0 \leq f_{x,y}(t) \leq c_{x,y}(t)$ represents the actual amount of data transmitted from node x to node y at slot t; and for any real number a, the operator $|a|_+ = a$ if $a > 0$, $|a|_+ = 0$ otherwise.

5.2.3 Mobility Pattern and Social Network of Human Relays

The following human mobility and social network properties are explored by our OBSEA algorithm.

Pairwise Inter-Contact Time. Let $ICT_{x,y}$ be the inter-contact time (i.e. the time elapsed between two successive contacts as shown in Figure 5.2) between a pair of relays $x,\ y \in \mathcal{R}$. The distribution of pair wise inter-contact times between mobile relays has great impact on data forwarding [38]. We do not assume any special distribution of $ICT_{x,y},\ x, y \in \mathcal{R}$ (e.g. power-law [38], exponential [67], or power-law head and exponential tail [102]). In addition, we generalise the concept of inter-contact time from each pair of pure mobile relays in $\mathcal{R}$ to that of all nodes in $\mathcal{N}$. Specifically, if two static nodes $x, y \in \mathcal{S} \cup \mathcal{D}$ are always in contact (i.e. $c_{x,y}(t) > 0$ for all $t \geq 1$), $ICT_{x,y} = 0$; otherwise (i.e. $c_{x,y}(t) = 0$ for all $t \geq 1$)), $ICT_{x,y} = \infty$.

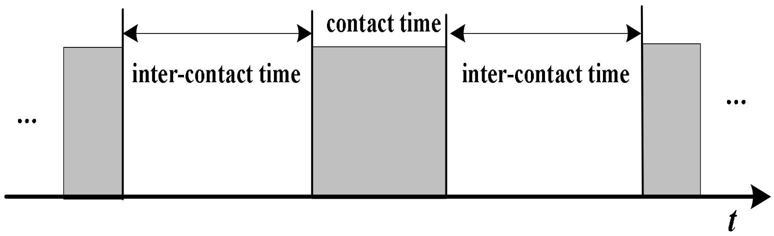


Figure 5.2: The concept of contact time and inter-contact time between a pair of nodes.

Overlapping Communities and Centrality. As shown in Figure 5.1 (c), we assume that there is an underlying social network that consists of all mobile relays in $\mathcal{R}$. We model the social network as an undirected graph $G^{social}(\mathcal{R},\ \mathcal{L}^{social})$, where $\mathcal{L}^{social}$ represents the set of social ties between mobile relays, which can be defined by inter-contact time or contact probability between each pair of nodes. A social tie between two mobile relays x and y in $\mathcal{R}$ is considered to exist, if $ICT_{x,y}$ is smaller than a predetermined threshold $\mathrm{ICT}^{\max}$. Each mobile relay $x \in \mathcal{R}$ maintains and updates its social neighbour table $\mathcal{N}_x^{social} \subseteq \mathcal{R}$, which is the set of mobile relays that share social ties with x, i.e. for each relay $y \in \mathcal{N}_x^{social}$, $ICT_{x,y} \leq \mathrm{ICT}^{\max}$. Using social neighbour table of each relay, the social network $G^{social}(\mathcal{R}, \ \mathcal{L}^{social})$ can be established in a distributed way at runtime.

It has been observed that a social network always exhibits overlapping community structures [86, 150, 199] and heterogeneous centrality [24, 66, 84]. In the WSN-HR context, *overlapping community structure* means that mobile relays in the same community (a set of mobile relays) meet each other much more frequently than that in different communities, and a mobile relay may belong to multiple communities. *Heterogeneous centrality* indicates that few mobile relays (e.g. postmen) meet a large number of other relays, but many mobile relays only meet a small number of others. We explore these two useful social network features in our OBSEA algorithm.

Mathematically, we use a tuple $(\mathbb{C}_x, \boldsymbol{h}_x)$ to represent the social profile of a mobile relay $x \in \mathcal{R}$, where $\mathbb{C}_x$ is the set of community(ies) node x belongs to, and $\boldsymbol{h}_x$ is a $|\mathbb{C}_x|$–dimensional vector, where each entry h_x^i, $1 \leq i \leq |\mathbb{C}_x|$ represents the local centrality [66, 199] of x in its i^{th} community $\mathcal{C}_x^i \in \mathbb{C}_x$. Specifically, the h_x^i is measured as the number of social ties between x and other mobile relays in $\mathcal{C}_x^i \in \mathbb{C}_x$, i.e. $h_x^i = |\mathcal{C}_x^i \cap \mathcal{N}_x^{social}|$. For instance, in Figure 5.1 (c), the local centralities of relays x and y in community 4 are 3 and 4 respectively. It can also be seen that for k-clique community structure of social network $G^{social}(\mathcal{R}, \ \mathcal{L}^{social})$, the local centrality of every mobile relay in every community should be no less than $k - 1$ [150].

Spatial Regularity of Human Mobility. Recent observations [74, 80] of real human traces demonstrate that people in a given community normally move within some certain geographic areas much more frequently than other locations (e.g. students in the same department normally visit their department building with a much higher probability than other places). An example of such community-location mappings are illustrated in Figure 5.1 (b) and (c). Based on this property, we establish social awareness for each static node $x \in \mathcal{S} \cup D$, by using a variable $\mathbb{C}_x^{static}$, which represents the set of community(ies) associated with the geographic area where x is located in. Take Figure 5.1 for instance, since sink 1 sensor 1 are located within the area associated with communities 1 and 4 respectively, $\mathbb{C}_{\text{sink1}}^{static} = \{\text{community } 1\}$ and $\mathbb{C}_{\text{sensor1}}^{static} = \{\text{community } 4\}$.

5.2.4 A Social Forwarding Metric: Sink-Aware Centrality

Based on social networking features, we define a novel metric, Sink-Aware (SA) centrality, to measure the potential ability of a mobile relay for delivering sensed data to the sink. Let $\mathbb{A}^{sink}$ be the set of static clusters, each of which contains at least a sink:

$$\mathbb{A}^{sink} = \bigcup_{\mathcal{A}\in\mathbb{A},\ \mathcal{A}\cap\mathcal{D}\neq\emptyset} \{\mathcal{A}\} \tag{5.2}$$

where $\mathbb{A}$ is the set of all static clusters. In Figure 5.1, for instance, $\mathbb{A} = \{$static clusters 1–4$\}$ and $\mathbb{A}^{sink} = \{$static clusters 2 and 3$\}$. We then define the global sink-aware community set $\mathbb{C}^{sink}$ as

$$\mathbb{C}^{sink} = \bigcup_{x\in\mathcal{A},\ \mathcal{A}\in\mathbb{A}^{sink}} \mathbb{C}_x^{static} \tag{5.3}$$

For instance, $\mathbb{C}^{sink} = \{$ communities 1, 3, and 4 $\}$ in Figure 5.1.

Definition 5.2 (Sink-Aware (SA) Centrality)**.** *For a mobile relay $x \in \mathcal{R}$ with social profile $(\mathbb{C}_x, \boldsymbol{h}_x)$, its SA centrality H_x^{sink} is defined as*

$$H_x^{sink} = \sum_{\mathcal{C}_x^i\in\mathbb{C}^{sink}\cap\mathbb{C}_x} h_x^i \tag{5.4}$$

where h_x^i is the local centrality of x in community $\mathcal{C}_x^i$.

It can be seen that a mobile relay with a high global centrality (i.e. the sum of its all local centralities) does not always have a high SA centrality. For instance, consider a policeman as a mobile relay who can meet a large number of people and therefore has a high global centrality. However, he may rarely patrol the streets where the sinks are deployed, resulting in a low SA centrality. In addition, Figure 5.1 (c) also illustrates a numerical example, where the global centralities of mobile relays x and y are 9 and 7 respectively. Therefore, relay x has a higher global centrality than relay y. However, the SA centrality of x is lower than that of y, i.e. $H_x^{sink} = 6 < H_y^{sink} = 7$).

Let the maximal possible SA centrality over all mobile relays be $\mathrm{H}_{\mathrm{sink}}^{\max} = \max_{x\in\mathcal{R}} H_x^{sink}$. It is obvious that $\mathrm{H}_{\mathrm{sink}}^{\max} \leq |\mathcal{R}|$. Although static nodes do not have centrality concept, we still assign a SA centrality value for each static node $x \in \mathcal{S} \cup \mathcal{D}$,

$$H_x^{sink} = \begin{cases} |\mathcal{R}| & \text{if } x \in \mathcal{A} \in \mathbb{A}^{sink} \\ 0 & \text{otherwise} \end{cases}$$

to support our OBSEA algorithm that seamlessly combines all static nodes and mobile relays in the whole WSN-HR. For instance, in Figure 5.1, $H_x^{sink} = |\mathcal{R}| = 18$, if a static node x is in static clusters 2 and 3; $H_x^{sink} = |\mathcal{R}| = 0$, otherwise.

Compared with classic centrality definitions used in social networks and DTNs (e.g. degree and betweenness centralities [62, 84]), SA centrality provides the destination-awareness for mobile relays, and thus is more effective for WSN-HR. A distributed lightweight algorithm to establish SA centrality at runtime will be introduced in next section.

5.2.5 Economic Network Model: Pricing and Social Profits

To incentivise mobile relays to forward sensor data for the static WSN, we establish a virtual economic network for the WSN-HR. The static WSN $G^{wsn}(\mathcal{S} \cup \mathcal{D}, \mathcal{L}^{wsn})$ can be considered as the employer who pays the mobile relays (the employees) in credits, which can be used for online shopping (e.g. to buy Android/iPhone Apps online). Specifically, at each slot $t \geq 1$, the sensor data producing and trading processes are described as follows:

- When a sensor or sink node $y \in \mathcal{S} \cup \mathcal{D}$ receives (buy) $f_{x,y}(t)$ amount of sensor data from a mobile relay $x \in \mathcal{R}$, y pays $f_{x,y}(t)\lambda_x(t)$ amount of credits to x, where $\lambda_x(t)$ is the selling price per unit data (decided by the seller x).

- When a mobile relay x sells (transmits) data to another relay y, $x, y \in \mathcal{R}$, x will receive a payment of $f_{x,y}(t)(\lambda_x(t) - \lambda_y(t))$ amount credits from y.

- A mobile relay $y \in \mathcal{R}$ can receive data for free from any sensor node $x \in \mathcal{S}$.

- When a sensor node $x \in S$ collects environmental data at a sampling rate $r_x(t)$, $I_x(r_x(t))$ amount of revenue will be provided by the WSN, where $I_x(r_x(t))$ can be any differentiable, non-decreasing, non-negative, and concave utility function of $r_x(t)$.

The sub-network consists of all mobile relays (employees) which can be viewed as a free information market, and every relay trades sensor data with other relays to reap benefits from price difference; similar to the real-world business. The instantaneous profit of a mobile relay $x \in \mathcal{R}$ is defined as

$$\gamma_x^{relay}(t) = \sum_{y \in \mathcal{N}_x(t)} \lambda_x(t) f_{x,y}(t) - \sum_{y \in \mathcal{N}_x(t) \cap \mathcal{R}} \lambda_y(t) f_{y,x}(t) \tag{5.5}$$

where the first term on the right-hand side of (5.5) represents the total revenue of x by selling data to others, and the last term represents the total expenditure of x, i.e. the credits paid for data purchased from other relays in $\mathcal{R} \cap N_x(t)$. Similarly, we can define the instantaneous profit of the static WSN $G^{wsn}()$ as

$$\Gamma^{wsn}(t) = \sum_{x \in \mathcal{S}} I_x(r_x(t)) - \sum_{x \in \mathcal{S} \cup \mathcal{D}} \sum_{y \in \mathcal{N}_x(t) \cap \mathcal{R}} \lambda_y(t) f_{y,x}(t) \tag{5.6}$$

where $\sum_{x \in \mathcal{S}} I_x(r_x(t))$ represents the total instantaneous revenue of the WSN, and the last term of (5.6) represents the total expenditure of the WSN, i.e. the total credits paid for data purchase from mobile relays in $\{y | y \in \mathcal{R} \cap \mathcal{N}_x(t),\ x \in \mathcal{S} \cup \mathcal{D}\}$.

From (5.5) and (5.6), it is easy to verify that the instantaneous global social profit of the whole WSN-HR is

$$\Gamma(t) = \quad \sum_{x \in \mathcal{R}} \gamma_x^{relay}(t) + \Gamma^{wsn}(t) = \sum_{x \in S} I_x(r_x(t)) \tag{5.7}$$

This is because the sum of the internal payments of all nodes in $\mathcal{N}$ is equal to the sum of revenue earned from taking these payments. Therefore, the total social profit is the total external incomes of the WSN.

5.2.6 Social Profits Maximisation

Due to the arbitrary stochastic process of channel state (which may be non-ergodic) $\boldsymbol{c}(t)$, an infinite-horizon time-average social profits may not exist. Therefore, we consider a finite number of slots $t \in \{1,\ 2, ...,\ t_{end}\}$. The objective is to seek an algorithm to solve the following finite-horizon optimisation problem:

$$\textbf{max} \qquad \overline{\Gamma} = \frac{1}{t_{end}} \sum_{t=1}^{t_{end}} \Gamma(t) \tag{5.8}$$

$$\textbf{subject to} \tag{5.9}$$

$$0 \le r_x(t) \le \mathrm{r}^{\max} \quad x \in \mathcal{S}, 1 \le t \le t_{end} \tag{5.10}$$

$$Q_x(t) \le Q_x^{max},\ \forall x \in \mathcal{S} \cup R,\ 1 \le t \le t_{end} \tag{5.11}$$

$$0 \le f_{x,y}(t) \le c_{x,y}(t) \le \mathrm{c}^{\max} \ x, y \in \mathcal{N}, 1 \le t \le t_{end} \tag{5.12}$$

$$\sum_{y \in \mathcal{N}-\{x\}} \overline{f}_{x,y} \ge 1_{\{x \in \mathcal{S}\}} \overline{r}_x + \sum_{y \in \mathcal{N}-\{x\}} \overline{f}_{y,x}, x \in \mathcal{S} \cup \mathcal{R} \tag{5.13}$$

where $\overline{f}_{x,y} = \sum_{t=1}^{t_{end}} f_{x,y}(t)/t_{end}$; and the indicator function $1_{\{x \in \mathcal{S}\}} = 1$ if $x \in \mathcal{S}$, $1_{\{x \in \mathcal{S}\}} = 0$ otherwise. The objective (5.8) is to maximise the time average social profits of all mobile relays and the WSN during the finite-horizon of size t_{end}. The constraint (5.10) represents that the sample rate $r_x(t)$ is bounded by a constant value $\mathrm{r}^{\max} < \infty$, which is realistic for typical sensor nodes. The constraint (5.11) states that the queue backlog $Q_x(t)$ of a sensor node or a mobile relay x should be less than its buffer size Q_x^{max}. The constraint (5.12) represents that the actual amount of data forwarded over each link should not be greater than the capacity of this link. Constraint (5.13) states the the flow conservation law, i.e. node x's average total incoming data rate must not be greater than its average total outgoing data rate.

5.3 Opportunistic Backpressure with Social/Economic Awareness

In this section, we introduce our OBSEA algorithm and a simple distributed scheme to establish SA centrality.

5.3.1 SA Centrality Updating

The following simple GPS-free scheme can establish SA centrality for each mobile relays in a fully distributed way:

- Step 1. Due to the time-varying nature of human mobility patterns [66,80,199], each node operates steps 2–5 during every short-term period (e.g. 6 hours) to obtain the transient SA centrality rather than the long-term cumulative one. At the beginning of a short-term period, every mobile relay $x \in \mathcal{R}$ initialises its SA centrality as $H_x^{sink} = 0$.

- Step 2. Each mobile relay $x \in \mathcal{R}$ establishes the social profile $(\mathbb{C}_x, \boldsymbol{h}_x)$ during every period, by using a distributed community detection algorithm [86] and the social neighbour table $\mathcal{N}_x^{social}$ established online.

- Step 3. Each static node maintains a set $\mathcal{F}_x = \{y | \in \mathcal{R}, ICT_{x,y} < \mathrm{ICT}^{\max}\}$, i.e. the set of mobile relays that visit x frequently[3]. Then x can establish $\mathbb{C}_x^{static}$ as

$$\mathbb{C}_x^{static} = \bigcup_{y \in \mathcal{F}_x,\ (\mathcal{F}_x \cap \mathcal{C}) \in \mathbb{C}_y} \{\mathcal{C}\}$$

- Step 4. If a node $x \in \mathcal{S} \cup \mathcal{D}$ in a static cluster $\mathcal{A}$ that also contains one or multiple sinks, then x broadcasts $\mathbb{C}_x^{static}$ to all other nodes in $\mathcal{A}$. As a result, every static node in $\mathcal{A}$ can know the set $\mathbb{C}_{\mathcal{A}}^{static} = \bigcup_{y \in \mathcal{A}} \mathbb{C}_y^{static}$, i.e. the set of communities whose geographic areas static cluster A is located in. For instance, in Figure 5.1, $\mathbb{C}_{static\ cluster\ 2}^{static} = \{$ communities 1 and 2$\}$.

- Step 5. When a mobile relay $x \in \mathcal{R}$ visits a node y in a static cluster $\mathcal{A}$ that contains a sink, x checks whether y meets any node in $\mathcal{A}$ during current period. If not, x requires $\mathbb{C}_{\mathcal{A}}^{static}$ from y, and then updates its SA centrality as

$$H_x^{sink} = H_x^{sink} + \sum_{\mathcal{C}_x^i \in \mathbb{C}_{\mathcal{A}}^{static} \cap \mathbb{C}_x} h_x^i$$

[3]A static node $x \in \mathcal{S} \cup \mathcal{D}$ establishes its $\mathcal{F}_x$ based on the same way that a mobile relay $y \in \mathcal{R}$ establishes its social neighbour table $\mathcal{N}_y^{social}$.

5.3.2 OBSEA Algorithm

At each slot $t \geq 1$, each node $x \in \mathcal{N}$ first observes its current neighbours that it is in contact with $\mathcal{N}_x(t)$, the queue backlogs of itself and its contact neighbours, and channel capacities of all its outgoing links, $c_{x,y}(y), y \in \mathcal{N}_x(t)$. Then each node runs the OBSEA algorithm as follows:

Pricing. The selling price set by every node $x \in \mathcal{N}$ in slot t is:

$$\lambda_x(t) = (Q_x(t) + \alpha(\mathrm{H}_{\mathrm{sink}}^{\max} - H_x^{sink}))/\lambda_{scale} \tag{5.14}$$

where $\lambda_{scale} > 0$ is the price-scaling parameter that *does not* impact the global social profits, but controls the profit ratio between all mobile relays (all employees) and the static WSN (the employer), i.e. $\sum_{x\in\mathcal{R}} \gamma_x^{relay}(t)/\Gamma^{WSN}(t))$; and $\alpha \geq 0$ is the weighting parameter for SA centrality awareness in the routing component of OBSEA. When $\alpha = 0$, the routing of OBSEA is the pure queue-backlog aware (backpressure) algorithm; as $\alpha \to +\infty$, the routing of OBSEA tends to be based on SA centrality only. It is worth noting that the selling price $\lambda_x(t)$ *is always non-negative*, due to the non-negative values of α, $Q_x(t)$, λ_{scale}, and $\mathrm{H}_{\mathrm{sink}}^{\max} - H_x^{sink}$.

Rate Control. Each source node $x \in \mathcal{S}$ sets its data sampling rate $r_x(t)$ to maximise the following simple algorithm.

$$\textbf{max} \qquad I_x(r_x(t)) - r_x(t)Q_x(t)/V \tag{5.15}$$

subject to

$$0 \leq r_x(t) \leq \mathrm{r}^{\max} \tag{5.16}$$

where $V > 0$ is the predefined control parameter for the tradeoff between queue backlogs and social profits. Since $I_x(r_x(t))$ is concave, problem (5.15) adopts an unique maximiser as

$$\widetilde{r}_x(t) = \min[\max[I_x'^{-1}(Q_x(t)/V),\ 0],\ \mathrm{r}^{\max}]$$

where $I_x'^{-1}()$ represents the inverse function of the utility function $I_x()$'s first derivative.

Routing. Recall that Q_x^{max} is the data buffer size of node x. Each node $x \in \mathcal{S} \cup \mathcal{R}$ computes the price differential between itself and each of its instantaneous contact neighbour $y \in \mathcal{N}_x(t)$, by using (5.14). Then x computes the weight of instantaneous link as

$$w_{x,y}(t) = \begin{cases} (\lambda_x - \lambda_y)\lambda_{scale} & \text{if } Q_y(t) < Q_y^{max} - \eta_y(t) \\ 0 & \text{otherwise} \end{cases}$$

where $\eta_y(t) = \sum_{z \in \mathcal{N}_y(t)} c_{z,y}(t) + 1_{\{y \in \mathcal{S}\}} \mathrm{r}^{\max}$ is the largest possible amount of data that can be injected into node y at slot t. Then node x transmits $f_{x,y}(t)$ amount of data packets to y:

$$f_{x,y}(t) = \begin{cases} c_{x,y}(t) & \text{if } w_{x,y}(t) > 0 \\ 0 & \text{otherwise} \end{cases} \tag{5.17}$$

It is clear that sensor-data packets are dynamically forwarded hop-by-hop rather than through maintained end-to-end paths.

Incentive and Credit Transfer. Based on the actual produced and transmitted sensor data decided by above rate control and routing respectively, each node in $\mathcal{N}$ transfers the credits using the mechanism described in Subsection 5.2.5.

All the static nodes in the WSN are enforced to obey the rate control and routing rule without any incentive. Now we analyse why mobile relays are willing to follow above routing rule. For $w_{x,y} > 0$, there are two cases:

- *Both x and y are mobile relays.* Based on the rationally-selfish assumption of mobile relays, if $w_{x,y}(t) > 0$, x can achieve $w_{x,y}(t)$ per packet benefit by selling sensor data to y. For node y, although it pays credits for buying sensor packets, but its selling price will be increased due to the incremental of its queue backlog (see (5.14)). y can sell sensor-data packets to other mobile relays with lower selling price or to static nodes. Therefore, if $w_{x,y}(t) > 0$, both nodes are willing to trade the sensor data at this contact.

- *One is a static node and the other is a relay.* If x is a sensor node, y can get free data from x, which can be sold to others in the future; if x is a relay, it can get benefit by selling the data.

Queue Update. Queue backlog of each node $\in \mathcal{N}$ are updated according to (6.3).

Since every node $x \in \mathcal{N}$ requires only the information of its instantaneous neighbours in $\mathcal{N}_x(t)$, the OBSEA algorithm is fully distributed. In addition, OBSEA is based on only *current* knowledge of the network at current slot t and *does not require* any prediction capacity for future knowledge after slot t.

Performance Analysis of OBSEA can be found in Appendix B

5.3.3 Control Overhead

The control overhead of the OBSEA algorithm is discussed as follows:

- **Communication Overhead.** Since OBSEA is fully distributed, each node only transmits at most one beacon to communicate its local queue backlog and SA centrality at each slot.

- **Computational Overhead.** Since a node $x \in \mathcal{N}$ can be in contact with at most $|\mathcal{N}| - 1$ nodes at each slot, it is clear that both the SA centrality update and the operations of OBSEA require at most $O(|\mathcal{N}|)$ simple arithmetic calculations only. It is worth noting that this is loosely bounded, since a node can normally contact a small number of nodes at each slot due to the sparse density of the network.

- **Storage Overhead.** Each node needs to maintain its SA centrality and its social profile (at most $2|\mathcal{N}| - 1$) values. Therefore, the per node storage overhead is O($|\mathcal{N}|$).

5.4 Simulation

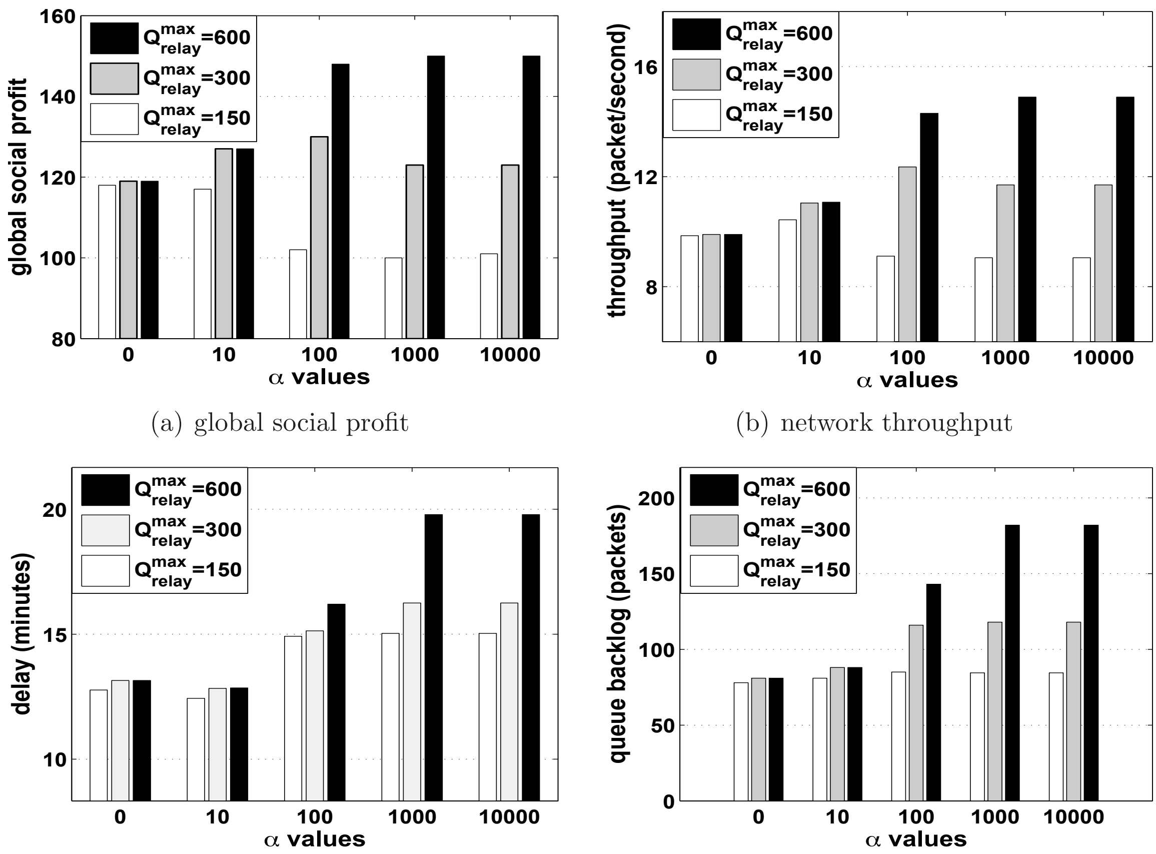
We implement the OBSEA algorithm in Castalia [3]; a realistic WSN simulator. Several real human mobility traces exist, such as the MIT reality [58] and the Infocom [38] traces, however, public GPS data for WSN is non-existent. To integrate mobile relays and static WSNs into a geographic area, therefore, we constructed a WSN-HR that consists of a random deployed static WSN and multiple mobile relays that follow the Heterogeneous Human Walk (HHW) mobility model [199]. HHW is a realistic human mobility model based on social network theory, which exhibits various features of real human mobility and social networks.

The size of the geographic area was set as 1.7×1.7 km^2 which is approximately the same size of the City of London. The total number of nodes was set as 100, which consists of 17 sensor nodes, 3 sinks, and 80 mobile relays. We set the duration of a slot as 1 seconds and ran the simulation for 10^6 slots (the equivalent of about 11 days). We considered the 4-clique community structure for the HHW mobility model and set the parameters as $\mathrm{PR}^{\mathrm{Csize}} = 1.2$, $\mathrm{PR}^{\mathrm{Osize}} = 2$, $\mathrm{PR}^{\mathrm{Csize}} = 1.2$, $\mathrm{PR}^{\mathrm{MN}} = 2$, and $\mathrm{PR}^{\mathrm{Local}} = 2$, according to the observations of real social networks and human mobility traces. In addition, the speed of each mobile relay was randomly distributed between 0 and 10 m/s (between walking and urban vehicular speeds.).

The transmission ranges of all nodes were set as 50 metres, and the data forwarding rate (capacity) of each instantaneous contact link is randomly selected between 1 and 20 packets per second. We set $\mathrm{r}^{\mathrm{max}} = 10$, $\eta^{max} = 50$, and $\mathrm{H}^{\mathrm{max}}_{\mathrm{sink}} = 80$. For each sensor node $x \in \mathcal{S}$, we set its buffer size $Q_x^{max} = 150$, and utility function $I(r_x(t)) = 20\ln(1 + r_x(t))$. The profit-backlog tradeoff parameter V was set as 5 according to (B.13).

Figure 5.3 shows the average global social profit, network throughput, end-to-end delay, and queue backlogs with different weighting parameter α in (5.14), where $\mathrm{Q}^{\mathrm{max}}_{\mathrm{relay}}$ represents the data buffer size of each mobile relay, i.e. $Q_x^{max} = \mathrm{Q}^{\mathrm{max}}_{\mathrm{relay}}, \ \forall \ x \in \mathcal{R}$. Here, $\mathrm{Q}^{\mathrm{max}}_{\mathrm{relay}}$ can be understood as either the physical memory size of mobile phones, or the memory space that the phone-user are willing to provide for the sensor data. We run each simulation five times and the results are highly close. The results shown in Figure 5.3 are the average of the same simulations.

When $Q_{relay}^{max} = 150$, OBSEA algorithm with small α values (large weight for queue-backlog awareness) perform better than that with large α values (large weight for social awareness). When $Q_{relay}^{max} = 300$, OBSEA with $\alpha = 100$ achieves better global social profit and throughput than that with other α values. Finally, when $Q_{relay}^{max} = 600$, larger weight for social awareness achieves better social profit and throughput. The results above demonstrate that neither the pure backlog-aware scheme ($\alpha = 0$) nor the (approximately) pure social-aware scheme ($\alpha = 10000$) can achieve the optimal performance in all network conditions. In addition, it can be seen that OBSEA can adapt to different network settings, by simply adjusting the weighting parameter α.WSN-HR



(a) global social profit

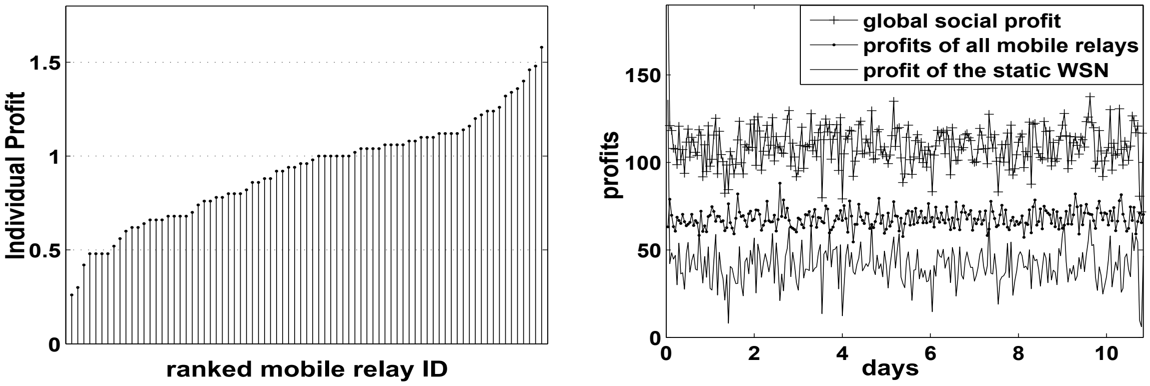
(b) network throughput

(c) per packet end-to-end delay

(d) average queue backlog

Figure 5.3: Long-term average simulation results for different α and Q_{relay}^{max}. (a) The object of this chapter $\overline{\Gamma}$, which can be transferred to various units such as pounds or dollars per day. (b) The sum of the packet receiving rates of all sinks. (c) The average sensor-to-sink delivery delay of all generated sensor-data packets during the simulation. (d) Average queue backlog of all sensor nodes and mobile relays during the simulation.

Figure 5.4 (a) shows that the profit of every mobile relay is positive, which means that every phone user achieves benefits through relaying sensor data for the static WSNs. The phone users just need to allocate some memory space $Q_{\text{relay}}^{\text{max}}$ and cost little power consumption for the short-range data transmissions to achieve such benefits. Figure 5.4 (b) shows that the 'win-win situation' (positive benefits) is achieved by both static WSNs and mobile relays during every hour of the simulation time. We run the other simulations with different parameter settings, all of which shows similar features of the results in Figures 5.3 and 5.4.



(a) average profits of individual relays (b) dynamic social profits for 10 days

Figure 5.4: Distribution of average individual profit and dynamic global social profit with $Q_{\text{relay}}^{\text{max}} = 600$, $\alpha = 100$, and $\lambda_{scale} = 10^4$.

5.5 Summary

In this chapter, we combine network science principles and Lyapunov network optimisation theory to develop a data collection scheme for sparse sensor networks with mobile phones. Without making any assumptions regarding the topology, mobility, and channel conditions of the network, we devise an approach for joint rate control, opportunistic routing, and resource pricing, which maximises the global social profit of the network. By exploiting the social and economic behaviours of mobile phone users, a lightweight algorithm (OBSEA) is proposed that is fully distributed and scalable. Simulation results show that OBSEA is adaptive to different network scenarios and outperforms pure backpressure and pure social-aware schemes.

Due to selfishness of mobile phone users secure pricing is another main challenge these networks. Every selfish user can attempt to affect the network by providing wrong information in terms of price, connectivity etc. to achieve maximum profit for itself, which may result in inefficiency of the system.

In the next chapter, we have designed a more strong economic system that includes *faithful* sensor data market design; thus discouraging phone users to subvert the market through misinformation, and incentivization based on the benefits a user brings to the system.

Chapter 6

Faithful Data Collection in Mobile Phone Sensing Systems using Taxation

The use of sensor-enabled smart phones is considered to be a promising solution to large-scale urban data collection. In previous chapters we investigated cost-effective data collection solution for WSNs with Mobile Phones (WSN-MP) using hybrid cellular and opportunistic short-range communications. We also established economic model to incentivise mobile users to participate in the network. Due to the human involvement in virual economic network, every selfish device owner can attempt to manipulate the network to maximize own benefits e.g.,by reporting wrong cost of its resources or by generating fake data. This will reduce the efficiency of the network. This chapter focuses on developing faithful market design for Mobile Phone Sensing System (MPSS) that consists of mobile phones who generate the sensor data and forward data to sinks (Wifi Access Point or cellular base-station).

We first develop an adaptive and distribute algorithm OptMPSS to maximize phone user financial rewards accounting for their costs across the MPSS.we then propose BMT to incentivize phone users to participate, while not subverting the behavior of OptMPSS using mechanism design theory.We show that our proven incentive compatible approaches achieve an asymptotically optimal gross profit for all phone users. Experiments with Android phones and trace-driven simulations verify our theoretical analysis and demonstrate that our approach manages to improve the system performance significantly (around 100%) while confirming that our system achieves incentive compatibility, individual rationality, and server profitability.

S	Server.
$\mathcal{N}$	The set of mobile phones.
$r_i(t)$	The sensing rate of mobile phone $i \in \mathcal{N}$ at slot t.
$\mu_{i,j}(t)$	capacity from a phone $i \in \mathcal{N}$ to a phone $j \in \mathcal{N}$ or to the server $j = S$ at slot t.
$\mathcal{N}_i(t)$	Time-varying *temporary neighbor table* of phone i.
$f_{i,j}(t)$	Amount of forwarded data from phone i to its current neighbor $j \in \mathcal{N}_i(t)$ at slot t.
$\boldsymbol{x}_i(t)$	Vector for sensing and data forwarding actions of a phone $i \in \mathcal{N}$ at slot t.
$Q_i(t)$	The queue backlog of phone $i \in \mathcal{N}$ at slot t.
$f_i^{out}(t)$	Total numbers of outgoing packets from phone i at slot t.
$f_i^{in}(t)$	Total numbers of incoming packets for phone i at slot t.
$Q_S(t)$	Queue backlog of the server.
$p_i^s(t) > 0$	Sensing price per packet of phone i at slot t.
$p_{i,j}^t(t)$	Transmission price for phone i to send a sensor data packet to a neighbor j.
$cost_i(t)$	Total cost of each phone $i \in \mathcal{N}$.
$\boldsymbol{p}_i(t)$	Vector for price profile of phone i at slot t.
$\overline{r}_i$	Average sensing rate of phone i over the time horizon $[1, \mathrm{t_{end}}]$.
$v_i(\overline{r}_i)$	Revenue function to indicate the time-average contribution level of phone i to the MPSS
α	Percentage of the *global social revenue* that is allocated to all phones.
φ_i	Time-average *gross profit* of each phone $i \in \mathcal{N}$.
$\overline{cost}_i$	Time-average of $cost_i(t)$ over time horizon $1 \leq t \leq \mathrm{t_{end}}$.
$\overline{f}_i^{in}$	Time-average of $f_i^{in}(t)$ over time horizon $1 \leq t \leq \mathrm{t_{end}}$.
$\overline{f}_i^{out}$	Time-average of $f_i^{out}(t)$ over time horizon $1 \leq t \leq \mathrm{t_{end}}$.
$\overline{r}_i$	Average sensing rate for each phone $i \in \mathcal{N}$ at slot $\mathrm{t_{end}}$.
$w_{i,j}(t)$	The routing weight of wireless link (i,j) at slot t.
θ_i	Private type (parameters) of each phone i.
Θ_i	Set of all possible types θ_i.
$\mathbf{x}(\boldsymbol{\theta})$	Joint rate control and routing decisions of the MPSS during the whole time horizon.
$\boldsymbol{\Theta}$	Set of all possible $\boldsymbol{\theta}$.
$\widehat{\theta}_i$	Reported type (parameters) of each phone to the server S.
$\widehat{\boldsymbol{\theta}}$	Reported types of all phones.
$\boldsymbol{\lambda}$	Monetary transfer function.
u_i	**Net profit** of each phone user.
u_S	Time-average **server profit**.
$\mathbf{x}^{vcg}(\widehat{\boldsymbol{\theta}})$	social decision rule of the VCG mechanism.

Table 6.1: Summary of symbols used in Chapter 6.

MPSSs	Mobile Phone Sensing Systems.
IoTs	Internet of Things.
BMT	Backpressure Meet Taxes.
VCG	Vickrey-Clarke-Groves.

Table 6.2: Summary of abbreviations used in Chapter 6.

6.1 Introduction

Proliferation of these sensor-rich mobile devices along with collection of ubiquitous sensors are envisioned to constitute a powerful mobile sensing system which can be used to understand and analyse many interesting phenomena of the physical world. Urban sensing is different from existing sensor networks by bringing people in the loop because people are no longer just consumers of sensed data but also are source of sensed data. Mobility of humans and increasing short range communication capabilities of smart phones (LTE, bluetooth, wifi-direct) have paved the way for opportunistic networking in urban sensing applications. Mobile Phone Sensing Systems (MPSSs) are characterized by social-based mobility and participation of people for data generation and communication. However, to build such a MPSS with hybrid cellular and opportunistic short-range communications, the following research issues must be addressed:

Networking Issues. As discussed in previous chapters, it is a challenge to perform the sensing and opportunistic multi-hop data transmission tasks that are adaptive to the time-varying and potentially unpredictable network states, including fluctuating wireless channel quality; intermittent connectivity caused by phone user movement; heterogeneous transmission and sensing costs across mobile phones; 3G/4G mobile data costs; and the opportunistic availability of nearby free Internet access points.

Economic Issues. To encourage the phone users to participate the MPSS, they should be properly rewarded to cover sensing and transmission costs [107,116,198] as discussed in chapter 4 and 5. In addition, the self-interest phone users may try to maximize their benefits strategically by misreporting their local state parameters. For instance, in order to prolong battery lifetime, a phone user may hide that she is connected to a free WiFi router to avoid relaying other nearby phone data to the server. This would result in significant performance degradation of the system.The problem of building mechanism and protocols that can tolerate selfish behaviour is an important issue in the design of networking protocols and distributed systems for mobile urban sensing systems. Therefore, incentivization to force faithful behaviour is a key issue for MPSS, which is much more challenging compared to pure cellular networks.

6.1.1 Our Approach

In this chapter, we present theoretical and practical studies to address above two issues. Our contributions are summarized as follows:

1. We formulate a finite-horizon stochastic optimization problem for continuous data collection in MPSS using hybrid cellular and opportunistic short-range communications. The objective of the formulated problem is to maximize the *global gross profit*, i.e. the total financial rewards of all phone users after costs incurred by performing the sensing and transmission tasks are deducted.

2. We develop a lightweight joint sensing rate control and dynamic routing algorithm, OptMPSS to solves the data collection problem in a fully distributed and therefore scalable way.

3. We propose a fully distributed mechanism, Backpressure Meet Taxes (BMT), to incentivize phone users to faithfully implement OptMPSS, by imposing taxes or providing subsidies for each phone user, depending on her impact on the rest of phone users in the MPSS. We prove that BMT manages to achieve asymptotic incentive compatibility [179]. To our knowledge, BMT is the first approach that integrating algorithmic mechanism design theory [91, 146, 169] to the stochastic Lynapunov optimization framework [69, 142]. Besides MPSS, this method developed for BMT also has a great potential to be applied to other stochastic distributed systems with self-interest and strategic users.

4. Through experiments with WiFi-direct-enabled Android devices and extensive simulations with real human mobility trace [5], we demonstrate that system performance can be significantly improved by exploiting low-cost short-range communications, in terms of global social profits

and phone users' costs. Evaluation results also show that each phone user can always get a positive *net profit* (i.e. gross profit plus subsidies or minus taxes) and the server never incurs a deficit (i.e. the server always obtain a positive profit). Furthermore, each phone user cannot increase her net profit improvement by lying about her private parameters. These results demonstrate that BMT can achieve individual rationality, server profitability, and incentive compatibility (faithful implementation) in practice.

6.1.2 Related Work

Recently, several incentive-based mechanisms have been proposed for MPSS [107,116,130,198]. [198] develops platform-centric and user-centric schemes based on a Stackelbergy game and auction theory respectively. [107] proposes a mechanism based on a Bayesian game to minimize participation costs while ensuring certain service qualities, by determining the level of user participation (i.e. sensing rate). However, all of these schemes focus on MPSS with pure cellular radios only, which cannot be directly used in MPSS with hybrid cellular and multi-hop short-range communications.

The explosive growth of cellular traffic has motivated an increase in research into cellular traffic offload using other forms of opportunistic connectivity, including WiFi [53, 118, 211] and Bluetooth [79]. However, none of these focus on MPSS. EffSense [190], considers MPSS with the same hybrid wireless networks as us. However, this heuristic-based scheme does not provide any performance guarantees, and does not consider incentivisation for the strategic and self-interest phone users.

Stochastic Lyapunov optimization [69, 142] provides elegant and powerful theoretical tools to derive backpressure style cross-layer network optimization and control algorithms. Due to their adaptiveness to network dynamics, several backpressure rate control and routing schemes [15, 158] have been proposed for opportunistic mobile networks. However, again none of them focus on MPSS nor do they account for incentivisation and the strategic behaviors of phone users.

Mechanism design [91, 146, 169] is concerned with how to make a global decision with desirable properties in systems consisting of strategic self-interest individuals who have private information. Recent theoretical work [152, 179] on distributed Vickrey-Clarke-Groves (VCG) mechanisms enables the faithful implementation of algorithms producing desired outcomes in a distributed way. However, these approaches focus on deterministic rather than stochastic systems. Our work combines threads of all the above works in a novel way.

6.1.3 Chapter Organization

The next section presents the system model and design objectives. Section 6.3 describes the OptMPSS algorithm. Mechanism design models are established in Section 6.4. Section 6.5 presents the BMT algorithm, then Section 6.6 discusses the performance evaluation. Finally, we conclude the chapter in Section 6.7.

6.2 System Model and Objectives

As shown in Fig. 6.1, we consider a MPSS that consists of a server S and a set of mobile phones $\mathcal{N}$ collecting urban sensing data. The MPSS operates during a finite time horizon (e.g. a week) with discrete time slots $t = \{1,\ 2,\ ..., \mathrm{t_{end}}\},\ \mathrm{t_{end}} < +\infty$. Every phone can communicate with the server S through 3G/ 4G cellular radios, or through the low-cost WiFi when it passes a WiFi router (slots 1-3 in Fig. 6.1). In addition, phones in immediate proximity can communicate with each other, using short-range communications such as WiFi direct (slot 2) and Bluetooth 4.0 (slot 3).

6.2.1 Sensing and Communication Models

At each time slot t, each phone i produces $0 \leq r_i(t) \leq \mathrm{r_{max}}$ sensor data packets, where the finite sensing rate upper bound $\mathrm{r_{max}} \leq \infty$ is defined by the specific mobile sensing application. Let $\mu_{i,j}(t) \geq 0$ be the channel capacity from a phone $i \in \mathcal{N}$ to a phone $j \in \mathcal{N}$ or to the server

Figure 6.1: An illustrative example of a MPSS with hybrid cellular and short-range communications, consisting of 4 phones and a server.

$j = S$ at slot t, i.e. the maximum (integer) number of data packets that can be successfully transmitted from i to j during slot t. It can be seen that $\mu_{i,j}(t)$ may vary significantly over the time slots, due to the stochastic phone user movement and wireless channel qualities. In practice, all possible channel capacities (i.e. $\mu_{i,j}(t), \forall i,\ j,\ t$) must have an finite upper bound $\mu^{\max} \leq \infty$, determined by the finite data rate of the wireless transceivers .

Each phone i maintains a time-varying *temporary neighbor table* $\mathcal{N}_i(t)$, consisting of the server S (if currently connected) and the phones in proximity at slot t:

$$\mathcal{N}_i(t) := \{j : j \in \mathcal{N} \cup \{S\}, \mu_{i,j}(t) > 0\} \tag{6.1}$$

In practice, $\mathcal{N}_i(t)$ can be established by using neighbor discovery schemes such as [23]. Denote $0 \leq f_{i,j}(t) \leq \mu_{i,j}(t)$ as the amount of forwarded data from phone i to its current neighbor $j \in \mathcal{N}_i(t)$ at slot t. We use a vector

$$\boldsymbol{x}_i(t) = (r_i(t),\ f_{i,j}(t),\ j \in \mathcal{N}_i(t)) \tag{6.2}$$

to represent the sensing and data forwarding actions of a phone $i \in \mathcal{N}$ at slot t.

6.2.2 Queue Dynamics

Each phone i maintains a data queue with size $Q_i(t) \geq 0$ to store the sensing data collected by itself and received from other phones. Considering sensing and data forwarding dynamics, the queue backlog of each phone i updates as follows:

$$Q_i(t+1) = |Q_i(t) - f_i^{out}(t)|_+ + r_i(t) + f_i^{in}(t) \tag{6.3}$$

where for any real number a, the operator $|a|_+ = a$ if $a > 0$; $|a|_+ = 0$ otherwise. $f_i^{out}(t) = \sum_{j \in \mathcal{N}_i(t)} f_{i,j}(t)$, and $f_i^{in}(t) = \sum_{j \in \mathcal{N}_i(t)} f_{j,i}(t)$ represent the total numbers of outgoing and incoming packets of phone i at slot t respectively. It is worth noting that the queue backlog of the server S is always zero, i.e. $Q_S(t) = 0, \ \forall t$, since it is the destination of all sensor data packets.

6.2.3 Costs of Phones

At each slot t, each phone incurs costs due to sensing and data transmission. Let $p_i^s(t) > 0$ be the per packet sensing price of phone i at slot t . Therefore, the sensing cost of phone i at slot t is $p_i^s(t) r_i(t)$. In practice, the sensing price $p_i^s(t)$ depends on both MPSS application requirements, and the available resources that phone i has at time t such as remaining battery energy. Denote $p_{i,j}^t(t)$ as the per packet transmission price for phone i to send a sensor data packet to a temporary neighbor j. If j is the server S, then $p_{i,j}^t(t)$ depends on financial cellular costs, the availability of a nearby WiFi router, and the remaining battery level of phone i. Specifically, price $p_{i,j}^t(t)$ are normally significantly smaller when i sends data to the server through WiFi than through cellular radios. If j is another phone, the transmission price $p_{i,j}^t(t)$ mainly depends on battery concerns of the users of phones i and j. We can see that the data transmission price is highly dynamic and heterogeneous across different wireless transmission links.

It is worth noting that although sensing and transmission prices are influenced by various practical aspects, they can be all normalized to monetary values (e.g. US dollar or credits per packet), estimated by each phone user herself rather then the server. The detailed estimation in practice is out of the scope of this work, but our BMT scheme can guarantee that faithful estimation is the best strategy for each self-interest phone user.

For a given slot t, the total cost of each phone $i \in \mathcal{N}$ can be computed as

$$cost_i(t) = \boldsymbol{p}_i(t)\boldsymbol{x}_i^{\intercal}(t) \tag{6.4}$$

where the vector

$$\boldsymbol{p}_i(t) = (p_i^s(t),\ p_{i,j}^t(t),\ j \in \mathcal{N}_i(t) \tag{6.5}$$

characterizes the price profile of phone i at slot t.

6.2.4 Revenue and Gross Profit Maximization

During the complete time horizon, the server obtains totally $\mathrm{t_{end}} \sum_{i\in\mathcal{N}} v_i(\overline{r}_i)$ amount of monetary revenue (*i.e. global social revenue*), by selling the collected mobile sensing date to external MPSS users. Here, $\overline{r}_i$ represents the average sensing rate of phone i over the time horizon $[1, \mathrm{t_{end}}]$, and the revenue function $v_i(\overline{r}_i)$ can be any concave (includes linear), differentiable, and non-decreasing function of $\overline{r}_i$. The revenue function may differ across mobile phones, depending on specific sensing applications and the Quality of Information (QoI) of sensor data produced by each phone i [107, 126]. Therefore, $v_i(\overline{r}_i)$ indicates the time-average contribution level of phone i to the MPSS.

At slot $\mathrm{t_{end}}$, the server computes $v_i(\overline{r}_i)$ and makes a payment $\alpha \mathrm{t_{end}} v_i(\overline{r}_i)$ to each phone i (shown in Fig. 6.1), where the system parameter $0 < \alpha \leq 1$ is the percentage of the *global social revenue* $\mathrm{t_{end}} \sum_{i\in\mathcal{N}} v_i(\overline{r}_i)$ that is allocated to all phones. As a result, the time-average *gross profit* of each phone user is given by:

$$\varphi_i = \alpha v_i(\overline{r}_i) - \overline{cost}_i,\ \forall i \in \mathcal{N} \tag{6.6}$$

where $\overline{cost}_i$ is the time-average of $cost_i(t)$ over time horizon $1 \leq t \leq \mathrm{t_{end}}$.

We call the time-average aggregated gross profit of all mobile phones $\sum_{i\in\mathcal{N}} \varphi_i$, as the *global* *gross profit*. The MPSS aims to maximize the global gross profit by solving the following finite-horizon stochastic problem:

$$\underset{\boldsymbol{x}_i(t), i\in\mathcal{N}}{\textbf{maximize}} \qquad \Phi = \sum_{i\in\mathcal{N}} \varphi_i \tag{6.7}$$

s.t.

$$r_i(t) < \mathrm{r}_{\max},\ i \in \mathcal{N}\ 1 \leq t \leq \mathrm{t}_{\mathrm{end}} \tag{6.8}$$

$$f_{i,j}(t) \leq \mu_{i,j}(t),\ i \in \mathcal{N}, j \in \mathcal{N}_i(t),\ 1 \leq t \leq \mathrm{t}_{\mathrm{end}} \tag{6.9}$$

$$\overline{r}_i + \overline{f}_i^{in} - \overline{f}_i^{out} = 0,\ i \in \mathcal{N} \tag{6.10}$$

where $\overline{f}_i^{in}$ and $\overline{f}_i^{out}$ are the time-averages of $f_i^{in}(t)$ and $f_i^{out}(t)$ over time horizon $1 \leq t \leq \mathrm{t}_{\mathrm{end}}$ respectively. Constraint (6.10) states the flow conservation law, i.e. the average total incoming and outgoing data rate should be equal for each phone. This constraint also ensures that the server will know the average sensing rate $\overline{r}_i$ for each phone $i \in \mathcal{N}$ at slot $\mathrm{t}_{\mathrm{end}}$. Section 6.3 will develop OptMPSS, an optimal distributed solution to problem (6.7)-(6.10). However, because each phone owner i is only interested in maximizing her own profit φ_i rather than the global gross profit of the system Φ, the optimal solution to problem (6.7)-(6.10) cannot be implemented without proper incentivization mechanism that encourages phone owners to apply the OptMPSS. All variables regarding the incentivization mechanism such as *net profit* will be defined in Section 6.4.

6.2.5 Objective

The objective of this work is to develop an algorithm that can achieve the following desired properties:

1. **Global Gross Profit Optimality.** The algorithm should be the optimal solution to problem (6.7)-(6.10).

2. **Adaptiveness.** The algorithm should be adaptive to all possible dynamic network states,

including time-varying and heterogeneous sensing and transmission cost across phones; wireless link qualities; and network connectivity (e.g. including extremely dense networks where all phones can always communicate with each other, to extremely sparse networks where short-range communication is rare or not available).

3. **No Prediction Requirement.** The desired algorithm is based on the current system state only, and does not require the prediction of any future MPSS information.

4. **Distributed and Real-time Operations.** The computational and communication overheads of the algorithm should be lightweight for real-time operations of each phone.

5. **Individual Rationality.** Each participating phone user should obtain a non-negative *net profit*, which is formally defined in Equation (6.17).

6. **Server Profitability.** The server S should not incur a deficit, which means a non-negative *server profit* (formally defined in Equation (6.18)) should be achieved.

7. **Incentive Compatibility.** Adopting the action suggested by the proposed algorithm should be the best strategy for each phone user, regardless others' actions. An important corollary of incentive compatibility is that using hybrid cellular and (opportunistic) short-range communications will always result in a same or increased net profit for each phone, compared with using the cellular communications alone.

6.3 The OptMPSS Algorithm

In this section, we develop an fully distributed algorithm, OptMPSS to optimize global gross profit (6.7), by controlling the action $\boldsymbol{x}_i^\intercal(t)$ of each phone $i \in \mathcal{N}$ at every slot $1 \leq t \leq \mathrm{t}_{\mathrm{end}}$: its sensing rate $r_i(t)$ and the data forwarding rate $f_{i,j}(t)$ to each of its temporary neighbors $j \in \mathcal{N}_i(t)$. Initially, in this section, we assume that all phone users are willing to truthfully implement the OptMPSS algorithm. We will relax this assumption in later sections.

6.3.1 Distributed Operations of OptMPSS

At each slot $1 \leq t \leq \mathrm{t_{end}}$, each phone $i \in \mathcal{N}$ operates as follows:

1. Sensing Rate Control. Phone i sets its sensing rate $r_i(t)$ as

$$r_i(t) = \min(\mathrm{r_{max}}, \alpha v_i'^{-1}(\frac{Q_i(t) + Vp_i^s(t)}{V})) \tag{6.11}$$

where $v_i'^{-1}()$ represents the inverse function of revenue function $v_i()$'s first derivative, and $V > 0$ is a system parameter defined by the server.

2. Opportunistic Routing and Data forwarding. Phone i computes a weight $w_{i,j}(t)$ for each temporary neighbors $j \in \mathcal{N}_i(t)$ as

$$w_{i,j}(t) = (Q_i(t) - Q_j(t))\mu_{i,j}(t) - Vp_{i,j}^t(t) \tag{6.12}$$

Based on $w_{i,j}(t)$, i sets the forwarding rate $f_{i,j}(t)$ for each of its temporary neighbor $j \in \mathcal{N}_i(t)$ as:

$$f_{i,j}(t) = \begin{cases} \mu_{i,j}(t) & \text{if } w_{i,j}(t) > 0 \\ 0 & \text{otherwise} \end{cases} \tag{6.13}$$

Remark 1. Since every node $i \in \mathcal{N}$ requires only the information of its temporary neighbors in $\mathcal{N}_x(t)$, OptMPSS algorithm is fully distributed. In addition, OptMPSS requires *current* knowledge of the network only for slot t and *does not require* any future knowledge after slot t. At each slot, each phone broadcasts a one-hop beacon message to communicate its queue backlog to its current temporary neighbors and performs simple arithmetic calculations. Therefore, the per slot per node communication of OptMPSS is $O(1)$ with respect to the network size $|\mathcal{N}|$.

Proof of the asymptotical optimality of OptMPSS can be found in appendix C

6.4 Mechanism Design for Faithful MPSS

A key property of MPSS is that all parameters local to each phone are private and not observable to other phones and the server. Consequently, this provide the phone users with the opportunity to subvert the system by miscommunicating their local parameters. In this section, we briefly discuss algorithmic mechanism design [146, 169], which studies faithful implementation of an intended algorithm in a system with a center and a set of individuals with private parameters. MPSS can be viewed as such a system where the center is a server S and the individuals are mobile phones, and we aim to design a mechanism to faithfully implement the intended OptMPSS algorithm.

6.4.1 Centralized Mechanism Design

Although we focus on distributed mechanism design, for readability, we first discuss direct revelation (centralized) mechanisms [152] in the context of MPSS.

Efficient Social Decision

For each phone i, define its private type (parameters) as

$$\theta_i = (\boldsymbol{p}_i(t), \mu_{i,j}(t), j \in \mathcal{N}_i(t), 1 \leq t \leq \mathrm{t_{end}}) \in \Theta_i$$

where Θ_i represents the set of all possible types θ_i. Denote the private types of all phones as $\boldsymbol{\theta} = (\theta_1,\ \ldots,\ \theta_{|\mathcal{N}|}(t)) \in (\Theta_1 \times, \ldots\ \times \Theta_{|\mathcal{N}|}) = \boldsymbol{\Theta}$, where the type space $\boldsymbol{\Theta}$ represents the set of all possible $\boldsymbol{\theta}$.

Let $\mathbf{x}(\boldsymbol{\theta})$ represent the joint rate control and routing decisions of the MPSS during the whole time horizon $1 \leq t \leq \mathrm{t}_{\mathrm{end}}$.

$$\begin{aligned}\mathbf{x} &= (\boldsymbol{x}_i(t),\ i \in \mathcal{N},\ 1 \leq t \leq \mathrm{t}_{\mathrm{end}}) \\ &= (r_i(t), f_{i,j}(t), i \in \mathcal{N}, j \in \mathcal{N}_i(t), 1 \leq t \leq \mathrm{t}_{\mathrm{end}}) \\ &\in \mathcal{X}\end{aligned}$$

where $\mathcal{X}$ represents the set of all possible rate control and routing decisions. It is easy to see that the gross profit of each phone $\varphi_i,\ i \in \mathcal{N}$ depends on its private type θ_i and the decision $\mathbf{x}$. Therefore, we can rewrite φ_i as $\varphi_i(\mathbf{x}, \theta_i)$. In mechanism theory, The function $\mathbf{x} : \boldsymbol{\Theta} \to \mathcal{X}$ is called a *social decision.*

Definition 1 [Efficient Social Decision] *A decision* $\mathbf{x}_{opt}$ *is said to be efficient if*

$$\sum_{i \in \mathcal{N}} \varphi_i(\mathbf{x}_{opt}, \theta_i) \geq \sum_{i \in \mathcal{N}} \varphi_i(\mathbf{x}, \theta_i) \tag{6.14}$$

for all $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ *and for all* $\mathbf{x} \in \mathcal{X}$. It can be seen that the sensing rate control and routing decisions made by OptMPSS is the efficient social decision when $V \to \infty$.

In order to make an efficient social decision $\mathbf{x}(\widehat{\boldsymbol{\theta}})$ in a centralized way, each phone is asked to report its type, denoted as $\widehat{\theta}_i$,

$$\widehat{\theta}_i = (\widehat{\boldsymbol{p}}_i(t), \widehat{\mu}_{i,j}(t), j \in \widehat{\mathcal{N}}_i(t), 1 \leq t \leq \mathrm{t}_{\mathrm{end}}) \in \Theta_i$$

to the server S, where $\widehat{\boldsymbol{\theta}} \in \boldsymbol{\Theta}$ represents the reported types of all phones.

Since each phone user i exhibits strategic behaviors in reality, he or she may be untruthful and report a type value $\widehat{\theta}_i$ that is different from the real type (i.e. $\widehat{\theta}_i \neq \theta_i$), in order to derive an alternative social decision $\mathbf{x}'(\widehat{\boldsymbol{\theta}})$ that results in a better gross profit $\varphi_i(\mathbf{x}'(\widehat{\boldsymbol{\theta}}), \theta_i) > \varphi_i(\mathbf{x}_{opt}, \theta_i)$.

Tax, Subsidy, Net Profit, and Server Profit

In order to make an efficient social decision, server S introduces a monetary transfer function $\boldsymbol{\lambda} : \boldsymbol{\Theta} \rightarrow \mathbb{R}^{|\mathcal{N}|}$

$$\boldsymbol{\lambda}(\widehat{\boldsymbol{\theta}}) = (\lambda_1(\widehat{\boldsymbol{\theta}}),\ ..., \lambda_{|\mathcal{N}|}(\widehat{\boldsymbol{\theta}})) \tag{6.15}$$

to encourage the phone users to report their true types. Based on the announcement of a phone i's type $\widehat{\theta}_i(t)$, the function $\lambda_i(\widehat{\theta}_i)$, where this is negative this represents a *tax* that is imposed on phone i, or where positive a *subsidy* is paid to i. The combined social decision and monetary transfer function $(\mathbf{x}(\widehat{\boldsymbol{\theta}}), \boldsymbol{\lambda}(\widehat{\boldsymbol{\theta}}))$ is referred to as the *social choice function* [91]:

$$g : \Theta \rightarrow \mathcal{X} \times \mathbb{R}^{|\mathcal{N}|} \tag{6.16}$$

As a result, the **net profit** of each phone user is defined as

$$u_i(\theta_i, \mathbf{x}(\widehat{\boldsymbol{\theta}}), \lambda_i(\widehat{\boldsymbol{\theta}})) = \varphi_i(\mathbf{x}(\widehat{\boldsymbol{\theta}}), \theta_i) + \lambda_i(\widehat{\boldsymbol{\theta}}) \tag{6.17}$$

In our MPSS model, the time-average **server profit** u_S can be formally defined as

$$u_S = (1 - \alpha) \sum_{i \in \mathcal{N}} v_i(\overline{r}_i) - \sum_{i \in \mathcal{N}} \lambda_i(\widehat{\boldsymbol{\theta}}) \tag{6.18}$$

VCG Mechanisms

A direct revelation mechanism is defined as $(g, \boldsymbol{\Theta})$, with a strategy (type) space $\boldsymbol{\Theta}$ and social choice function g. A mechanism defines a non-cooperative game with incomplete information as each phone has no knowledge of the types of other phones.

Definition 2 [Incentive Compatibility] *A direct revelation mechanism $g(\boldsymbol{\theta})$ is dominant strategy incentive compatible, if the reported $\theta_i(t)$ is a dominant strategy for each phone $i \in \mathcal{N}$:*

$$u_i(g(\theta_i, \theta_{-i}), \theta_i) \geq u_i(g(\widehat{\theta}_i, \theta_{-i}), \theta_i), \forall \theta_{-i}, \theta_i, \widehat{\theta}_i \neq \theta_i$$

, *where* $\theta_{-i}(t)$ *represents the types of all other phones* $j \in \mathcal{N} - \{i\}$. *In this case we say the social choice function is implemented in ex-post Nash equilibria.*

The Vickrey-Clarke-Groves (VCG) mechanism is well-known for the desired property of incentive compatibility [91]. The social decision rule of the VCG mechanism is given by

$$\mathbf{x}^{vcg}(\widehat{\boldsymbol{\theta}}) = \underset{\mathbf{x}(\widehat{\boldsymbol{\theta}}) \in \mathcal{X}}{\arg\max} \sum_{i \in \mathcal{N}} \varphi_i(\mathbf{x}(\widehat{\boldsymbol{\theta}}), \widehat{\theta}_i) \tag{6.19}$$

and the monetary transfer function of each phone $i \in \mathcal{N}$ is:

$$\lambda_i^{vcg}(\widehat{\theta}_i) = \underbrace{\sum_{j \neq i} \varphi_j(\mathbf{x}^{vcg}(\widehat{\boldsymbol{\theta}}), \widehat{\theta}_j)}_{\text{(a)}} - \underbrace{\max_{\mathbf{x}_{-i} \in \mathcal{X}_{-i}} \sum_{j \neq i} \varphi_j(\mathbf{x}_{-i}, \widehat{\theta}_j)}_{\text{(b)}} \tag{6.20}$$

where the term (a) corresponds to the gross profit of all phones excluding i (i.e. all phones in $\mathcal{N} - \{i\}$) when an efficient social decision has been made, and term (b) represents the maximum global gross profit achievable for all phones in $\mathcal{N} - \{i\}$, without i's presence in the MPSS. Therefore, the monetary transfer $\lambda_i^{vcg}(\widehat{\theta}_i)$ represents the impact (either loss or increase) in value that is imposed on all other individuals (i.e. marginal social impact) due to the social decision that has been updated resulting from i's presence in the MPSSs.

6.5 The BMT Algorithm

By applying distributed mechanism design to OptMPSS, we develop BMT, an on-line and fully distributed algorithm that calculates the marginal social impact (for computing the VCG tax) of each phone in parallel with the operations of OptMPSS.

6.5.1 Distributed Operations of BMT

During the complete time horizon $1 \leq t \leq \mathrm{t_{end}}$, the BMT algorithm operates as follows:

1. Initialization. At beginning of slot $t = 1$.

- *The server*: S broadcasts the set $\mathcal{N}$, revenue function $\alpha v_i()$, and system parameters V and $\mathrm{r}^{\max}$ to each phone $i \in \mathcal{N}$.

- *Mobile phones*: Besides storing its data queue holding data, each phone i initializes a *virtual queue length* (a non-negative integer number) $Q_i^{-j}(t)$ for each of all other phones $j \in \mathcal{N},\ j \neq i$, where $Q_i^{-j}(t)$ means the queue length of phone i without phone j's presence. The initial lengths of all virtual queues are set as zero.

2. At each slot $1 \leq t \leq \mathrm{t}_{\mathrm{end}}$

Distributed Sensing and Routing. Each phone $i \in \mathcal{N}$ adopts the OptMPSS algorithm for optimal distributed sensing and data forwarding.

Distributed Marginal Social Impact Computation. In parallel, each phone $i \in \mathcal{N}$ computes the *virtual* sensing rate $r_i^{-j}(t)$ and the *virtual* cost $cost_i^{-j}(t)$ for each of all other phones $j \in \mathcal{N},\ j \neq i$ based on the corresponding virtual queue length $Q_i^{-j}(t)$,

$$r_i^{-j}(t) = \min(\mathrm{r}_{\max}, \alpha v_i'^{-1}(\frac{Q_i^{-j}(t) + p_i^s(t)}{V})) \tag{6.21}$$

and

$$cost_i^{-j}(t) = p_i^s(t) r_i^{-j}(t) + \sum_{k \in \mathcal{N}_i(t),\ k \neq j} f_{i,k}^{-j} c_{i,k}^t(t) \tag{6.22}$$

where the *virtual* forwarding rate is

$$f_{i,k}^{-j}(t) = \begin{cases} \mu_{i,k}(t) & \text{if } w_{i,k}^{-j}(t) > 0 \\ 0 & \text{otherwise} \end{cases} \tag{6.23}$$

where for each $k \in \mathcal{N}_i(t),\ k \neq j$, the *virtual* weight is

$$w_{i,k}^{-j}(t) = (Q_i^{-j}(t) - Q_k^{-j}(t))\mu_{i,k}(t) - V p_{i,k}^t(t) \tag{6.24}$$

The virtual queue lengths $Q_i^{-j}(t), \ \forall j \in \mathcal{N} - \{i\}$ are updated as

$$\begin{aligned} Q_i^{-j}(t+1) &= |Q_i^{-j}(t) - \sum_{k \in \mathcal{N}_i(t)} f_{i,k}^{-j}(t)|_+ + r_i^{-j}(t) \\ &\quad + \sum_{k \in \mathcal{N}_i(t)} f_{k,i}^{-j}(t) \end{aligned}$$

The average virtual sensing rates and virtual costs for all $j \neq i, \ j \in \mathcal{N}$ are updated as

$$\overline{r}_i^{-j} = (r_i^{-j} + (t-1)\overline{r}_i^{-j})/t$$
$$\overline{cost}_i^{-j} = (cost_i^{-j} + (t-1)\overline{cost}_i^{-j})/t$$

3. At slot $t = \mathrm{t_{end}}$.

- *Each phone* i reports the average virtual sensing rates $\overline{r}_i^{-j}$ and virtual cost $\overline{cost}_i^{-j}$ for all $j \in \mathcal{N}, \ j \neq i$ to the server.
- *The server* The server can compute the VCG tax

$$\begin{aligned} \lambda_i^{vcg} &= \sum_{j \in \mathcal{N} - \{i\}} (\alpha v_j(\overline{r}_j^{-i}) - \overline{cost}_j^{-i}) \\ &\quad - \sum_{j \in \mathcal{N} - \{i\}} (\alpha v_j(\overline{r}_j) - \overline{cost}_j) \end{aligned}$$

  of each phone $i \in \mathcal{N}$. Finally, the server makes a payment of $\mathrm{t_{end}}(\alpha v_i(\overline{r}_i) + \lambda_i^{vcg}(\widehat{\theta}_i))$ to phone i.

From a global view of point, the BMT algorithm runs in total one *real* and $|\mathcal{N}|$ *virtual* OptMPSS algorithms in parallel during the time horizon of the MPSS: the real OptMPSS algorithm makes the actual sensing rate and routing decisions at each slot, while $|\mathcal{N}|$ virtual OptMPSS algorithms simulate $\mathcal{N}$ virtual marginal societies with absence of each phone $i \in \mathcal{N}$ to compute the final tax or subsidy for each phone in a fully distributed way.

Remark 2. It worth noting that the $|\mathcal{N}| - 1$ *virtual* queue lengths maintained in each phone are integer numbers rather than real data packet queues, which results in negligible storage

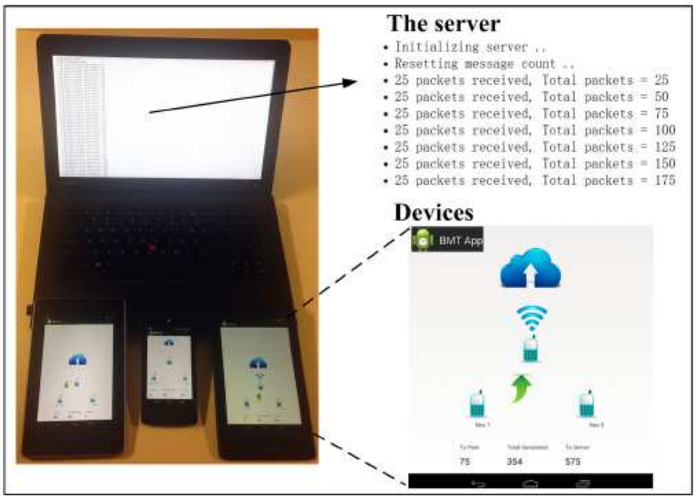


Figure 6.2: Experiment Prototype Illustration.

overheads for mobile phones (e.g. only several KB storage overhead for a MPSS with thousands of phones with several GB RAMs). In addition, BMT requires each node to transmit $O(|\mathcal{N}|)$ bytes of information (its real and maintained virtual queue backlogs) to its current neighbors only, while also performing $O(|\mathcal{N}|)$ simple arithmetic calculations. This is still realistic for today's smart phones using short-range radios such as WiFi direct that can achieve up to 250Mbps data transmission rate. Due to its distributed operations and light overheads, BMT has a great potential to be applied in large-scale MPSS.

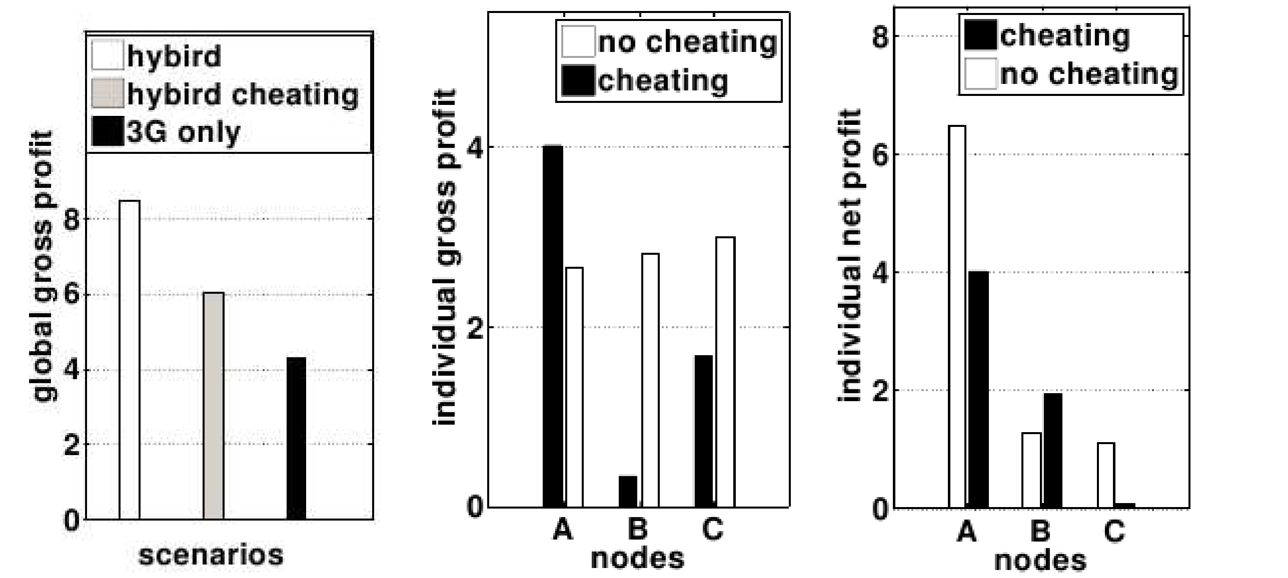
Proof of Asymptotic Incentive Compatibility of BMT can be found in Appendix C.

6.6 Evaluation

In this section, we evaluate the performance of the BMT algorithm via both prototype experiments and simulations using real human mobility traces.

6.6.1 Experiments Based on Android Device

We implemented the BMT algorithm in Android OS 4.3, and developed an application called BMT App. We constructed a proof-of-concept MPSS with three WiFi-direct enabled Android



(a) (b) (c)

Figure 6.3: Experiment results: (a) time-average global gross profits for the three experiments. (b) and (c) shows the impact of device A's cheating action on the time-average individual gross and net profits of every device respectively.

devices (i.e. a Nexus 5 phone and two Nexus 7 tablets) and a server implemented in NODE.JS (http://nodejs.org/), as shown in Fig.6.2. The duration of a slot was set as two seconds and the duration of each experiment was 10 minutes. At the each slot, the BMT App run a discovery phase to update the temporary neighbor table (e.g. the server and nearby devices), and then performed the sensing, routing, and marginal social impact computation tasks as defined by the BMT algorithm. Each device was held by a researcher moved around our lab.

We use the revenue function $v_i(\overline{r}_i) = 4\ln(1+\overline{r}_i)$ for each $i \in \mathcal{N}$, and set the system parameters $V = 100$ and $\alpha = 0.5$. We use WiFi direct as the short-range radios. The channel capacities of all wireless radios are set at 25 packets per slot. The sensing prices and transmission prices (in credits per data packet) for the data sent by WiFi direct were set as 0.1, for all three Android devices A, B, and C. The 3G transmission prices of A, B and C were set as 0.1, 1, and 1.5 respectively[1].

As shown in Fig.6.3 (a), the time-average global gross profit of MPSS using hybrid 3G and WiFi direct communications is approximately twice of that using 3G radios alone, which demonstrates

[1]This implies that A may have an unlimited mobile data budget, while B and C may adopt a limited monthly contract or 'pay as you go' tariff rate.

that significant performance improvement can be achieved by using hybrid cellular and short-range radio.

In addition, we also evaluated the incentive compatibility of BMT. Since device A had a much lower 3G prices than B and C, the routing decisions made by BMT would require A to relay the sensor data packets collected by B and C (when A passed them) to the server, in order to maximize the global gross profit. However, this would result in an individual gross profits reduction for A (due to the relaying cost), and therefore (the owner of) A may not want to faithfully adopt the distributed routing actions suggested by BMT. To check whether BMT can avoid this, we mimic a quite intuitive cheating behaviour for A, i.e. disabling its WiFi direct radio.

Fig. 6.3 (b) shows that the individual gross profit for each device before and after A's cheating action. It can be seen that this untruthful behavior can indeed improve A's individual gross profit, but results in a significant degeneration of global gross profit for the whole system, as shown in Fig. 6.3(a). As shown in Fig.6.3 (c), however, A eventually missed the opportunity of obtaining approximately 36% more net profit due to cheating. This means that A would be better off relaying sensor data from other devices than misinforming the network. This demonstrates that in practice BMT can achieve the highly desired incentive compatibility property. In addition, the server profit and net profit of each device and server profit were positive in all experiments , which demonstrates that BMT can achieve individual rationality and server profitability in practice.

6.6.2 Trace-driven Simulations

To evaluate the practical performance of BMT at scale, we established simulations using the real human trace collected from Infocom05 (41 nodes for 3 days) [5]. In all simulations, each phone has both 3G, WiFi, and WiFi direct radios. When a phone meets a free WiFi router, it sends data through the WiFi radio rather than 3G. In each simulation, a power-law distributed random variable was assigned to each phone to simulate the heterogeneous free WiFi access probability across phones, observed from real human mobility traces [38]. The sensing and

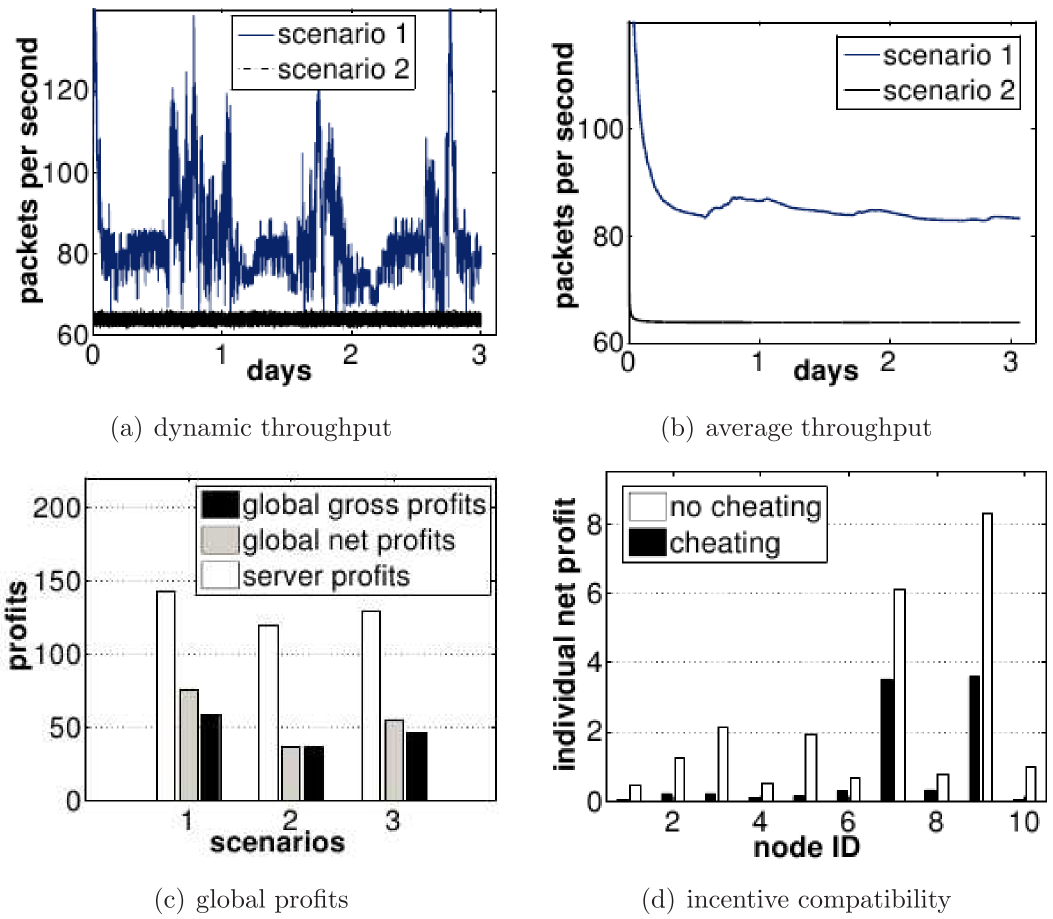
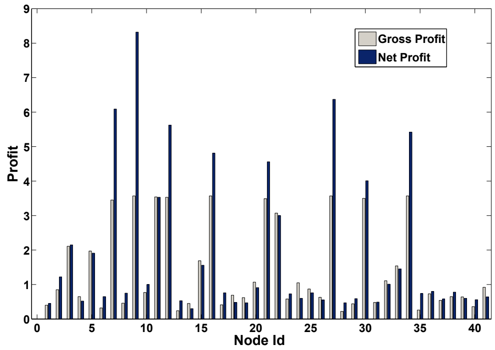
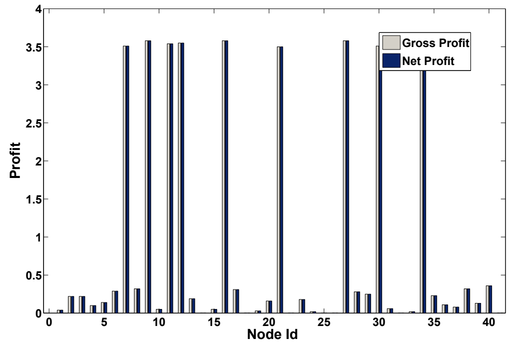


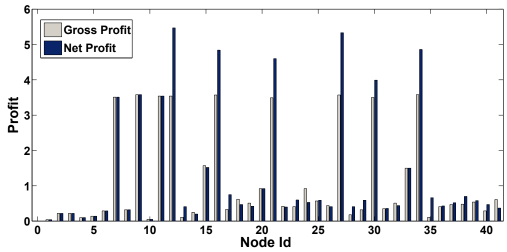
Figure 6.4: Simulation results of BMT.



(a) Scenario 1: Wifi



(b) Scenario 2: No Wifi



(c) Scenario 3: Cheating

Figure 6.5: Individual Profits of Nodes.

WiFi/Wifi-direct transmission prices for each phone were dynamically set between 0 and 0.01 for each slot (representing channel quality variation) at each slot, while 3G transmission price of each phone was randomly set between 0.1 and 0.5 at the beginning of each simulation and remained constant over slots. We set $r_{max} = 5$ and the duration of a slot as one second. All other parameters in the simulations were set the same as in the prototype experiment represented earlier.

We run three simulations for three different scenarios:

- **Scenario 1:** MPSS with all wireless radios which show the normal operation of the network where each user is willing to participate and data can be forwarded either using Wifi Direct radio or cellular radio. The system can achieve optimal performance through utilization of all available communication resources.
- **Scenario 2:** MPSS without WiFi direct where users are not willing to cooperate with each other and all data is forwarded using only cellular radios. This results in increased cost of the system as cellular communication is expensive.

- **Scenario 3:** cheating actions in MPSS with all wireless radios, where the users of phone 1 to 10 try to hide their WiFi direct abilities in effort to increase their own profit without cooperating with other users. The users who are involved in cheating report that they only have cellular radio therefore they cannot receive or forward data to other users.

Fig. 6.4 (a) and (b) show the dynamic and time-average throughput (i.e. total sensing data packets produced by all phones at each slot) of MPSS respectively, which demonstrate that the BMT algorithm is adaptive to network dynamics (e.g. mobility) and manages to converge to the time-average optimal. This also show higher packet rate when both Wifi Direct radio and cellular radio are used which results in high system performance.

This is also shown in Fig. 6.4 (c), where global profits are higher in scenario 1. This demonstrate that using multi-hop opportunistic short-range communications can significantly improve network performance.

Fig. 6.4 (d) show individual profits of users 1-10 in scenario 1 and 3. It verifies that the users earn less profit in scenario 3 where they cheated than scenario 1 under normal network operations. This show that BMT can achieve incentive compatibility and server profitability in practice.

Fig. 6.5 shows individual gross and net profit in above three scenarios. The gross and net profit of each device were positive in all scenarios , which demonstrates that BMT achieves individual rationality and effectively incentivizes the users to participate in the network.

6.7 Conclusion

In this chapter we investigate a cost-effective data collection solution for Mobile Phone Sensing Systems (MPSS) that utilize hybrid cellular and opportunistic short-range wireless communications. We formulate a stochastic optimization problem for mobile sensor data collection, and develop OptMPSS, a scalable joint sensing rate control and routing algorithm to solve the formulated optimization problem in a fully distributed and scalable way. In order to encourage phone users to faithfully apply the OptMPSS algorithm's control suggestions, we propose BMT, a joint networking and taxing scheme, based on combing Lyapunov stochastic optimization and distributed mechanism design theories. We prove that BMT achieves *asymptotical optimality* and *incentive compatibility*.

In order to evaluate the practical performance of BMT, we developed BMT App, an Android application that implements BMT algorithm for WiFi-direct-enabled devices. Trough real-world experiments and realistic trace-driven simulations, we demonstrate that BMT can efficiently exploit low-cost short-range communications, which significantly improves the global gross profit of the MPSS (around 100%). In addition, we also verify the incentive compatibility of BMT by mimicking potential cheating behaviors of phone users. Evaluation results show that the sensing and routing actions suggested by BMT are the best choice for each individual phone user. In addition, evaluation results also demonstrate that our approach can achieve *individual rationality* and *server profitability* in practice.

In the future, we plan to construct a larger-scale practical MPSS with real phone users to test BMT App in a non-lab environment, and investigate how to achieve strict individual rationality and server profitability guarantees in theory.

Chapter 7

Conclusion

7.1 Summary of Achievements

There are many cases in which the classic WSN deployment, consisting of a single static sink node, is infeasible. An alternative is to use mobile sinks, such as wireless devices carried by people, robots, or vehicles to collect data from statically-deployed sensors in an opportunistic way. For example, a smart sustainable city typically requires deployments of sensor nodes over a large space to ensure sensing coverage. This means that the sensors are sparsely spread and the network may even be disconnected. Mobile sinks can solve this problem by visiting some nodes in network to collect data. These nodes close to the trajectory of mobile sink act as relay nodes for other nodes in the network. All the other nodes in the network can forward their data to relay nodes in multi-hop manner. The detection of such relay nodes is important in WSNs with mobile sinks.

Chapter 3 propose a novel routing metric CA-ETX to estimate the packet transmission delay caused by both link unreliability and intermittent connectivity. An opportunistic shortest path routing scheme, OSPR, is also developed to demonstrate the efficiency of CA-ETX. CA-ETX can seamlessly and synchronously work with ETX, illustrating that existing ETX-based routing protocols, such as standard CTP and IETF IPv6 Routing Protocol RPL, can be easily applied to WSN-MSs using CA-ETX.

We developed a novel joint routing and scheduling algorithm Opportunistic Backpressure Collection (OBC) that integrates mobility-awareness for opportunistic data collection in WSNs with mobile sinks (WSN-MSs). OBC requires each node to maintain only a one-hop neighbour table by periodically broadcasting a small one-hop beacon (several bytes). The communication overhead of OBC is therefore O(1) with respect to the number of sensor nodes and mobile sinks, which demonstrates that it has a much better scalability than current mobility-aware schemes that require explicit routing structure maintenance [108, 115, 123]. In addition, OBC does not require mobility prediction. It is well-known that backpressure algorithms suffer from large end-to-end delay and unnecessary packet transmissions (which significantly increase the energy cost) due to routing loops. By using a lightweight mobility-aware scheme, OBC significantly mitigates this problem and therefore achieves a significantly improved performance in end-to-end delay and energy consumption. We prove the throughput optimality of OBC and also implement OBC in TinyOS 2.1 [1] and a realistic WSN simulator, Castalia [3]. The results of both the testbed experiments and extensive simulations show that the shortest path routing with CA-ETX and OBC can achieve significant performance improvements in terms of end-to-end delay, storage overheads, and energy consumption, compared with state-of-the-art approaches.

After sensing or collecting data from sensors, a mobile phone can send this data via the Internet using two approaches; (1) To cellular base station through expensive and bandwidth limited 3G or 4G cellular communication radios. (2) To a static base-station (e.g. WiFi router) via other mobile relay nodes through short-range opportunistic communication, which has a great network capacity.

Also, the owners of these devices may not be willing to collect sensor data (i.e. act as a mobile sink) and forward this data to others devices or Internet.

To address these problem, we considered a Mobile Phone Sensing System (MPSS) consisting of mobile phone for sensing and communication in **Chapter 4.** We devolop a network architecture to provide cost-effective networking service for MPSS by seamlessly integrating short-range communication with 3G/4G. This architecture considers properties of data packet and performs cost-benefit analysis to chooses the best approach for forwarding of data packet. Based on the guiding principles of the proposed network architecture, we develop a joint pricing and routing scheme aiming to provide cost-effective networking services for both real-time and delay-tolerant data of MPSS applications. To achieve effective incentivization for the phone users, we modelled the MPSS as an economic system where sensor data is a commodity produced by sensors and end users are consumers whereas everyone can contribute to the system by providing goods or services to achieve benefits. This scheme achieves throughput optimality for big sensor data transmission and minimizes costs of phone users. show how the performance of MPSS is affected by not only urban sensor data properties and communication infrastructures, but as also the social and economic behaviours of people. Through simulations, we demonstrate that our scheme adapts to user's preferences and privacy. Furthermore, it is able to self configure, self adapts and self heals under different dynamic network conditions.

Although there are many wireless devices around us in the real world, every owner of these devices may not be a very good choice to relay the sensor data. A postman can be a good relay, who is popular and visit sink frequently whereas a geek will be bad relay, who is lazy to move or far from any sink.

To solve this problem, **Chapter 5** exploit social behaviors of humans for efficient data transmission. We proposes a new sensing architecture, WSN with Human Relays (WSN-HR), consisting of a static sensor network combined with mobile phone users for delay-tolerant sensing applications in future smart and sustainable cities. By integrating social and economic behaviours of citizens into sensor networking, a joint rate control, routing, and resource pricing algorithm (OBSEA) is developed to reduce end-to-end delay and incentivise the participation of phone users. By exploiting mobility patterns and the underlying social networks of human relays, a novel data forwarding metric Sink-Aware (SA) centrality is proposed to measure the potential sensor data forwarding ability of mobile relays. SA centrality significantly improves the performancu of the network in terms of end-to-end delay. We also established a virtual economic network for sensor data production and trading to incentivise people to serve as data relays using their phones. Each mobile relay acquires profit by dynamically adjusting the selling price of its maintained sensor data and then trading (transmitting and receiving) data with other nodes opportunistically at each moment of contact. We formalise a optimisation problem to maximise the global social profits of all nodes in a WSN-HR. We *do not* make any probabilistic/stochastic assumptions (e.g. specific probability distributions or ergodicity) for the network conditions (e.g. mobility, topology, and wireless channel), and thus is suitable for arbitrary dynamic evolution process of WSN-HR. The lightweight OBSEA solves the problem using only current and local information. This means that OBSEA is *fully distributed* and *does not require* any prediction capacity, thereby maximising the practical application of the work.

We evaluate the performance of OBSEA using the Castali simulator [3] and a realistic mobility model [199]. Simulation results demonstrate that OBSEA is adaptive to different network settings and outperforms both pure backpressure routing and pure social-aware forwarding schemes, in terms of global social profit, data buffer efficiencies, and end-to-end delay. In addition, the results show that a 'win-win situation' (positive profits) can be achieved by both the network and all mobile phone owners.

Every mobile owner can attempt to maximize his profits through his strategic actions. They may use information learned from participation in MPSS to manipulate the network for their benefit. e.g. by dropping packets, reporting wrong information regarding their cellular cost or turning off short-range connectivity etc. This behaviour may increase the benefits of the user but it would result in significant performance degradation of the network.

In chapter6, we developed a well designed mechanism that will incentivize users into behaving according to networks wishes. We considered a MPSS that utilize hybrid cellular and opportunistic short-range wireless communications for mobile sensor data collection. To solve the data collection problem, we first develop a joint sensing, rate control and routing algorithm OptMPSS, which is fully distributed and scalable. OptMUSS maximizes the *global gross profit*, i.e. the total financial rewards of all phone users after costs incurred by performing the sensing and transmission tasks are deducted. OptMPSS is fully distributed and its operation depend on the local information of its neighbours such as queue length. In MPSS, all parameters local to each phone are private and not observable to other phones and the server. Consequently, phone users can subvert the OptMUSS system by miscommunicating their local parameters.

We develop a joint networking and taxing scheme BMT by applying distributed mechanism design to OptMPSS. BMT calculates the social impact for computing the tax of each phone in parallel with the operations of OptMPSS. BMT provides the subsidy to a phone user if he has positive effect on the network otherwise it imposes the tax on the phone user. BMT ensures incentive compatibility, so that all of the participants are better off when they truthfully reveal any private parameter required in OptMPSS operation.

To evaluate the practical performance , we developed an Android application BMTApp that implements BMT algorithm for WiFi-direct-enabled devices. To evaluate large-scale scenarios, we established simulations using the real human mobility traces. The results show that that BMT can improves the global gross profit of the MUSS up to 100 percent by exploiting low-cost short-range communications. In addition, we also verify the incentive compatibility of BMT by introducing potential strategic behaviours of some phone users in the network. Evaluation results show that the sensing and routing actions suggested by BMT are the best choice for each individual phone user and they will not be able to increase their profit through reporting wrong information.

7.2 Future Work

This section will outline some important and promising extensions of the work presented in this thesis

7.2.1 Utilizing Multiple Short-range radios

Current phones can have multiple short-range radios such as LTE direct, WiFi direct, and Bluetooth, which have different properties in terms of neighbor discovery time, throughput, energy consumption, and transmission range. In order to fully exploiting performance gain by using short-range communications, all possible available radios rather than a single one should be considered. For instance, WiFi direct has a high throughput but a large energy cost for neighbor discovery. If we use WiFi direct as the only short-range radio, the battery will run out of energy fast. To solve this issue, we can use bluetooth which has a much less energy cost for the always-on neighbor discovery, and WiFi direct can be used for data transmission whenever a contact opportunity is discovered by bluetooth. Therefore, it is highly desirable to develop optimal networking and incentivisation algorithms for such multi-radio MPSS.

7.2.2 Cost-effective Networking Support Beyond MPSS

It can been seen that our social/economic-aware algorithms and protocols for MPSS can be easily extended to general networks consist of human-carried devices. Besides sensor data, all kinds of Internet traffic can be supported by using the hybrid short-range and long-range communications. This will not only significantly reduce the cost of individual users, but also the congestion of the cellular networks. To this end, three key issues should be addressed in the future work.

- Classification schemes should be designed to rank different types of data traffics according to their delay tolerance.
- Algorithms with strong Quality of Service (QoS) guarantees (e.g. end-to-end delay and bandwidth) should be developed to support real-time applications such as HTTP video streams.
- Develop theoretical-optimal pricing algorithm and seamlessly combine this with the existing price policy provided by Internet service providers(ISPs).

7.2.3 Secure and Rapid Communication

With the increase in short range communication capabilities of smart phones, a large number of mobile devices can form an ubiquitous opportunistic network: Contact duration can not be precisely determined for mobile devices that meet opportunistically for an undetermined time. It ranges from few seconds for pedestrian users or few minutes for passengers of a train to hours for devices held by adjacent colleagues. Then connections between nearby devices have to be established for short range data communication. How to effectively establish secure short range communication connections between mobile devices, or device pairing, is a fundamental but challenging problem to solve in future networks based on mobile devices.

we will focus on the device pairing problem in opportunistic networks based on mobile devices to develop a device pairing mechanism that has the following properties:

- **Secure.** Security has the first priority in communications involving personal devices, e.g. smart phones or tablets. Attackers may disguise themselves as potential relays or sinks which can compromise the integrity of the system. To solve this issue, both devices should be aware of whom it is communicating with, therefore mutual authentication is necessary before the short range connection is established. Furthermore, short-range communication should be confidential for any third party trying to eavesdrop or modify transmitted data. Therefore a shared secret is necessary for content encryption.

- **Rapid.** In opportunistic scenarios, devices are encountered for an undetermined time. The short range connection may be intermittent, and reconnection could be frequent. The device pairing should be as quick as possible for devices to exchange data in time, and to efficiently reconnect to alternative devices.

Appendix A

Theorems for Chapter 3

The following appendix presents the results of the joint work between Shusen Yang and Usman Adeel.

Theorem below demonstrates the throughput optimality of OBC.

Theorem A.1. *Given any arriving traffic (sensing rate vector)* $\boldsymbol{r}$ *such that* $\boldsymbol{r} + \varepsilon \in \Lambda$ *for any* $\epsilon > 0$*, the OBC algorithm can stabilise all queues, i.e.*

$$\limsup_{K\to\infty} \frac{1}{K} \sum_{t=1}^{K} \sum_{x\in N} \mathbb{E}[Q_x(t)] < \infty$$

Proof of Theorem A.1. Let the N-dimensional vector $\boldsymbol{Q}(t)$ be the queue backlogs of all nodes in the WSN-MSs at slot t. Define $\triangle Q_x(t) = Q_x(t+1) - Q_x(t)$. According to (6.3), we have$\triangle Q_x(t) \leq r_x + f_x^{in}(t) - f_x^{out}(t)$. Define a constant value $W = \frac{1}{\varphi^{\min}}|N|(\mathrm{r}^{\max} + 2|N|\mathrm{c}^{\max})^2$ where $\mathrm{r}^{\max} = \max_{x\in N^s} r_x$. We then define the Lyapunov function $V(t) = \sum_{x\in N} Q_x^2(t)/\varphi_x$ and consider the its 1-slot drift:

$$\begin{aligned}
&\triangle_1 V(t)\\
&= V(t+1) - V(t)\\
&= \sum_{x\in N}(2Q_x(t)\triangle Q_x(t) + \triangle Q_x^2(t))/\varphi_x\\
&\le W + 2\sum_{x\in N} Q_x(t)\triangle Q_x(t)/\varphi_x\\
&\le W + 2\sum_{x\in N} Q_x(t)(r_x + f_x^{in}(t) - f_x^{out}(t))/\varphi_x\\
&= W + 2\sum_{x\in N}(\frac{Q_x(t)}{\varphi_x}r_x - \sum_{y\in N_x(t)}(\frac{Q_x(t)}{\varphi_x} - \frac{Q_y(t)}{\varphi_y})f_{x,y}(t))
\end{aligned}$$

It is clear that OBC choose $f_{x,y}(t), \forall t \ge 0$ to minimize the right-hand side of above inequality over all possible other algorithms. Hence we have

$$\triangle_1 V(t) \le W + 2\sum_{x\in N}\frac{1}{\varphi_x}Q_x(t)(r_x + \widetilde{f_x^{in}}(t) - \widetilde{f_x^{out}}(t)) \tag{A.1}$$

where $\widetilde{f_x^{in}}(t)$ and $\widetilde{f_x^{in}}(t)$ are the routing and scheduling decision made by any other algorithm $\widetilde{\xi}$ which is independent of queue backlogs.

Since the channel capacity $\boldsymbol{c}(t)$ is a discrete finite state ergodic Markov chain, we use a sequence $T_i, i \ge 0$ to represents recurrence times to the initial state $\boldsymbol{c}(0)$. It is clear that $T_i, i \ge 0$ is a i.i.d. sequence with $\mathbb{E}[T_i] = 1/\pi_{\boldsymbol{c}(0)}$. Also, it is known that the first and second moments of sequence T_i are finite, which are denoted as $\mathbb{E}[T]$ and $\mathbb{E}[T^2]$ respectively. Finally, we define $s_i = \sum_{\tau=0}^{i-1} T_\tau$, i.e. the time of the ith revisitation to channel state $\boldsymbol{c}(0)$. Consider the variable T_i-slots drift of the Lyapunov function

$$
\begin{aligned}
&\triangle_{T_i} V(s_i) \\
&= V(s_{i+1}) - V(s_i) \\
&= \sum_{t=s_i}^{s_i+T_i-1} (V(t+1) - V(t)) \\
&\leq \frac{W}{2}(T_i^2 + T_i) \\
&+2 \sum_{x \in N} \frac{Q_x(s_i)}{\varphi_x} \sum_{t=s_i}^{s_i+T_i-1} (r_x + \widetilde{f_x^{in}}(t) - \widetilde{f_x^{out}}(t))
\end{aligned}
\tag{A.2}
$$

where the equality is because of (A.1) and the fact for any $s_i \leq t \leq s_i + T_i - 1$,

$$
|Q_x(t) - Q_x(s_i)|/\varphi_x \leq (t - s_i)W/|N| \tag{A.3}
$$

Now we consider the conditional expectation of the variable T_i-slots drift (A.2)

$$
\begin{aligned}
&\mathbb{E}[\triangle_{T_i} V(s_i) | \boldsymbol{Q}(s_i)] \\
&\leq \mathbb{E}[\frac{W}{2}(T_i^2 + T_i) + 2 \sum_{x \in N} \frac{Q_x(s_i)}{\varphi_x} \sum_{t=s_i}^{s_i+T_i-1} (r_x \\
&\qquad\qquad + \widetilde{f_x^{in}}(t) - \widetilde{f_x^{out}}(t)) | \boldsymbol{Q}(s_i)] \\
&=_{(a)} \frac{W}{2}(\mathbb{E}[T^2] + \mathbb{E}[T]) \\
&\quad +2 \sum_{x \in N} \frac{Q_x(s_i)}{\varphi_x} \mathbb{E}[\sum_{t=s_i}^{s_i+T_i-1} (r_x + \widetilde{f_x^{in}}(t) - \widetilde{f_x^{out}}(t))] \\
&=_{(b)} \frac{W}{2}(\mathbb{E}[T^2] + \mathbb{E}[T]) \\
&\quad +2 \sum_{x \in N} \frac{Q_x(s_i)}{\varphi_x} \mathbb{E}[T](r_x + \mathbb{E}[\widetilde{f_x^{in}}(t) - \widetilde{f_x^{out}}(t))]
\end{aligned}
\tag{A.4}
$$

where the equality (a) is because both recurrence time T_i and the algorithm $\widetilde{\xi}$ are independent of queue backlogs $\boldsymbol{Q}(s_i)$; and the equality (b) is based on the renewal reward theory. Consider

(A.4), $\varphi_x \leq \varphi^{\min}$ and the fact that $\boldsymbol{r}+\varepsilon$ is inside the capacity region, we have

$$\begin{aligned}&\mathbb{E}[\triangle_{T_i} V(s_i)|\boldsymbol{Q}(s_i)]\\&\quad\leq \frac{W}{2}(\mathbb{E}[T^2]+\mathbb{E}[T]) - \frac{2\varepsilon\mathbb{E}[T]}{\varphi^{\min}}\sum_{x\in N} Q_x(s_i)\end{aligned} \tag{A.5}$$

Taking expectations of the above, summing the resulting telescoping series over $i \in \{0, 1, ...I - 1\}$, dividing by $2\varepsilon\mathbb{E}[T]/\varphi^{\min}$,rearranging the terms, and using the fact that $V(0) = 0$ and $V(s_i \geq 0)$, $\forall i$, we have:

$$\sum_{i=0}^{I-1}\sum_{x\in N}\mathbb{E}[Q_x(s_i)] \leq \frac{IW\varphi^{\min}(\mathbb{E}[T^2]+\mathbb{E}[T])}{4\varepsilon\mathbb{E}[T]} \tag{A.6}$$

Consider (A.3), we have

$$\sum_{t=s_i}^{s_i+T_i-1}\sum_{x\in N} Q_x(t) \leq T_i\sum_{x\in N} Q_x(s_i) + \frac{\varphi^{\max}W(T_i^2 - T_i)}{2} \tag{A.7}$$

Combine(A.6) and (A.7), and we have

$$\begin{aligned}\sum_{t=0}^{s_i+T_i-1}\sum_{x\in N}\mathbb{E}[Q_x(t)] \quad\leq\quad & \frac{IW\varphi^{\min}(\mathbb{E}[T^2]+\mathbb{E}[T])}{4\varepsilon}\\ &+\frac{I\varphi^{\max}W(\mathbb{E}[T^2]-\mathbb{E}[T])}{2}\end{aligned} \tag{A.8}$$

Let $K = s_i + T_i - 1$, dividing both sides by K, taking an expectation and lim sup over both sides, we have

$$\begin{aligned}&\limsup_{K\to\infty}\frac{1}{K}\sum_{t=0}^{K}\sum_{x\in N}\mathbb{E}[Q_x(t)]\\&\leq \frac{W\varphi^{\min}(\mathbb{E}[T^2]+\mathbb{E}[T])}{4\varepsilon\mathbb{E}[T]} + \frac{\varphi^{\max}W(\mathbb{E}[T^2]-\mathbb{E}[T])}{2\mathbb{E}[T]}\\&< \infty\end{aligned}$$

This completes the proof of Theorem A.1. $\square$

Appendix B

Performance Analysis for Chapter5

The following appendix presents the results of the joint work between Shusen Yang and Usman Adeel.

B.1 Bounded Queues

Memory is a key resource for both sensor nodes and mobile relays. Theorem B.1 below shows that all data queue backlogs are deterministically bounded.

Theorem B.1. *Suppose the initial queue backlogs $Q_x(1) = 0$, $\forall x \in \mathcal{S} \cup \mathcal{R}$, then $Q_x(t)$ is always less than its buffer size $\leq Q_x^{max}$, $\forall t \geq 1$, if V satisfy:*

$$Q_x^{max} \geq \begin{cases} VI'_x(0) + \eta^{\max} & x \in \mathcal{S} \\ \eta^{\max} & x \in \mathcal{R} \end{cases} \tag{B.1}$$

where $\eta^{\max} = \max_{x \in \mathcal{S} \cup \mathcal{R}, t \geq 1} \eta_x(t) \leq |\mathcal{N}| \mathrm{c}^{\max}$. The proof of Theorem B.1 can be found below B.1.1. In practice, $I'_x(0)$ and Q_x^{max}, $x \in \mathcal{S} \cup \mathcal{R}$ are normally fixed and can be determined in advance. It is also easy to estimate $\eta^{\max}$ based on the data rate of wireless transceiver [1].

[1]Although we ignore wireless interference and assume full-duplex wireless radio for simplicity, most current wireless transceivers are half-duplex. Therefore, each node $x \in \mathcal{S} \cup R$ can simply set $\eta^{\max}$ as the data rate (in packets per slot) of its wireless transceiver.

Therefore, the parameter V can be set to guarantee inequality (B.1). For instance, if the utility function of a sensor node x is chosen as $I_x(r_x(t)) = \ln(r_x(t)+1)$, then $I'_x(0) = 1$ and V should be not larger than $Q_x^{max} - \eta^{\max}$.

B.1.1 Bounded Queues and The Supporting Lemma

We first propose Lemma B.1 to support the proof of Theorem B.1.

Lemma B.1. *Considering a sensor node $x \in \mathcal{S}$ at a slot $t \geq 1$, if $Q_x(t) \geq V I'_x(0)$, then the rate controller of OBSEA algorithm sets $r_x(t) = 0$.*

Proof. Since $I_x(r(t))$ is concave, its first derivative $I'_x(r(t))$ is a monotonically decreasing function of $r(t)$. Therefore, we have

$$I_x(r(t)) \leq I_x(0) + I'_x(0) r_x(t), \ \forall\ 0 \leq r_x(t) \leq \mathrm{r}^{\max} \tag{B.2}$$

Considering (B.2) and the objective of the rate controller (5.15), we have for any $0 \leq r_x(t) \leq \mathrm{r}^{\max}$

$$\begin{aligned}
&I_x(r(t)) - Q_x(t) r_x(t)/V \\
&\leq I_x(0) + I'_x(0) r_x(t) - Q_x(t)/V \\
&= I_x(0) - r_x(t)(Q_x(t)/V - I'_x(0)) \\
&\leq I_x(0)
\end{aligned} \tag{B.3}$$

Inequality (B.3) holds only when $r_x(t) = 0$, as $Q_x(t)/V - I'_x(0) > 0$ (the condition of Lemma B.1). Then, the rate controller must set $r_x(t) = 0$ to maximise (5.15). □

Proof of Theorem B.1. We will prove that Theorem B.1 by using mathematical induction. From the supposition of Theorem B.1, we have $Q_x(1) \leq \mathrm{Q}^{\max}$ hold at slot 1. For all $t > 1$, suppose $Q_x(t) \leq Q_x^{max}$ for a slot $t \geq 2$, then there are two cases :

- **Case 1.** $Q_x(t) \le Q_x^{max} - \eta_x(t), x \in \mathcal{S} \cup \mathcal{R}$. Since $\eta_x(t)$ is the maximum possible amount of data that can be injected in to node x at slot t, it is clear that $Q_x(t+1) < Q_x^{max}$ at slot $t+1$, according to (6.3).

- **Case 2.** $Q_x^{max} - \eta_x(t) < Q_x(t) \le Q_x^{max}$, $x \in \mathcal{S} \cup \mathcal{R}$. In this case, the link weight $w_{z,x}(t)$ will be assigned as 0, for all x's instantaneous neighbours $z \in \mathcal{N}_x(t)$, according to (5.17). Hence, no data will be transmitted to x, according to (5.17). Therefore, if x is a mobile relay, then $Q_x(t+1) \le Q_x(t) \le Q_x^{max}, x \in \mathcal{R}$. Further, if x is a sensor node, since $Q_x(t) > Q_x^{max} - \eta_x(t) \ge VI'_x(0) + \eta^{\max} - \eta_x(t) > VI'_x(0)$, we have $r_x(t) = 0$, according to Lemma B.1 and the condition of Theorem B.1. Therefore, we can see that $Q_x(t+1) \le Q_x(t) \le Q_x^{max}$, $x \in \mathcal{S}$. In summary, $Q_x(t+1) \le Q_x^{max}, \forall x \in \mathcal{S} \cup \mathcal{R}$.

Because $Q_x(t+1) \le Q_x^{max}, \forall x \in \mathcal{S} \cup \mathcal{R}$ in both cases, we can conclude that $Q_x(t) \le Q_x^{max}$ for all $t \ge 1$. □

B.2 Social Profits performance

To derive the performance bounds of our OBSEA scheme, we divide the duration $1 \le t \le t_{end}$ into K frames with size of T slots as shown in Figure B.1. We assume that there exists an *ideal* algorithm that fully knows the network information (i.e. the mobility trace and channel capacity) for the future T slots. Based on the future knowledge, the ideal algorithm solves the following optimisation problem:

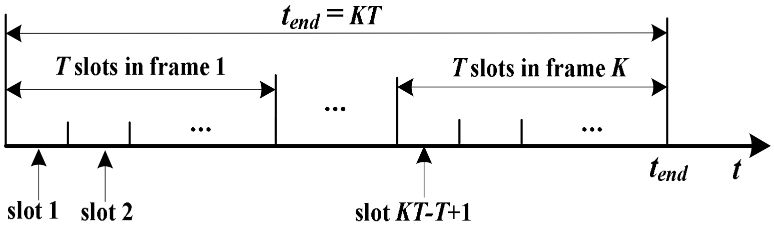


Figure B.1: Illustration of slots and frames.

$$\textbf{max} \quad \frac{1}{T}\sum_{t=kT}^{kT-T+1}\Gamma(t) \tag{B.4}$$

$$\textbf{subject to} \tag{B.5}$$

$$0 \le r_x(t) \le \mathrm{r}^{\max}\ \forall x \in \mathcal{S},\ \forall t \tag{B.6}$$

$$Q_x(t) \le Q_x^{max},\ \forall x \in \mathcal{S}\cup R,\ \forall t \tag{B.7}$$

$$0 \le f_{x,y}(t) \le c_{x,y}(t),\ \forall x \in \mathcal{N},\ y \in \mathcal{N}_x(t),\ \forall t \tag{B.8}$$

$$\frac{1}{T}\sum_{t=kT-T+1}^{kT}\Big(\sum_{y\in\mathcal{N}-\{x\}} f_{x,y}(t) - 1_{\{x\in\mathcal{S}\}}r_x(t) - \sum_{y\in\mathcal{N}-\{x\}} f_{y,x}(t)\Big) \ge 0, \forall x \in \mathcal{S}\cup\mathcal{R}, \forall k \le K \tag{B.9}$$

The objective (B.4) demonstrates that the ideal algorithm optimises the social profits over each frame $1 \le k \le K$. Define the $\Gamma_k^*(T)$ be the optimal social profits of problem (B.4) for the k^{th} T-slot frame. Let $r_x^*(t), x \in \mathcal{S}$ and $f_{x,y}^*(t), x \in \mathcal{N}, y \in \mathcal{N}_x(t)$ respectively be the rate control and routing decisions of the ideal algorithm that achieve $\Gamma_k^*(T)$. Due to the requirement of complete future knowledge, it is impossible to design such an ideal algorithm to achieve $\Gamma_k^*(T)$ in practice. We use $\Gamma_k^*(T)$ as a performance baseline to evaluate our OBSEA algorithm.

Theorem B.2. *The average social profits of OBSEA algorithm satisfies:*

$$\overline{\Gamma} = \frac{1}{KT}\sum_{t=1}^{KT}\Gamma(t) \ge \frac{1}{K}\sum_{k=1}^{K}\Gamma_k^*(T) - \frac{MT+Z}{V} \tag{B.10}$$

where

$$M = \frac{1}{2}|\mathcal{N}|^2(\mathrm{c}^{\max} + \mathrm{r}^{\max})^2 \tag{B.11}$$

$$Z = |\mathcal{N}|^2\mathrm{c}^{\max}(2\alpha\mathrm{H}_{\text{sink}}^{\max} + \eta^{\max}) \tag{B.12}$$

The proof of Theorem B.2 can be found below B.2.1. Inequality (B.10) demonstrates that the average social profits of our OBSEA algorithm will not be smaller than that of the ideal algorithm minus a term $(MT+Z)/V$ during a finite horizon with size t_{end}. In addition, the constraint (B.9) is more stringent than the constraint (6.10), since constraint (B.9) states the total amount of data injected to a node must be less than or equal to the total amount of data departures from this node, over each T-slot frame, rather than over the total K frames.

Therefore, $\frac{1}{K}\sum_{k=1}^{K}\Gamma_k^*(T)$ should be not greater than the optimal solution to problem (6.7). Therefore, the global social profit achieved by our OBSEA algorithm is not necessarily smaller than that of the ideal algorithm $\frac{1}{K}\sum_{k=1}^{K}\Gamma_k^*(T)$.

As M and Z are constant, parameter V can be set as large as desired to enforce $(MT+Z)/V$ to be arbitrarily small, but resulting a large risk of packet loss caused by data buffer overflow, according to Theorem B.1. In practice, V can be chosen as

$$V = \min_{x\in\mathcal{S}}(Q_x^{max} - \eta^{\max})/I_x'(0) \tag{B.13}$$

to maximise the worst-case global social profit bounds while guaranteeing no packet loss caused by buffer overflow.

B.2.1 Proof of Theorem B.2

Proof. Let $\boldsymbol{Q}(t)$ be the vector of all queues maintained at all nodes in $\mathcal{N}$. To simplify the proof, we assume the initial queue backlogs $\boldsymbol{Q}(0) = 0$. Define the Lyapunov function $L(\boldsymbol{Q}(t))$ as

$$L(\boldsymbol{Q}(t)) = \frac{1}{2}\sum_{x\in\mathcal{N}} Q_x^2(t) \tag{B.14}$$

Then we define the T-slot sample-path drift as

$$\triangle_T L(\boldsymbol{Q}(t)) = L(\boldsymbol{Q}(t+T)) - L(\boldsymbol{Q}(t)) \tag{B.15}$$

Denote $\Delta H_{x,y}^{sink} = H_x^{sink} - H_y^{sink}$ and $\Delta Q_{x,y}(t) = Q_x(t) - Q_y(t)$. We first consider 1-slot drift-plus-penalty for each slot $1 \leq t \leq t_{end}$, we have

$$
\begin{aligned}
&\triangle_1 L(\boldsymbol{Q}(t)) - V\Gamma(t) \\
&= L(\boldsymbol{Q}(t+1)) - L(\boldsymbol{Q}(t)) - V\Gamma(t) \\
&= \frac{1}{2}\sum_{x\in\mathcal{N}} Q_x^2(t+1) - \frac{1}{2}\sum_{x\in\mathcal{N}} Q_x^2(t) - V\Gamma(t) \\
&\leq_a M + \sum_{x\in\mathcal{N}} Q_x(t)(r_x(t)1_{x\in\mathcal{S}} + \sum_{y\in\mathcal{N}_x(t)} f_{y,x}(t) - \sum_{y\in\mathcal{N}_x(t)} f_{x,y}(t)) - V\Gamma(t) \\
&= M + \sum_{x\in S} r_s(t)Q_s(t) - V\Gamma(t) - \sum_{x\in\mathcal{N}}\sum_{y\in\mathcal{N}_x(t)} f_{x,y}(t)(Q_x(t) - Q_y(t)) \\
&= M \ + \sum_{x\in\mathcal{S}} r_x(t)Q_s(t) - VI_x(t) - \alpha \sum_{x\in\mathcal{N},y\in\mathcal{N}_x(t)} f_{x,y}(t)\Delta H_{x,y}^{sink} \\
&\quad - \sum_{x\in\mathcal{N},y\in\mathcal{N}_x(t)} 1_{\{Q_y(t)>Q_y^{max}\}} f_{x,y}(t)(\triangle Q_{x,y}(t) - \alpha\triangle H_{x,y}^{sink}(t)) \\
&\quad - \sum_{x\in\mathcal{N},y\in\mathcal{N}_x(t)} 1_{\{Q_y(t)\leq Q_y^{max}\}} f_{x,y}(t)w_{x,y}(t) \\
&\leq_b M + Z - V\sum_{x\in\mathcal{S}}(I_x(t) - r_x(t)Q_s(t)/V) - \sum_{x\in\mathcal{N},y\in\mathcal{N}_x(t)} 1_{\{Q_y(t)\leq Q_y^{max}\}} f_{x,y}(t)w_{x,y}(t)
\end{aligned} \tag{B.16}
$$

where inequality $\leq_a$ followed by the fact that $M \geq \sum_{x\in\mathcal{N}}(Q_x(t+1)-Q_x(t))^2$, $\forall t$; and inequality $\leq_b$ is because of the following fact

$$
\begin{aligned}
Z \quad &= |\mathcal{N}|^2\mathrm{c}^{\max}(2\alpha\mathrm{H}_{\mathrm{sink}}^{\max} + \eta^{\max}) \\
&= \alpha|\mathcal{N}|^2(\mathrm{c}^{\max}\mathrm{H}_{\mathrm{sink}}^{\max}) + |\mathcal{N}|^2\mathrm{c}^{\max}(\eta^{\max} + \alpha\mathrm{H}_{\mathrm{sink}}^{\max}) \\
&\geq -\alpha \sum_{x\in\mathcal{N},y\in\mathcal{N}_x(t)} f_{x,y}(t)\Delta H_{x,y}^{sink} - \sum_{x\in\mathcal{N},y\in\mathcal{N}_x(t)} 1_{\{Q_y(t)>Q_y^{max}\}} f_{x,y}(t)(\triangle Q_{x,y}(t) - \alpha\triangle H_{x,y}^{sink}(t))
\end{aligned}
$$

It is easy to see that our OBSEA algorithm greedily minimises the right-hand side of inequality (C.8) at every slot t, i.e. the rate controller minimises the third term of the right-hand side of inequality (C.8), and the routing component minimises the last term.

Then we define the T-slot sample-path drift-plus-penalty as

$$
\begin{aligned}
&\triangle_T L(\boldsymbol{Q}(t)) - V \sum_{t=kT-T+1}^{kT} \Gamma(t) \\
&= L(\boldsymbol{Q}(t+T)) - L(\boldsymbol{Q}(t)) - V \sum_{t=kT-T+1}^{kT} \Gamma(t) \\
&= \sum_{t=kT-T+1}^{kT} (L(\boldsymbol{Q}(t+1)) - L(\boldsymbol{Q}(t))) - V \sum_{t=kT-T+1}^{kT} \Gamma(t) \\
&= \sum_{t=kT-T+1}^{kT} (\triangle_1 L(\boldsymbol{Q}(t)) - V\Gamma(t)) \\
&\leq_a MT + ZT + \sqrt{2M}\frac{T(T-1)}{2} + \sum_{x\in\mathcal{N}} \sum_{t=kT-T+1}^{kT} 1_{\{Q_x(t)\leq Q_x^{max}\}} Q_x(t)(r_x(t)1_{x\in\mathcal{S}} \\
&\qquad\qquad + \sum_{y\in\mathcal{N}-\{x\}} f_{x,y}(t) - \sum_{y\in\mathcal{N}-\{x\}} f_{y,x}(t)) - V \sum_{t=kT-T+1}^{kT} \Gamma(t) \\
&\leq_b MT^2 + ZT + \sum_{x\in\mathcal{N}} \sum_{t=kT-T+1}^{kT} 1_{\{Q_x(t)\leq Q_x^{max}\}} Q_x(t)(r_x^*(t)1_{x\in\mathcal{S}} \\
&\qquad\qquad + \sum_{y\in\mathcal{N}-\{x\}} f_{x,y}^*(t) - \sum_{y\in\mathcal{N}-\{x\}} f_{y,x}^*(t)) - VT\Gamma^*(t) \\
&\leq_c MT^2 + ZT - VT\Gamma^j *_k (T)
\end{aligned} \tag{B.17}
$$

where the inequality $\leq_a$ is based on inequality (C.8), the sum of $\triangle_1 L(\boldsymbol{Q}(t)) - V\Gamma(t)$ over T slots, and the fact that the each queue backlog does not change by more than $(t-(kT-T+1))(\mathrm{r}^{\max}+\mathrm{c}^{\max})$ for any slot $kT-T+1 \leq t \leq kT$; the inequality $\leq_b$ is because $M \geq \sqrt{M}$ (M is a non-negative integer number), and our OBSEA algorithm minimises the right-hand side of $\leq_a$ over all possible rate control and routing decisions, including the decisions of ideal algorithm, $r_x^*(t), x \in \mathcal{S}$ and $f_{x,y}^*(t), x \in \mathcal{N}, y \in \mathcal{N}_x(t)$ that achieves $\Gamma_k^*(T)$; the inequality $\leq_c$ follows from the fact that the decisions $r_x^*(t)$ and $f_{x,y}^*(t)$ satisfy constraints (B.7) and (B.9).

Taking a telescopic sum of the inequality (C.9) over $k \in \{1, ..., K\}$ and dividing both side by VKT, we get

$$
L(\boldsymbol{Q}(KT+T)) - L(\boldsymbol{Q}(0)) - \frac{1}{KT}\sum_{t=1}^{KT}\Gamma(t) \leq \frac{MT+Z}{V} - \frac{1}{K}\sum_{k=1}^{K}\Gamma_k^*(T) \tag{B.18}
$$

Consider $L(\boldsymbol{Q}(1)) = 0$ and $L(\boldsymbol{Q}(KT+T)) \geq 0$, we have

$$\frac{1}{KT}\sum_{t=1}^{KT}\Gamma(t) \geq \frac{1}{K}\sum_{k=1}^{K}\Gamma_k^*(T) - \frac{MT+Z}{V}$$

□

Appendix C

Theorems for Chapter 6

The following appendix presents the results of the joint work between Shusen Yang and Usman Adeel.

C.1 Asymptotical Optimality of OptMPSS

To prove the optimality of OptMPSS, We divide the time horizon of the MPSS, $1 \leq t \leq \mathrm{t}_{\mathrm{end}}$, into K successive frames with size T slots (i.e. $\mathrm{t}_{\mathrm{end}} = KT$). We assume that there exists an *ideal* algorithm operating at the first slot of each frame $t = (k-1)T + 1, 1 \leq k \leq K$, which can obtain full information regarding the dynamics of the MPSS for future T slots (which is impossible in practice). Based on future knowledge, the ideal algorithm solves problem (6.7)-(6.10) over each frame $[(k-1)T + 1, kT]$, $1 \leq k \leq K$ rather than the whole horizon $[1, \mathrm{t}_{\mathrm{end}}]$. Note that when $T = \mathrm{t}_{\mathrm{end}}$, the ideal algorithm becomes the optimal solution of the original problem (6.7)-(6.10) . Let $\Phi^{ideal}(k, T)$ denote the optimal global gross profit computed by the ideal algorithm over each frame $1 \leq k \leq K$.

Theorem C.1. *The time-average global gross profit computed by OptMPSS satisfies:*

$$\Phi^{OptMPSS} \geq \frac{1}{K}\sum_{k=1}^{K}\Phi^{ideal}(k,T) - \frac{MT}{V} \tag{C.1}$$

where $M = |\mathcal{N}|(\mathrm{r}_{\max} + \mu_{\max})^2/2$ *is a constant value.*

Proof of Theorem C.1. Theorem C.1 can be proved by using sample-path based Lyapunov optimization theory. We first present the formalized optimization problem that the ideal algorithm aims to optimize at the beginning of each frame $k, 1 \leq k \leq K$.

$$\max_{\boldsymbol{x}_i(t), i\in\mathcal{N}} \quad \sum_{i\in\mathcal{N}}(\alpha v(\overline{r}_i(k,T) - \overline{cost}_i(k,T))) \tag{C.2}$$

s.t.

$$r_i(t) < \mathrm{r}_{\max},\ i \in \mathcal{N}\ kT - T + 1 \leq t \leq kT \tag{C.3}$$

$$f_{i,j}(t) \leq \mu_{i,j}(t),\ i \in \mathcal{N}, j \in \mathcal{N}_i(t),\ kT - T + 1 \leq t \leq kT \tag{C.4}$$

$$\overline{r}_i(k,T) + \overline{f}_i^{in}(k,T) - \overline{f}_i^{out} = 0,\ kT - T + 1 \leq t \leq kT \tag{C.5}$$

where [35]

$$\overline{r}_i(k,T) = \sum_{k=kT-T+1}^{kT} r_i(t) \tag{C.6}$$

which represents the time-average sensing rate during the kth T-slot frame. Other time-average variables during the kth frame are defined in a similar way.

Now we prove Theorem C.1. Let $\boldsymbol{Q}(t)$ be the vector of all queues maintained at all phones in $\mathcal{N}$ and the server S. To simplify the proof, we assume the initial queue backlogs $\boldsymbol{Q}(0) = 0$. Define the Lyapunov function $L(\boldsymbol{Q}(t))$ as

$$L(t) = \frac{1}{2}\sum_{i\in\mathcal{N}\cup\{S\}} Q_i^2(t) = \frac{1}{2}\sum_{i\in\mathcal{N}} Q_i^2(t) \tag{C.7}$$

Let $\varphi_i(t) = \alpha v_i(r_i(t)) - cost_i(t)$ and $\Phi(t) = \sum_{i\in\mathcal{N}}\varphi_i(t)$. We first consider 1-slot Lyapunov drift plus penalty for each slot $1 \leq t \leq \mathrm{t}_{\mathrm{end}}$:

$$\begin{aligned}
&\triangle_1 L(t) - V\Phi(t) \\
&= L(\boldsymbol{Q}(t+1)) - L(\boldsymbol{Q}(t)) - V\Phi(t) \\
&= \frac{1}{2}\sum_{i\in\mathcal{N}} Q_x^2(t+1) - \frac{1}{2}\sum_{x\in\mathcal{N}} Q_x^2(t) - V\Gamma(t) \\
&\leq_a M + \sum_{i\in\mathcal{N}} Q_i(t)(r_i(t) + f_i^{in}(t) - f_i^{out}(t)) - V\sum_{i\in\mathcal{N}}\varphi_i(t) \\
&= M - \underbrace{\sum_{i\in\mathcal{N}}(V\alpha v_i(r_i(t)) - Q_i(t)r_i(t) - Vc_i^s(t)r_i(t))}_{\text{sensing rate control}} \\
&\qquad - \underbrace{\sum_{(i,j)\in\mathcal{L}}(Q_i(t) - Q_j(t)) - Vc_{i,j}^t f_{i,j}(t)}_{\text{routing and data forwarding}}
\end{aligned} \tag{C.8}$$

where inequality $\leq_a$ followed by the fact that

$$\begin{aligned}
M &= \frac{1}{2}|\mathcal{N}|(\mathrm{r}_{\max} + \mu_{\max})^2 \\
&\geq \frac{1}{2}\sum_{i\in\mathcal{N}}(Q_i(t+1) - Q_i(t))^2, \ \forall 1 \leq t \leq \mathrm{t}_{\mathrm{end}}
\end{aligned}$$

It can be verified that the OptMPSS algorithm minimizes the right-hand side of inequality (C.8) by making sensing rate control and routing decisions at each slot $1 \leq t \leq \mathrm{t}_{\mathrm{end}}$. Then we consider the T-slot sample-path drift-plus-penalty for a frame $k,\ 1 \leq k \leq K$:

$$
\begin{aligned}
&\triangle_T L(t) - V \sum_{t=kT-T+1}^{kT} \Phi(t) \\
&= L(kT) - L(kT-T+1) - V \sum_{t=kT-T+1}^{kT} \Phi(t) \\
&= \sum_{t=kT-T+1}^{kT} (\triangle_1 L(t) - \Phi(t)) \\
&\leq_a MT + \sqrt{2M}\frac{T(T-1)}{2} \\
&\quad + \sum_{t=kT-T+1}^{kT} (\sum_{i\in\mathcal{N}} Q_i(t)(r_i(t) + f_i^{in}(t) - f_i^{out}(t)) \\
&\quad - V \sum_{i\in\mathcal{N}} \varphi_i(t)) \\
&\leq_b MT^2 - VT\Phi^{ideal}(k,T) + \sum_{t=kT-T+1}^{kT} \sum_{i\in\mathcal{N}} Q_i(t)(r_i(t) \\
&\quad + \sum_{j\in\mathcal{N}_i(t)} (f_{j,i}^{ideal}(t) - f_{i,j}^{ideal}(t)))
\end{aligned}
\tag{C.9}
$$

where the inequality $\leq_a$ is based on inequality (C.8), the sum of $\triangle_1 L(t) - V\Phi(t)$ over T slots of the kth frame, and the fact that the each queue backlog does not change by more than $(t-(kT-T+1))(\mathrm{r}^{\max}+\mu^{\max})$ for any slot $kT-T+1 \leq t \leq kT$; the inequality $\leq_b$ follows from $M \geq \sqrt{M}, \forall M \geq 1$, and the fact that our OptMPSS algorithm minimizes the right-hand side of inequality $\leq_a$ over all possible sensing rate control and routing decisions, including the decisions of the ideal algorithm: $r_i^{ideal}(t)$ and $f_{i,j}^{ideal}(t)$, $\forall i \in \mathcal{N}\ ,j \in \mathcal{N}_i(t)$, which achieves the optimal gross profits of the ideal algorithm $\Phi^{ideal}(k,T)$. Consider (C.9) and the fact that $r_i^{ideal}(t)$ and $f_{i,j}^{ideal}(t)$ satisfy constraint (C.5), we have

$$
\begin{aligned}
&\triangle_T L(t) - V \sum_{t=kT-T+1}^{kT} \Phi^{OptMPSS}(t) \\
&\leq MT^2 - VT\Phi^{ideal}(k,T)
\end{aligned}
\tag{C.10}
$$

Taking a telescopic sum of the inequality (C.10) over $k \in \{1, ..., K\}$ and dividing both side by VKT, we get

$$\begin{aligned} L(KT+T) - L(0) - \frac{1}{KT}\sum_{t=1}^{KT}\Phi^{OptMPSS}(t) \\ \leq \frac{MT}{V} - \frac{1}{K}\sum_{k=1}^{K}\Phi^{ideal}(k,T) \end{aligned}$$

Consider $L(1) = 0$ and $L(KT+T) \geq 0$, we have

$$\begin{aligned} &\Phi^{OptMPSS} \\ &= \frac{1}{KT}\sum_{t=1}^{KT}\Phi^{OptMPSS}(t) \\ &\geq \frac{1}{K}\sum_{k=1}^{K}\Phi(k,T) - \frac{MT}{V} \end{aligned}$$

This completes the proof the Theorem C.1. □ □

Inequality (C.1) shows that parameter V can be set as large as desired to force MT/V to be arbitrarily small. Specifically, Theorem C.1 also demonstrates that when $T = \mathrm{t_{end}}$, the optimal average global gross profit can be asymptotically achieved by OptMPSS, as $V \to \infty$.

C.2 Asymptotic Incentive Compatibility of BMT

To prove that BMT achieves asymptotic incentive compatibility, We first introduce a new definition and a supportuing lemma.

Definition 5 [Asymptotically Efficient Social Decision] *For a distributed mechanism $dM = (h, \Pi, A)$, a social decision $h_{\boldsymbol{x}}(\boldsymbol{\theta})$ made by the suggested algorithm A is said to be asymptotically efficient if*

$$\sum_{i\in\mathcal{N}}\varphi_i(h_{\mathbf{x}} \circ A(\boldsymbol{\theta}), \theta_i) \geq \sum_{i\in\mathcal{N}}\varphi_i(h_{\mathbf{x}} \circ A'(\boldsymbol{\theta}), \theta_i) - \varepsilon(V)$$

for all $\boldsymbol{\theta} \in \boldsymbol{\Theta}$ and for all $A' \in \Pi$, where $\varepsilon(V) > 0$ and $\varepsilon(V) \to 0$ as $V \to \infty$.

Lemma 1. *The social decision made by BMT algorithm, $\boldsymbol{x}^{bmt}$ is asymptotically efficient.*

Proof. **Proof.** Since the distributed social decision (e.g. sensing rate control and routing decisions) made by BMT is the same as that of OptMPSS, this Lemma obviously holds when $\varepsilon(V) = MT/V$ and frame size $T = \mathrm{t}_{\text{end}}$, according to Theorem 1. □

□

Theorem C.2. *BMT achieves asymptotic incentive compatibility.*

Proof of Theorem C.2. We prove Theorem C.2 by contradiction. Consider a distributed mechanism $dM = (h, \Pi, A^{bmt})$, where $A^{bmt} = (A_1^{bmt}, ..., A_{|\mathcal{N}|}^{bmt})$ is the distributed strategy of each phone allocated by BMT algorithm. Suppose that BMT is not asymptotically incentive compatible, i.e. $\exists i \in \mathcal{N}, A_i' \neq A_i^{bmt}$ such that

$$\begin{aligned}
& u_i(h(A_i^{bmt}(\theta_i), A_{-i}^{bmt}(\theta_{-i}), \theta_i) + \varepsilon(V) \\
&=_{(a)} \varphi_i(h_{\mathbf{x}}(A_i^{bmt}(\theta_i), A_{-i}^{bmt}(\theta_{-i})), \theta_i) \\
&\qquad\qquad + h_{\lambda_i}(A_i^{bmt}(\theta_i), A_{-i}^{bmt}(\theta_{-i})) + \varepsilon(V) \\
&=_{(b)} \sum_{i \in \mathcal{N}} \varphi_i(h_{\mathbf{x}}(A_i^{bmt}(\theta_i), A_{-i}^{bmt}(\theta_{-i})), \theta_i) \\
&\qquad\qquad\qquad - \max \sum_{j \neq i} \varphi_j + \varepsilon(V) \\
&< \sum_{i \in \mathcal{N}} \varphi_i(h_{\mathbf{x}}(A_i'(\theta_i), A_{-i}^{bmt}(\theta_{-i}), \theta_i) - \max \sum_{j \neq i} \varphi_j \\
&= u_i(h_{\mathbf{x}}(A_i'(\theta_i), A_{-i}^{bmt}(\theta_{-i}), \theta_i)
\end{aligned} \tag{C.11}$$

where equalities (a) and (b) follow the definitions of net profit and VCG tax respectively. It can be seen that inequality (C.11) implies that

$$\sum_{i \in \mathcal{N}} \varphi_i(h_{\mathbf{x}} \circ A^{bmt}(\boldsymbol{\theta}), \theta_i) < \sum_{i \in \mathcal{N}} \varphi_i(h_{\mathbf{x}} \circ A'(\boldsymbol{\theta}), \theta_i) - \varepsilon(V)$$

where $A' = (A_1^{bmt}, ..., A_i', ..., A_{|\mathcal{N}|}^{bmt})$. This contradicts the asymptotically social efficiency of BMT, i.e. Lemma 1. $\square$ $\square$

Bibliography

[1] http://www.tinyos.net/.

[2] http://www.contiki-os.org/.

[3] http://castalia.npc.nicta.com.au/.

[4] Greenorbs: http://www.greenorbs.org/.

[5] http://crawdad.cs.dartmouth.edu/ crawdad/cambridge/haggle/.

[6] http://www.qualcomm.com/research/projects/lte-direct.

[7] http://www.sensorly.com.

[8] micaz mote,. http://www.openautomation.net/uploadsproductos/micaz_datasheet.pdf.

[9] *RSSI-Based Forwarding for Multihop Wireless Sensor Networks*, EUNICE '09, Berlin, Heidelberg, 2009. Springer-Verlag.

[10] Billmonitor "the national billmonitor mobile report". http://www.billmonitor.com/national-billmonitor-mobile-report, 2011.

[11] Rpl: Ipv6 routing protocol for low-power and lossy networks. 2012.

[12] Muhammad Abdulla and Robert Simon. Characteristics of common mobility models for opportunistic networks. In *Proceedings of the 2nd ACM workshop on Performance monitoring and measurement of heterogeneous wireless and wired networks*, PM2HW2N '07, pages 105–109, New York, NY, USA, 2007. ACM.

[13] U. Akyol, M. Andrews, P. Gupta, J. Hobby, I. Saniee, and A. Stolyar. Joint scheduling and congestion control in mobile ad-hoc networks. In *Proc. IEEE Infocom*, pages 619–627, 2008.

[14] Jude Allred, Ahmad Bilal Hasan, Saroch Panichsakul, William Pisano, Peter Gray, Jyh Huang, Richard Han, Dale Lawrence, and Kamran Mohseni. Sensorflock: an airborne wireless sensor network of micro-air vehicles. In *Proceedings of the 5th international conference on Embedded networked sensor systems*, pages 117–129. ACM, 2007.

[15] Majed Alresaini, Maheswaran Sathiamoorthy, Bhaskar Krishnamachari, and Michael J Neely. Backpressure with adaptive redundancy (bwar). In *Proc. IEEE Infocom*, pages 2300–2308, 2012.

[16] Waleed Alsalih, Hossam Hassanein, and Selim Akl. Placement of multiple mobile data collectors in wireless sensor networks. *Ad Hoc Networks*, 8(4):378–390, 2010.

[17] Habib M Ammari and Sajal K Das. Mission-oriented k-coverage in mobile wireless sensor networks. In *Distributed Computing and Networking*, pages 92–103. Springer, 2010.

[18] G. Anastasi, M. Conti, and M. Di Francesco. Reliable and energy-efficient data collection in sparse sensor networks with mobile elements. *Performance Evaluation*, 66(12):791–810, 2009.

[19] G. Anastasi, M. Conti, E. Gregori, C. Spagoni, and G. Valente. Motes sensor networks in dynamic scenarios: an experimental study for pervasive applications in urban environments. *Journal of Ubiquitous Computing and Intelligence*, 1(1):9–16, 2007.

[20] G. Anastasi, M. Conti, E. Monaldi, and A. Passarella. An adaptive data-transfer protocol for sensor networks with Data Mules. In *World of Wireless, Mobile and Multimedia Networks, 2007. WoWMoM 2007. IEEE International Symposium on a*, pages 1–8. IEEE, 2007.

[21] Giuseppe Anastasi, Eleonora Borgia, Marco Conti, and Enrico Gregori. A Hybrid Adaptive Protocol for Reliable Data Delivery in WSNs with Multiple Mobile Sinks. *The Computer Journal*, 54(2):213–229, 2011.

[22] Luigi Atzori, Antonio Iera, and Giacomo Morabito. The internet of things: A survey. *Computer Networks*, 54(15):2787–2805, 2010.

[23] Mehedi Bakht, Matt Trower, and Robin Hilary Kravets. Searchlight: won't you be my neighbor? In *Proceedings of the 18th annual international conference on Mobile computing and networking*, pages 185–196, 2012.

[24] A.L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286(5439):509–512, 1999.

[25] Stefano Basagni, Ladislau Bölöni, Petrika Gjanci, Chiara Petrioli, Cynthia A Phillips, and Damla Turgut. Maximizing the value of sensed information in underwater wireless sensor networks via an autonomous underwater vehicle.

[26] Stefano Basagni, Alessio Carosi, Chiara Petrioli, and Cynthia A Phillips. Coordinated and controlled mobility of multiple sinks for maximizing the lifetime of wireless sensor networks. *Wireless Networks*, 17(3):759–778, 2011.

[27] Greg Bigwood and Tristan Henderson. Ironman: Using social networks to add incentives and reputation to opportunistic networks. In *Privacy, security, risk and trust (passat), 2011 ieee third international conference on and 2011 ieee third international conference on social computing (socialcom)*, pages 65–72. IEEE, 2011.

[28] L. Boloni and D. Turgut. Should I send now or send later? A decision-theoretic approach to transmission scheduling in sensor networks with mobile sinks. *Wireless Communications and Mobile Computing*, 8(3):385–403, 2008.

[29] A. Brzezinski, G. Zussman, and E. Modiano. Enabling distributed throughput maximization in wireless mesh networks: a partitioning approach. In *Proc. ACM Mobicom*, pages 26–37, 2006.

[30] M. Buettner, G.V. Yee, E. Anderson, and R. Han. X-MAC: a short preamble MAC protocol for duty-cycled wireless sensor networks. In *Proceedings of the 4th international conference on Embedded networked sensor systems*, pages 307–320. ACM, 2006.

[31] J. Burrell, T. Brooke, and R. Beckwith. Vineyard computing: Sensor networks in agricultural production. *IEEE Pervasive Comput.*, 3(1):38–45, 2004.

[32] Levente Buttyán, László Dóra, Márk Félegyházi, and István Vajda. Barter trade improves message delivery in opportunistic networks. *Ad Hoc Networks*, 8(1):1–14, 2010.

[33] Levente Buttyán and Jean-Pierre Hubaux. Enforcing service availability in mobile ad-hoc wans. In *Proceedings of the 1st ACM international symposium on Mobile ad hoc networking & computing*, pages 87–96. IEEE Press, 2000.

[34] Han Cai and Do Young Eun. Crossing over the bounded domain: from exponential to power-law inter-meeting time in manet. In *Proceedings of the 13th annual ACM international conference on Mobile computing and networking*, pages 159–170. ACM, 2007.

[35] Andrew Campbell et al. The rise of people-centric sensing. *IEEE Internet Comput.*, 12(4):12–21, 2008.

[36] Daniel Camps-Mur, Andres Garcia-Saavedra, and Pablo Serrano. Device to device communications with wifi direct: overview and experimentation. *IEEE Wireless Commun.*, 20(3):96–104, 2013.

[37] H. Cao, K.W. Parker, and A. Arora. O-MAC: a receiver centric power management protocol. In *Network Protocols, 2006. ICNP'06. Proceedings of the 2006 14th IEEE International Conference on*, pages 311–320. IEEE, 2007.

[38] A. Chaintreau, P. Hui, J. Crowcroft, C. Diot, R. Gass, and J. Scott. Impact of human mobility on opportunistic forwarding algorithms. *IEEE Trans. Mobi. Comput.*, 6(6):606–620, 2007.

[39] Augustin Chaintreau, Pan Hui, Jon Crowcroft, Christophe Diot, Richard Gass, and James Scott. Impact of human mobility on the design of opportunistic forwarding algorithms. In *Proc. IEEE Infocom*, 2006.

[40] Tony F Chan, Gene H Golub, and Randall J LeVeque. Algorithms for computing the sample variance: Analysis and recommendations. *The American Statistician*, 37(3):242–247, 1983.

[41] C.K. Chau and P. Basu. Analysis of latency of stateless opportunistic forwarding in intermittently connected networks. *IEEE/ACM Trans. Netw.*, 19(4):1111–1124, 2011.

[42] L. Chen, S.H. Low, M. Chiang, and J.C. Doyle. Cross-layer congestion control, routing and scheduling design in ad hoc wireless networks. In *Proc. IEEE Infocom*, pages 1–13, 2006.

[43] M. Chiang, S.H. Low, A.R. Calderbank, and J.C. Doyle. Layering as optimization decomposition: A mathematical theory of network architectures. *Proc. IEEE*, 95(1):255–312, 2007.

[44] T Cisco. Cisco visual networking index: Global mobile data traffic forecast update, 2012–2017. *Cisco Public Information*, 2013.

[45] M. Conti, S. Giordano, M. May, and A. Passarella. From opportunistic networks to opportunistic computing. *IEEE Commun. Mag.*, 48(9):126–139, 2010.

[46] M. Conti and M. Kumar. Opportunities in opportunistic computing. *Computer*, 43(1):42–50, 2010.

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

[48] Elizabeth M Daly and Mads Haahr. Social network analysis for routing in disconnected delay-tolerant manets. In *Proceedings of the 8th ACM international symposium on Mobile ad hoc networking and computing*, pages 32–40. ACM, 2007.

[49] Douglas SJ De Couto, Daniel Aguayo, John Bicket, and Robert Morris. A high-throughput path metric for multi-hop wireless routing. In *Proc. Mobicom*, 2003.

[50] Michael Demmer and Kevin Fall. Dtlsr: delay tolerant routing for developing regions. In *Proceedings of the 2007 workshop on Networked systems for developing regions*, page 5. ACM, 2007.

[51] M. Di Francesco, S.K. Das, and G. Anastasi. Data collection in wireless sensor networks with mobile elements: a survey. *ACM Trans. Sensor Networks*, 8(1):7, 2011.

[52] Suhas N Diggavi, Matthias Grossglauser, and David N C Tse. Even one-dimensional mobility increases ad hoc wireless capacity. In *Information Theory, 2002. Proceedings. 2002 IEEE International Symposium on*, page 352. IEEE, 2002.

[53] Wei Dong, Swati Rallapalli, Rittwik Jana, Lili Qiu, KK Ramakrishnan, Leo Razoumov, Yin Zhang, and Tae Won Cho. ideal: Incentivized dynamic cellular offloading via auctions. In *Proc. IEEE INFOCOM*, pages 755–763, 2013.

[54] Klaus Doppler, Mika Rinne, Carl Wijting, Cassio B Ribeiro, and Klaus Hugl. Device-to-device communication as an underlay to lte-advanced networks. *IEEE Communications Magazine*, 47(12):42–49, 2009.

[55] M. Dunbabin, P. Corke, I. Vasilescu, and D. Rus. Data muling over underwater wireless sensor networks using an autonomous underwater vehicle. In *Proc. IEEE ICRA*, pages 2091–2098, 2006.

[56] A. Dvir and A.V. Vasilakos. Backpressure-based routing protocol for dtns. In *ACM SIGCOMM Computer Communication Review*, volume 40, pages 405–406, 2010.

[57] A. Dvir and A.V. Vasilakos. Backpressure-based routing protocol for dtns. In *ACM SIGCOMM Computer Communication Review*, volume 40, pages 405–406, 2010.

[58] N. Eagle and A. Pentland. Reality mining: sensing complex social systems. *Personal and Ubiquitous Computing*, 10(4):255–268, 2006.

[59] Albert Einstein. *Investigations on the Theory of the Brownian Movement.* Courier Dover Publications, 1956.

[60] K. Fall. A delay-tolerant network architecture for challenged internets. In *Proc. ACM Sigcomm*, pages 27–34, 2003.

[61] Jialu Fan, Yuan Du, Wei Gao, Jiming Chen, and Youxian Sun. Geography-aware active data dissemination in mobile social networks. In *Mobile Adhoc and Sensor Systems (MASS), 2010 IEEE 7th International Conference on*, pages 109–118. IEEE, 2010.

[62] L.C. Freeman. Centrality in social networks conceptual clarification. *Social networks*, 1(3):215–239, 1979.

[63] Raghu K Ganti, Fan Ye, and Hui Lei. Mobile crowdsensing: current state and future challenges. *IEEE Commun. Mag.*, 49(11):32–39, 2011.

[64] Raghu K Ganti, Fan Ye, and Hui Lei. Mobile crowdsensing: current state and future challenges. *IEEE Commun. Mag.*, 49(11):32–39, 2011.

[65] S. Gao, H.K. Zhang, and H.S. Xu. Efficient data gathering approach in sensor networks with path-fixed sinks. *Ruan Jian Xue Bao(Journal of Software)*, 21(1):147–162, 2010.

[66] W. Gao, G. Cao, T. La Porta, and J. Han. On exploiting transient social contact patterns for data forwarding in delay tolerant networks. *IEEE Trans. Mobi. Comput.*, (99):1–1, 2011.

[67] W. Gao, Q. Li, B. Zhao, and G. Cao. Social-aware multicast in disruption-tolerant networks. *IEEE/ACM Trans. Netw.*, (99):1–1.

[68] Wei Gao and Guohong Cao. User-centric data dissemination in disruption tolerant networks. In *Proc. IEEE Infocom*, pages 3119–3127, 2011.

[69] L. Georgiadis, M. Neely, M.J. Neely, and L. Tassiulas. *Resource allocation and cross-layer control in wireless networks*. Now Pub, 2006.

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

[71] D Gilstrap. Traffic and market report: On the pulse of the networked society. *Ericsson Traffic and Market Report*, 2012.

[72] O. Gnawali, R. Fonseca, K. Jamieson, D. Moss, and P. Levis. Collection tree protocol. In *Proc. ACM SenSys*, 2009.

[73] O. Gnawali, R. Fonseca, K. Jamieson, D. Moss, and P. Levis. Collection tree protocol. In *Proc. ACM Sensys*, pages 1–14, 2009.

[74] M.C. Gonzalez, C.A. Hidalgo, and A.L. Barabási. Understanding individual human mobility patterns. *Nature*, 453(7196):779–782, 2008.

[75] Kannan Govindan and Prasant Mohapatra. Trust computations and trust dynamics in mobile adhoc networks: a survey. *IEEE Communications Surveys & Tutorials*, 14(2):279–298, 2012.

[76] Matthias Grossglauser and David Tse. Mobility increases the capacity of ad-hoc wireless networks. In *INFOCOM 2001. Twentieth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, volume 3, pages 1360–1369. IEEE, 2001.

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

[78] Vishal Gupta and Mukesh Kumar Rohil. Enhancing wifi with ieee802. 11u for mobile data offloading. *International Journal of Mobile Network Communications & Telematics (IJMNCT)*, 2(4):19–29, 2012.

[79] Bo Han, Pan Hui, VS Anil Kumar, Madhav V Marathe, Jianhua Shao, and Aravind Srinivasan. Mobile data offloading through opportunistic communications and social participation. *IEEE Trans. Mobi. Comput.*, 11(5):821–834, 2012.

[80] W. Hsu, T. Spyropoulos, K. Psounis, and A. Helmy. Modeling time-variant user mobility in wireless mobile networks. In *Proc. IEEE Infocom*, pages 758–766, 2007.

[81] J.H. Huang, S. Amjad, and S. Mishra. Cenwits: a sensor-based loosely coupled search and rescue system using witnesses. In *Proc. ACM SenSys*, pages 180–191, 2005.

[82] L. Huang, S. Moeller, M.J. Neely, and B. Krishnamachari. Lifo-backpressure achieves near optimal utility-delay tradeoff. In *Proc. IEEE WiOpt*, pages 70–77, 2011.

[83] L. Huang and M.J. Neely. Delay reduction via lagrange multipliers in stochastic network optimization. *IEEE Trans. Autom. Control*, 56(4):842–857, 2011.

[84] P. Hui, J. Crowcroft, and E. Yoneki. Bubble rap: social-based forwarding in delay tolerant networks. In *Proc. ACM Mobihoc*, pages 241–250, 2008.

[85] P. Hui, J. Crowcroft, and E. Yoneki. Bubble rap: Social-based forwarding in delay-tolerant networks. *IEEE Trans. Mobi. Comput.*, 10(11):1576–1589, 2011.

[86] P. Hui, E. Yoneki, S.Y. Chan, and J. Crowcroft. Distributed community detection in delay tolerant networks. In *Proc. MobiArch*, 2007.

[87] Pan Hui, Augustin Chaintreau, James Scott, Richard Gass, Jon Crowcroft, and Christophe Diot. Pocket switched networks and human mobility in conference environments. In *Proceedings of the 2005 ACM SIGCOMM workshop on Delay-tolerant networking*, pages 244–251. ACM, 2005.

[88] Pan Hui, Kuang Xu, Victor OK Li, Jon Crowcroft, Vito Latora, and Pietro Lio. Selfishness, altruism and message spreading in mobile social networks. In *INFOCOM Workshops 2009, IEEE*, pages 1–6. IEEE, 2009.

[89] Bret Hull, Vladimir Bychkovsky, Yang Zhang, Kevin Chen, Michel Goraczko, Allen Miu, Eugene Shih, Hari Balakrishnan, and Samuel Madden. Cartel: a distributed mobile sensor computing system. In *Proc. ACM SenSys*, pages 125–138, 2006.

[90] C. Intanagonwiwat, R. Govindan, and D. Estrin. Directed diffusion: A scalable and robust communication paradigm for sensor networks. In *Proceedings of the 6th annual international conference on Mobile computing and networking*, pages 56–67. ACM, 2000.

[91] Matthew O Jackson. Mechanism theory. *The Encyclopedia of Life Support Systems*, 2000.

[92] Matthew O Jackson and Brian W Rogers. Meeting strangers and friends of friends: How random are social networks? *The American economic review*, pages 890–915, 2007.

[93] S. Jain, K. Fall, and R. Patra. Routing in a delay tolerant network. In *Proceedings of the 2004 conference on Applications, technologies, architectures, and protocols for computer communications*, pages 145–158. ACM, 2004.

[94] S. Jain, R.C. Shah, W. Brunette, G. Borriello, and S. Roy. Exploiting mobility for energy efficient data collection in wireless sensor networks. *Mobile Networks and Applications*, 11(3):327–339, 2006.

[95] D. Jea, A. Somasundara, and M. Srivastava. Multiple controlled mobile elements (data mules) for data collection in sensor networks. *Distributed Computing in Sensor Systems*, pages 244–257, 2005.

[96] David B Johnson and David A Maltz. Dynamic source routing in ad hoc wireless networks. In *Mobile computing*, pages 153–181. Springer, 1996.

[97] C. Joo, X. Lin, and N.B. Shroff. Understanding the capacity region of the greedy maximal scheduling algorithm in multihop wireless networks. *IEEE/ACM Trans. Network.*, 17(4):1132–1145, 2009.

[98] Philo Juang, Hidekazu Oki, Yong Wang, Margaret Martonosi, Li Shiuan Peh, and Daniel Rubenstein. Energy-efficient computing for wildlife tracking: Design tradeoffs and early experiences with zebranet. In *ACM Sigplan Notices*, volume 37, pages 96–107. ACM, 2002.

[99] H. Jun, M.H. Ammar, and E.W. Zegura. Power management in delay tolerant networks: A framework and knowledge-based mechanisms. In *IEEE Communications Society Conference on Sensor and Ad Hoc Communications and Networks (SECON)*. Citeseer, 2005.

[100] J.M. Kahn, R.H. Katz, and K.S.J. Pister. Next century challenges: mobile networking for Smart Dust. In *Proceedings of the 5th annual ACM/IEEE international conference on Mobile computing and networking*, pages 271–278. ACM, 1999.

[101] A. Kansal, A.A. Somasundara, D.D. Jea, M.B. Srivastava, and D. Estrin. Intelligent fluid infrastructure for embedded networks. In *Proceedings of the 2nd international conference on Mobile systems, applications, and services*, pages 111–124. ACM, 2004.

[102] T. Karagiannis, J.Y. Le Boudec, and M. Vojnovic. Power law and exponential decay of intercontact times between mobile devices. *IEEE Trans. Mobi.Comput.*, 9(10):1377–1390, 2010.

[103] F.P. Kelly, A.K. Maulloo, and D.K.H. Tan. Rate control for communication networks: shadow prices, proportional fairness and stability. *Journ. of the Operational Res. society*, 49(3):237–252, 1998.

[104] Wazir Zada Khan, Yang Xiang, Mohammed Y Aalsalem, and Quratulain Arshad. Mobile phone sensing systems: a survey. *IEEE Communications Surveys & Tutorials*, 15(1):402–427, 2013.

[105] H.S. Kim, T.F. Abdelzaher, and W.H. Kwon. Minimum-energy asynchronous dissemination to mobile sinks in wireless sensor networks. In *Proc. ACM SenSys*, pages 193–204, 2003.

[106] Minkyong Kim, David Kotz, and Songkuk Kim. Extracting a mobility model from real user traces. In *INFOCOM*, volume 6, pages 1–13, 2006.

[107] Iordanis Koutsopoulos. Optimal incentive-driven design of participatory sensing systems. In *Proc. IEEE INFOCOM*, pages 1402–1410, 2013.

[108] B. Kusy, H.J. Lee, M. Wicke, N. Milosavljevic, and L. Guibas. Predictive qos routing to mobile sinks in wireless sensor networks. In *Proc. IEEE/ACM IPSN*, pages 109–120, 2009.

[109] Nicholas D Lane, Emiliano Miluzzo, Hong Lu, Daniel Peebles, Tanzeem Choudhury, and Andrew T Campbell. A survey of mobile phone sensing. *IEEE Communicati.s Mag.*, 48(9):140–150, 2010.

[110] Nicholas D Lane, Emiliano Miluzzo, Hong Lu, Daniel Peebles, Tanzeem Choudhury, and Andrew T Campbell. A survey of mobile phone sensing. *IEEE Communications Magazine*, 48(9):140–150, 2010.

[111] L.B. Le, E. Modiano, and N.B. Shroff. Optimal control of wireless networks with finite buffers. In *Proc. IEEE Infocom*, pages 1–9, 2010.

[112] J.Y. Le Boudec and M. Vojnovic. Perfect simulation and stationarity of a class of mobility models. In *Proc. IEEE Infocom*, volume 4, pages 2743–2754, 2005.

[113] Jason LeBrun, Chen-Nee Chuah, Dipak Ghosal, and Michael Zhang. Knowledge-based opportunistic forwarding in vehicular wireless ad hoc networks. In *Vehicular technology conference, 2005. VTC 2005-Spring. 2005 IEEE 61st*, volume 4, pages 2289–2293. IEEE, 2005.

[114] H. Lee and A. Keshavarzian. Towards energy-optimal and reliable data collection via collision-free scheduling in wireless sensor networks. In *INFOCOM 2008. The 27th Conference on Computer Communications. IEEE*, pages 2029–2037. IEEE, 2008.

[115] H.J. Lee, M. Wicke, B. Kusy, O. Gnawali, and L. Guibas. Data stashing: energy-efficient information delivery to mobile sinks through trajectory prediction. In *Proc. ACM/IEEE IPSN*, pages 291–302, 2010.

[116] Juong-Sik Lee and Baik Hoh. Sell your experiences: a market mechanism based incentive for participatory sensing. In *Proc. IEEE PerCom 2010*, pages 60–68, 2010.

[117] J.W. Lee, B. Kusy, T. Azim, B. Shihada, and P. Levis. Whirlpool routing for mobility. In *Proc. ACM MobiHoc*, volume 10, pages 131–140, 2010.

[118] Kyunghan Lee, Joohyun Lee, Yung Yi, Injong Rhee, and Song Chong. Mobile data offloading: how much can wifi deliver? In *In Proc. CoNEXT*, page 26, 2010.

[119] Lei Lei, Zhangdui Zhong, Chuang Lin, and Xuemin Shen. Operator controlled device-to-device communications in lte-advanced networks. *IEEE Wireless Commun.*, 19(3):96–104, 2012.

[120] Na Li and Sajal K Das. Radon: reputation-assisted data forwarding in opportunistic networks. In *Proceedings of the Second International Workshop on Mobile Opportunistic Networking*, pages 8–14. ACM, 2010.

[121] Q. Li, S. Zhu, and G. Cao. Routing in socially selfish delay tolerant networks. In *Proc. IEEE Infocom*, pages 1–9, 2010.

[122] Q. Li, S. Zhu, and G. Cao. Routing in socially selfish delay tolerant networks. In *Proc. IEEE Infocom*, pages 1–9, 2010.

[123] Z. Li, M. Li, J. Wang, and Z. Cao. Ubiquitous data collection for mobile users in wireless sensor networks. In *Proc. IEEE Infocom*, pages 2246–2254, 2011.

[124] H. Lin, M. Lu, N. Milosavljevic, J. Gao, and L.J. Guibas. Composable information gradients in wireless sensor networks. In *Proc. IEEE/ACM IPSN*, pages 121–132, 2008.

[125] Anders Lindgren, Avri Doria, and Olov Schelén. Probabilistic routing in intermittently connected networks. *ACM SIGMOBILE mobile computing and communications review*, 7(3):19–20, 2003.

[126] Chi Harold Liu, Pan Hui, Joel W Branch, Chatschik Bisdikian, and Bo Yang. Efficient network management for context-aware participatory sensing. In *Proc. IEEE SECON*, pages 116–124, 2011.

[127] Cong Liu and Jie Wu. Scalable routing in cyclic mobile networks. *Parallel and Distributed Systems, IEEE Transactions on*, 20(9):1325–1338, 2009.

[128] Hengchang Liu, Shaohan Hu, Wei Zheng, Zhiheng Xie, Shiguang Wang, Pan Hui, and Tarek Abdelzaher. Efficient 3g budget utilization in mobile participatory sensing applications. In *Proc. IEEE INFOCOM*, pages 1411–1419, 2013.

[129] R. Lu, X. Lin, H. Zhu, X. Shen, and B. Preiss. Pi: A practical incentive protocol for delay tolerant networks. *IEEE Trans. Wireless Commun.*, 9(4):1483–1493, 2010.

[130] Tie Luo, Hwee-Pink Tan, and Lirong Xia. Profit-maximizing incentive for participatory sensing. In *Proc. IEEE Infocom*, 2014.

[131] M. Ma and Y. Yang. Sencar: An energy-efficient data gathering mechanism for large-scale multihop sensor networks. *IEEE Transactions on Parallel and Distributed Systems*, pages 1476–1488, 2007.

[132] R.T.B. Ma, S. Lee, J. Lui, and D.K.Y. Yau. Incentive and service differentiation in p2p networks: a game theoretic approach. *IEEE/ACM Trans. Netw.*, 14(5):978–991, 2006.

[133] Elke Mackensen, Matthias Lai, and Thomas M Wendt. Bluetooth low energy (ble) based wireless sensors. In *IEEE Sensors*, pages 1–4, 2012.

[134] Mateusz Malinowski, Matthew Moskwa, Mark Feldmeier, Mathew Laibowitz, and Joseph A Paradiso. Cargonet: a low-cost micropower sensor node exploiting quasi-passive wakeup for adaptive asychronous monitoring of exceptional events. In *Proceedings of the 5th international conference on Embedded networked sensor systems*, pages 145–159. ACM, 2007.

[135] Howard Margolis. *Selfishness, altruism, and rationality.* University of Chicago Press, 1984.

[136] Emiliano Miluzzo, Nicholas D Lane, Kristóf Fodor, Ronald Peterson, Hong Lu, Mirco Musolesi, Shane B Eisenman, Xiao Zheng, and Andrew T Campbell. Sensing meets mobile social networks: the design, implementation and evaluation of the cenceme application. In *Proc. ACM SenSys*, pages 337–350, 2008.

[137] S.H. Moeller, A. Sridharan, B. Krishnamachari, and O. Gnawali. Routing without routes: the backpressure collection protocol. In *Proc. ACM/IEEE IPSN*, 2010.

[138] Prashanth Mohan, Venkata N Padmanabhan, and Ramachandran Ramjee. Nericell: rich monitoring of road and traffic conditions using mobile smartphones. In *Proceedings of the 6th ACM conference on Embedded network sensor systems*, pages 323–336. ACM, 2008.

[139] Michael J Neely. Stochastic network optimization with application to communication and queueing systems. *Synthesis Lectures on Communication Networks*, 3(1):1–211, 2010.

[140] M.J. Neely. *Dynamic power allocation and routing for satellite and wireless networks with time varying channels.* PhD thesis, MIT, 2003.

[141] M.J. Neely. Optimal pricing in a free market wireless network. In *Proc. IEEE Infocom*, pages 213–221, 2007.

[142] M.J. Neely. Universal scheduling for networks with arbitrary traffic, channels, and mobility. In *Proc. IEEE CDC*, pages 1822–1829, 2010.

[143] M.J. Neely and L. Golubchik. Utility optimization for dynamic peer-to-peer networks with tit-for-tat constraints. In *Proc. IEEE Infocom*, pages 1458–1466, 2011.

[144] M.J. Neely, A.S. Tehrani, and A.G. Dimakis. Efficient algorithms for renewable energy allocation to delay tolerant consumers. In *Proc. IEEE SmartGridComm*, pages 549–554, 2010.

[145] M.J. Neely and R. Urgaonkar. Optimal backpressure routing for wireless networks with multi-receiver diversity. *Ad Hoc Networks*, 7(5):862–881, 2009.

[146] Noam Nisan and Amir Ronen. Algorithmic mechanism design. In *Proc. ACM STOC*, pages 129–140, 1999.

[147] D. Niyato and E. Hossain. Competitive pricing for spectrum sharing in cognitive radio networks: Dynamic game, inefficiency of nash equilibrium, and collusion. *IEEE J. Sel. Areas Commun.*, 26(1):192–202, 2008.

[148] M. O'Grady and G. O'Hare. How smart is your city? *Science*, 335(6076):1581–1582, 2012.

[149] Melek Onen, Abdullatif Shikfa, and Refik Molva. Optimistic fair exchange for secure forwarding. In *Mobile and Ubiquitous Systems: Networking & Services, 2007. MobiQuitous 2007. Fourth Annual International Conference on*, pages 1–5. IEEE, 2007.

[150] G. Palla, I. Derényi, I. Farkas, and T. Vicsek. Uncovering the overlapping community structure of complex networks in nature and society. *Nature*, 435(7043):814–818, 2005.

[151] U. Park and J. Heidemann. Data muling with mobile phones for sensornets. In *Porc. ACM Sensys*, pages 162–175, 2011.

[152] David C Parkes and Jeffrey Shneidman. Distributed implementations of vickrey-clarke-groves mechanisms. In *Proc. AAMAS*, pages 261–268, 2004.

[153] L. Pelusi, A. Passarella, and M. Conti. Opportunistic networking: data forwarding in disconnected mobile ad hoc networks. *Communications Magazine, IEEE*, 44(11):134–141, 2006.

[154] RADHA Poovendran. Cyber–physical systems: Close encounters between two parallel worlds [point of view]. *Proceedings of the IEEE*, 98(8):1363–1366, 2010.

[155] Kiran K Rachuri, Christos Efstratiou, Ilias Leontiadis, Cecilia Mascolo, and Peter J Rentfrow. Metis: Exploring mobile phone sensing offloading for efficiently supporting social sensing applications. In *Proc. IEEE PerCom*, pages 85–93, 2013.

[156] Rajib Kumar Rana, Chun Tung Chou, Salil S Kanhere, Nirupama Bulusu, and Wen Hu. Ear-phone: an end-to-end participatory urban noise mapping system. In *Proc. IPSN*, pages 105–116, 2010.

[157] M. Rasti, A.R. Sharafat, and B. Seyfe. Pareto-efficient and goal-driven power control in wireless networks: A game-theoretic approach with a novel pricing scheme. *IEEE/ACM Trans. Netw.*, 17(2):556–569, 2009.

[158] J. Ryu, V. Bhargava, N. Paine, and S. Shakkottai. Back-pressure routing and rate control for icns. In *Proc. ACM Mobicom*, pages 365–376, 2010.

[159] J. Ryu, L. Ying, and S. Shakkottai. Back-pressure routing for intermittently connected networks. In *Proc. IEEE Infocom*, pages 1–5, 2010.

[160] V.P. Sadaphal and B.N. Jain. Random and Periodic sleep schedules for target detection in sensor networks. *Journal of Computer Science and Technology*, 23(3):343–354, 2008.

[161] Hans Schaffers, Nicos Komninos, Marc Pallot, Brigitte Trousse, Michael Nilsson, and Alvaro Oliveira. Smart cities and the future internet: towards cooperation frameworks for open innovation. In *The future internet*, pages 431–446. 2011.

[162] John Scott and Peter J Carrington. *The SAGE handbook of social network analysis*. SAGE publications, 2011.

[163] Winston KG Seah, Zhi Ang Eu, and Hwee-Pink Tan. Wireless sensor networks powered by ambient energy harvesting (wsn-heap)-survey and challenges. In *Wireless Communication, Vehicular Technology, Information Theory and Aerospace & Electronic Systems Technology, 2009. Wireless VITAE 2009. 1st International Conference on*, pages 1–5. IEEE, 2009.

[164] Shitiz Sehgal, Salil S Kanhere, and Chun Tung Chou. Mobishop: Using mobile phones for sharing consumer pricing information. In *Demo Session of the Intl. Conference on Distributed Computing in Sensor Systems*, 2008.

[165] Kewei Sha, Guoxing Zhan, Weisong Shi, Mark Lumley, Clairy Wiholm, and Bengt Arnetz. Spa: a smart phone assisted chronic illness self-management system with participatory sensing. In *Proceedings of the 2nd International Workshop on Systems and Networking Support for Health Care and Assisted Living Environments*, page 5. ACM, 2008.

[166] G. Sharma, R.R. Mazumdar, and N.B. Shroff. On the complexity of scheduling in wireless networks. In *Proc. ACM Mobicom*, pages 227–238, 2006.

[167] Xiang Sheng, Jian Tang, and Weiyi Zhang. Energy-efficient collaborative sensing with mobile phones. In *IEEE INFOCOM*, pages 1916–1924, 2012.

[168] U. Shevade, H.H. Song, L. Qiu, and Y. Zhang. Incentive-aware routing in dtns. In *Proc. IEEE ICNP*, pages 238–247, 2008.

[169] Yoav Shoham and Kevin Leyton-Brown. *Multiagent systems: Algorithmic, game-theoretic, and logical foundations*. Cambridge University Press, 2009.

[170] A.A. Somasundara, A. Kansal, D.D. Jea, D. Estrin, and M.B. Srivastava. Controllably mobile infrastructure for low energy embedded networks. *IEEE Transactions on Mobile Computing*, pages 958–973, 2006.

[171] Thrasyvoulos Spyropoulos, Konstantinos Psounis, and Cauligi S Raghavendra. Spray and wait: an efficient routing scheme for intermittently connected mobile networks. In *Proceedings of the 2005 ACM SIGCOMM workshop on Delay-tolerant networking*, pages 252–259. ACM, 2005.

[172] A. Sridharan, S. Moeller, B. Krishnamachari, and M. Hsieh. Implementing backpressure-based rate control in wireless networks. In *Proc. IEEE ITA workshop*, pages 341–345, 2009.

[173] Avinash Sridharan and Bhaskar Krishnamachari. Explicit and precise rate control for wireless sensor networks. In *Proc. ACM SenSys*, pages 29–42, 2009.

[174] Ryo Sugihara and Rajesh K. Gupta. Optimal speed control of mobile node for data collection in sensor networks. *IEEE Transactions on Mobile Computing*, 9:127–139, 2010.

[175] Ryo Sugihara and Rajesh K. Gupta. Speed control and scheduling of data mules in sensor networks. *ACM Trans. Sen. Netw.*, 7:4:1–4:29, August 2010.

[176] Y. Sun, S. Du, O. Gurewitz, and D.B. Johnson. DW-MAC: a low latency, energy efficient demand-wakeup MAC protocol for wireless sensor networks. In *Proceedings of the 9th ACM international symposium on Mobile ad hoc networking and computing*, pages 53–62. ACM, 2008.

[177] Y. Sun, O. Gurewitz, and D.B. Johnson. RI-MAC: a receiver-initiated asynchronous duty cycle MAC protocol for dynamic traffic loads in wireless sensor networks. In *Proceedings of the 6th ACM conference on Embedded network sensor systems*, pages 1–14. ACM, 2008.

[178] S. Suthaharan, A. Chawade, R. Jana, and J. Deng. Energy Efficient DNA-Based Scheduling Scheme for Wireless Sensor Networks. *Wireless Algorithms, Systems, and Applications*, pages 459–468, 2009.

[179] Takashi Tanaka, Farhad Farokhi, and Cédric Langbort. Faithful implementations of distributed algorithms and control laws. *arXiv preprint arXiv:1309.4372*, 2013.

[180] Chayant Tantipathananandh, Tanya Berger-Wolf, and David Kempe. A framework for community identification in dynamic social networks. In *Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 717–726. ACM, 2007.

[181] L. Tassiulas and A. Ephremides. Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks. *IEEE Trans. Autom. Control*, 37(12):1936–1948, 1992.

[182] RD Team et al. Rpl: Routing protocol for low power and lossy networks. *IETF ROLL WG, IETF Internet-Draft*, 4, 2009.

[183] O. Tekdas, V. Isler, J.H. Lim, and A. Terzis. Using mobile robots to harvest data from sensor fields. *IEEE Wireless Commun.*, 16(1):22–28, 2009.

[184] Arvind Thiagarajan, Lenin Ravindranath, Katrina LaCurts, Samuel Madden, Hari Balakrishnan, Sivan Toledo, and Jakob Eriksson. Vtrack: accurate, energy-aware road traffic delay estimation using mobile phones. In *Proc. ACM SenSys*.

[185] Cristian Tuduce and Thomas Gross. A mobility model based on wlan traces and its validation. In *INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE*, volume 1, pages 664–674. IEEE, 2005.

[186] Amin Vahdat, David Becker, et al. Epidemic routing for partially connected ad hoc networks. Technical report, Technical Report CS-200006, Duke University, 2000.

[187] L. Van Hoesel and P. Havinga. A lightweight medium access protocol (LMAC) for wireless sensor networks. In *1st Int. Workshop on Networked Sensing Systems (INSS 2004)*. Citeseer, 2004.

[188] P. Venkitasubramaniam, S. Adireddy, and L. Tong. Sensor networks with mobile access: optimal random access and coding. *Selected Areas in Communications, IEEE Journal on*, 22(6):1058–1068, 2004.

[189] N. Vlajic and D. Stevanovic. Sink mobility in wireless sensor networks: a (mis) match between theory and practice. In *Proceedings of the 2009 International Conference on Wireless Communications and Mobile Computing: Connecting the World Wirelessly*, pages 386–393. ACM, 2009.

[190] Leye Wang, Daqing Zhang, and Haoyi Xiong. effsense: energy-efficient and cost-effective data uploading in mobile crowdsensing. In *Proc. ACM UbiComp*, pages 1075–1086, 2013.

[191] W. Wang, V. Srinivasan, and K.C. Chua. Using mobile relays to prolong the lifetime of wireless sensor networks. In *Proceedings of the 11th annual international conference on Mobile computing and networking*, pages 270–283. ACM, 2005.

[192] Lifei Wei, Haojin Zhu, Zhenfu Cao, and Xuemin Sherman Shen. Mobiid: a user-centric and social-aware reputation based incentive scheme for delay/disruption tolerant networks. In *Ad-hoc, mobile, and wireless networks*, pages 177–190. Springer, 2011.

[193] R. Wohlers, N. Trigoni, R. Zhang, and S. Ellwood. Twinroute: Energy-efficient data collection in fixed sensor networks with mobile sinks. In *Proc. IEEE MDM*, pages 192–201, 2009.

[194] K.J. Wong, B.S. Lee, B.C. Seet, G. Liu, and L. Zhu. Busnet: Model and usage of regular traffic patterns in mobile ad hoc networks for inter-vehicular communications. In *Proc. ICT*. Citeseer, 2003.

[195] Xiuchao Wu, Kenneth N Brown, and Cormac J Sreenan. Analysis of smartphone user mobility traces for opportunistic data collection in wireless sensor networks. *Pervasive and Mobile Computing*, 2013.

[196] G. Xing, T. Wang, Z. Xie, and W. Jia. Rendezvous planning in mobility-assisted wireless sensor networks. In *rtss*, pages 311–320. IEEE Computer Society, 2007.

[197] H. Xiong, R. Li, A. Eryilmaz, and E. Ekici. Delay-aware cross-layer design for network utility maximization in multi-hop networks. *IEEE J. Sel. Areas Commun.*, 29(5):951–959, 2011.

[198] Dejun Yang, Guoliang Xue, Xi Fang, and Jian Tang. Crowdsourcing to smartphones: incentive mechanism design for mobile phone sensing. In *Proc. ACM Mobicom*, pages 173–184, 2012.

[199] S. Yang, X. Yang, C. Zhang, and E. Spyrou. Using social network theory for modeling human mobility. *IEEE Network*, 24(5):6–13, 2010.

[200] Shusen Yang, Usman Adeel, and Julie A McCann. Selfish mules: Social profit maximization in sparse sensornets using rationally-selfish human relays. *IEEE J. Sel. Areas Commun.*, 31(6):1124–1134, 2013.

[201] W. Ye, J. Heidemann, and D. Estrin. An energy-efficient MAC protocol for wireless sensor networks. In *INFOCOM 2002. Twenty-First Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, volume 3, pages 1567–1576. IEEE, 2002.

[202] S. Yessad, F. Nait-Abdesselam, T. Taleb, and B. Bensaou. R-MAC: Reservation medium access control protocol for wireless sensor networks. In *Local Computer Networks, 2007. LCN 2007. 32nd IEEE Conference on*, pages 719–724. IEEE, 2007.

[203] Y. Yi and M. Chiang. Stochastic network utility maximisation–a tribute to kelly's paper published in this journal a decade ago. *European Transactions on Telecommunications*, 19(4):421–442, 2008.

[204] J. Yick and B. Mukherjee. Wireless sensor network survey. *Computer Networks*, 52(12):2292–2330, 2008.

[205] Jennifer Yick, Biswanath Mukherjee, and Dipak Ghosal. Wireless sensor network survey. *Computer networks*, 52(12):2292–2330, 2008.

[206] L. Ying, S. Shakkottai, A. Reddy, and S. Liu. On combining shortest-path and back-pressure routing over multihop wireless networks. *IEEE/ACM Trans. Netw.*, 19(3):841–854, 2011.

[207] Jungkeun Yoon, Brian D Noble, Mingyan Liu, and Minkyong Kim. Building realistic mobility models from coarse-grained traces. In *Proceedings of the 4th international conference on Mobile systems, applications and services*, pages 177–190. ACM, 2006.

[208] X. Zhang, J. Kurose, B.N. Levine, D. Towsley, and H. Zhang. Study of a bus-based disruption-tolerant network: mobility modeling and impact on routing. In *Proceedings of the 13th annual ACM international conference on Mobile computing and networking*, pages 195–206. ACM, 2007.

[209] X. Zhang, J. Kurose, B.N. Levine, D. Towsley, and H. Zhang. Study of a bus-based disruption-tolerant network: mobility modeling and impact on routing. In *Proc. ACM Mobicom*, pages 195–206, 2007.

[210] Sheng Zhong, Jiang Chen, and Yang Richard Yang. Sprite: A simple, cheat-proof, credit-based system for mobile ad-hoc networks. In *INFOCOM 2003. Twenty-Second Annual Joint Conference of the IEEE Computer and Communications. IEEE Societies*, volume 3, pages 1987–1997. IEEE, 2003.

[211] Xuejun Zhuo, Wei Gao, Guohong Cao, and Yiqi Dai. Win-coupon: An incentive framework for 3g traffic offloading. In *Proc. IEEE ICNP*, pages 206–215, 2011.