Energy Efficient Machine-Type Communications over Cellular Networks

A Battery Lifetime-Aware Cellular Network Design Framework

AMIN AZARI

Licentiate Thesis in Information and Communication Technology
School of Information and Communication Technology
KTH Royal Institute of Technology
Stockholm, Sweden 2016

KTH School of Information and
Communication Technology
SE-164 40 Kista
SWEDEN

TRITA-ICT 2016:34
ISBN 978-91-7729-162-6

Akademisk avhandling som med tillstånd av Kungl Tekniska högskolan framlägges till offentlig granskning för avläggande av licentiatexamen i Informations- och kommunikationsteknik fredagen den 2 december 2016 klockan 10:00 i Sal B, Electrum, Kungl Tekniska högskolan, Kistagången 16, Kista.

Tryck: Universitetsservice US AB

Abstract

Internet of Things (IoT) refers to the interconnection of uniquely identifiable smart devices which enables them to participate more actively in everyday life. Among large-scale applications, cheap and widely spread machine-to-machine (M2M) communications supported by cellular networks, also known as machine-type communications (MTC), will be one of the most important enablers for the success of IoT. As the existing cellular infrastructure has been optimized for a small number of long-lived communications sessions, serving a massive number of machine-type devices with extremely diverse quality of service requirements is a big challenge. Also, most machine nodes are battery-driven, and hence, long battery life is crucial for them especially when deployed in remote areas.

The present work is devoted to energy consumption modeling, battery lifetime analysis, and lifetime-aware network design for massive MTC over cellular networks. We first develop a realistic energy consumption model for MTC, based on which, network battery lifetime models are defined. To address the massive concurrent access issue and save energy in data transmission, we first consider cluster-based MTC and let machine devices organize themselves locally, create machine clusters, and communicate through the cluster-heads to the base-station (BS). Toward this end, we need to find where clustering is feasible, and how the clusters must be formed. These research problems as well as some other aspects of cluster-based MTC are investigated in this work, battery lifetime-aware solutions are derived, and performance evaluation for the proposed solutions are provided.

For direct communications of the unclustered nodes and cluster-heads with the BS, we investigate the potential benefit in lifetime-aware uplink scheduling and transmit power control. Analytic expressions are derived to demonstrate the impact of scheduling on the individual and network battery lifetimes. The derived expressions are subsequently employed in uplink scheduling and transmit power control for mixed-priority MTC traffic in order to maximize the network lifetime. Besides the main solutions, low-complexity solutions with limited feedback requirement are also investigated.

Finally, we investigate the impact of energy saving for the access network on battery lifetime of machine-type devices. We present a queuing system to model the uplink transmission of a green base station which serves two types of distinct traffics with strict requirements on delay and battery lifetime. Then, the energy-lifetime and energy-delay tradeoffs are introduced, and closed-form expressions for energy consumption of the base station, average experienced delay in data transmission, and expected battery lifetime of machine devices are derived. Numerical results show the impact of energy saving for the access network on the introduced tradeoffs, and figure out how to trade the tolerable lifetime/delay of the users for energy saving in the access network.

The derived solutions are finally extended to the existing LTE networks, and simulation results in the context of LTE are presented. The simulation results show that the proposed solutions can provide substantial network lifetime improvement and network maintenance cost reduction in comparison with the existing approaches.

Sammanfattning

Sakernas Internet hänvisar till sammankoppling av smarta enheter som gör det möjligt för dem att delta mer aktivt i det dagliga livet. Billig och utbredd maskin-till-maskin (M2M) kommunikation, även känd som maskintypisk kommunikation (MTC) stödd av mobilnät, förväntas spela en viktig roll i förverkligandet av sakernas Internet. I och med att befintlig mobil infrastruktur optimerats för ett litet antal långlivade kommunikationssessioner, blir servandet av ett överväldigande antal MTC-enheter med extremt olika servicekvalitetskrav en stor utmaning. Dessutom är de flesta maskinnoder batteridrivna. Lång batteritid är därför avgörande när de ska distribueras till avlägsna områden.

Detta arbete fokuserar på energiförbrukningsmodellering, batterilivslängdsanalys, och livslängdsmedveten nätverksdesign för massiv MTC över mobilnät. Vi utvecklar först en realistisk energiförbrukningsmodell för MTC som ligger till grund för definierandet av nätverkets batterilivslängdsmodeller. För att lösa den massiva samtidiga åtkomstproblematiken och spara energi i dataöverföringen så betraktar vi först klusterbaserad MTC och låter maskinerna organisera sig lokalt, skapa maskinkluster, och via klusterhuvuden kommunicera med basstationen. För detta ändamål behövs tas reda på var klustring är genomförbart och hur klustren bör formas. Dessa forskningsproblem tillsammans med några andra aspekter av klusterbaserad MTC undersöks i detta arbete, liksom batterilivslängdsvänliga lösningar och utvärdering av prestanda för de föreslagna lösningarna.

För direktkommunikation av icke-klustrade noder och klusterhuvuden med basstationen undersöks också nyttan av livstidsvänlig schemaläggning i upplänk och effektkontroll. Analytiska uttryck härleds för att visa på effekterna av schemaläggning i batterilivslängd på individuell- och nätverksnivå. De härledda uttrycken används sedan för upplänksschemaläggning och effektkontroll för blandad prioritet av MTC-trafik för att maximera nätverks livstid.

Slutligen undersöker vi effekterna på batterilivslängd av energibesparing i ett accessnät med maskintypiska anordningar. Vi presenterar ett kösystem för att modellera upplänksöverföring av en grön basstation som tjänar två typer av trafik med höga krav på fördröjning och batterilivslängd. Därefter presenteras avvägningar mellan energilivstid och energifördröjning, och uttryck i sluten form för energiförbrukning av basstationen, genomsnittlig fördröjning i dataöverföringen, och den förväntade livslängden av maskinanordningar härleds. Numeriska resultat visar på effekterna av energibesparingar för basstationen med de införda avvägningarna, samt räknar ut hur man växlar mellan användarnas tolerabla livstid/fördröjning mot energibesparingar i accessnätet.

De härledda lösningarna utvidgas slutligen till befintliga LTE-nät i syfte att utvärdera deras prestanda i praktiken. Simuleringsresultat visar att de föreslagna lösningarna kan ge betydande nätverkslivstidsförbättring och kostnadsminskning för nätunderhåll jämfört med befintliga metoder.

... To Montazar.

Acknowledgments

I am sure that it would not have been possible for me to reach to this step in my life without the help from God. To be honest, I entered Sweden to tolerate living abroad, specially in a cold region; however, after a short time I found it an amazing second home, which benefits from a hospitable society. I found the RS Lab a very exciting place for research, empowered by research scholars with a variety of research experiences. To name the people, first I would like to express my sincere gratitude to my main advisor, Associate Professor Guowang Miao for giving me the opportunity to take this position, encouraging me to do my best in the research, and for his kind support. I would like to offer my special thanks to Professor Jens Zander, my co advisor, for giving me the opportunity to work under his supervision, and benefit from his valuable comments. I am also grateful to all other members of RS Lab because of their support in dealing with lots of problems that a new student may deal with them abroad. Special thanks to all members who have contributed to my research by providing feedback in Friday seminars. Also, I am grateful to Dr. Joachim Sachs, Associate Professor Markus Hidell, and Associate Professor Cedomir Stefanovic for reviewing licentiate proposal, reviewing licentiate thesis, and accepting the role of opponent for my defense, respectively. I don't know how it is possible to express my gratefulness by some words to my wife, Mohaddeseh, for being here far from her parents to support me, but I am sure that our newborn son is a heavenly gift for her. Finally, I am always grateful to my parents Fatemeh and Alimadad, brothers Mohammad Kazem and Iman, and sister Zahra for their support, and to all others who have prayed for me.

Life is the unique scene for presentation of everyone's art.
Everyone sangs his song and leaves the scene, but scene is fixed.
Blessed is the one whose song is remembered by the people.
Jaleh Esfahani, Persian poet, 1921-2007

زندگی صحنه یکتای هنرمندی ماست. هرکسی نغمه خود خواند و از صحنه رود،
صحنه پیوسته به جاست. خرم آن نغمه که مردم بسپارند به یاد. - ژاله اصفهانی -

Contents

List of Figures

Part I

Thesis Overview

Chapter 1

Introduction

Internet of Things (IoT) refers to the ever-growing network of smart physical objects that are capable of sensing or acting on their environment, and the communications between these objects and other Internet-enabled systems. IoT enables smart devices to participate more actively in everyday life, business, industry, and health care. Among large-scale applications, cheap and widely spread machine-to-machine (M2M) communications supported by cellular networks, also known as machine-type communications (MTC), will be one of the key important enablers for the success of IoT [1]. M2M communications, where IoT will evolve from it [2], means the communications of machine devices with reduced or without human intervention [3], and is applicable to data gathering, smart metering, health and presence monitoring, and so on [4]. The continuing growth in demand from cellular-based MTC, 34-fold from 2014 to 2019 [5], encourages mobile networks operators (MNOs) to investigate new radio access technologies for accommodating MTC traffic in cellular networks in order to decrease the revenue gap [6], [7]. To this end, the use cases, characteristics, and requirements of MTC should be well investigated.

1.1 Background

1.1.1 Characteristics of MTC

To facilitate network design, the 3GPP and IEEE have defined specific features for MTC like low-mobility, time-controlled, and infrequent transmission [8]-[9]. The low-mobility feature corresponds to the applications in which machine devices do not move or move infrequently within a certain area. The time-controlled feature is intended for use with MTC applications that can tolerate to send data only during defined time intervals. Machine-type communications are also characterized by the massive number of concurrent active devices, and short payload size. However, there are some MTC applications like video surveillance which occasionally require long payload size data transmissions. Moreover, in most MTC applications machine devices are battery driven and once deployed; their batteries will never be replaced.

Hence, long battery life is crucial for them especially for devices deployed in remote areas. The characteristics of MTC traffic are also fundamentally different from those of human-oriented-communications (HoC) because a great portion of the exchanged information packets in MTC are not generated by human behavior any more. Finally, machine-type communications are known for their vastly diverse quality-of-service (QoS) requirements [10]. For example, the tolerated delay for data transmission may vary from milliseconds to hours.

1.1.2 Main Challenges in Enabling MTC

As mentioned, most smart devices like sensors are battery-driven and energy-efficient transmission is crucial for them. Telecommunications industry has spent a great deal of resources investigating how to realize high-throughput infrastructure but forgotten about scalable low-power low-rate systems in which, energy efficiency and battery lifetime are crucially important. Regarding the fundamental differences between MTC and human-oriented communications, many research works have been launched to understand how existing cellular infrastructure needs to change to provide large-scale energy-efficient access for cellular MTC [11]-[17]. To enable MTC over cellular networks, at least three major issues must be addressed [18].

The first issue is providing a scalable medium access protocol (MAC) which is able to serve a massive number of concurrent access requests. The existing cellular networks have been designed and optimized for a small number of long-lived HoC sessions, and are mainly downlink-oriented, i.e. the amount of uplink traffic is normally lower than the downlink traffic in serving HoC traffic. When massive MTC services meet cellular networks, it is expected that cellular networks cannot survive with a massive number of short-lived MTC sessions in the uplink direction. Hence, network congestion, including radio network congestion and signaling network congestion, is likely to happen.

The second issue which must be tackled is energy efficiency in small data transmissions. Consider the case that MTC traffic is to be sent through the LTE infrastructure. Then, around 34 bytes overhead data for user datagram protocol (UDP), Internet protocol (IP), packet data convergence protocol, radio link control, and the MAC overhead must be added to the original data [19]. When LTE is used for data-hungry applications installed in smart-phones charged in a daily manner, this level of overhead is reasonable. However, in MTC applications where machine nodes have only a few bits of data to transmit, and the energy is not renewable, this level of overhead makes the communications inefficient.

The third issue concerns satisfying QoS for mission-critical MTC. Some MTC applications like autonomous control require a very short delay with extremely high reliability. As this level of delay and its respective reliability cannot be achieved using existing cellular infrastructure, enabling mission-critical MTC over cellular networks would require revolutionary access schemes [20], which is a hot topic in the 5G standardization.

Further than the above mentioned node level challenges, at the network level the fundamental tradeoffs that serving massive MTC services introduces to cellular networks have not been investigated. Given the limited set of resources available at the network side, a decision on the amount of resources that can be allocated or shared by new MTC services is challenging because it can affect QoS of existing subscribers, energy consumption of the base stations, and battery lifetime of machine-type devices. Overcoming the above mentioned challenges pave the way to realize battery-friendly cellular networks that can support massive IoT services.

Apart from the cellular solutions, some standardization efforts like IEEE 802.15.4 and WirelessHART, and some non-standardized efforts like SigFox aim at providing scalable connectivity for energy-limited smart devices [21]. For example, IEEE 802.15.4 specifies the physical and MAC layer for communications using low-power, low-rate, and low-complexity radio frequency transmissions in a wireless personal area network (WPAN). In this thesis, we mainly focus on cellular-based solutions for enabling massive IoT services.

1.1.3 Thesis Focus and High Level Research Questions

Motivated by the fact that the number of connected machine-type devices in 2020 is expected to be nearly 7 billion [22], the introduction of massive MTC services is an opportunity for mobile network operators to decrease their revenue gap by providing new services like Internet of Things over cellular networks. Then, in this thesis we will mainly focus on battery lifetime of device side and energy consumption of network side of cellular networks, and seek energy saving solutions for serving a massive number of MTC devices. Therefore, the main goal of this thesis is:
to incorporate battery lifetime-awareness into the design of future cellular networks in order to identify solutions enabling serving a massive number of energy limited devices with minimum increase in network deployment and operational costs and without degrading human-type users perceived quality of service.

Based on this main goal, two high level research problems addressed in this dissertation are as follows:

- HQ1: How should cellular networks be deployed and operated in a battery lifetime-aware manner?
- HQ2: What are the consequences of battery lifetime-aware solutions on the QoS of other users and energy consumption of the access network?

1.2 Related Works

To provide massive machine access, curb undesirable energy wastage, reduce the costs, and extend the battery lifetime of machine devices, a variety of techniques have been proposed in literature. In this chapter, we explore the necessary background for the thesis. More detailed related works can be found in the appended papers A-D.

1.2.1 MAC Design for MTC

1.2.1.1 MAC Design for Direct MTC

Random access channel (RACH) of the LTE-Advanced is the typical way for machine nodes to access the base station [23]. The capacity limits of RACH for serving MTC and a survey of improved alternatives have been studied in [24]. Network congestion, including radio network congestion and signaling network congestion as defined in [25], is likely to happen when a massive number of MTC devices try to connect to the BS. Therefore, several solutions have been proposed in [18] to reduce congestion in overloaded scenarios. Among these solutions, access class barring (ACB) has attracted lots of attentions in literature [26]. In this scheme, when a device needs cellular connectivity, it decides to access the network with probability q_a or defer its access with probability $1 - q_a$. The corresponding access probability q_a for each class of nodes/traffic is broadcasted by the base station (BS). The BS can introduce one or several access classes depending on the granularity of the control needed over MTC [18]. In [27], the authors propose to divide each communications frame into two periods: one for contention-based resource reservation, and the other for data transmission of successful nodes in resource reservation. While the proposed schemes in [27] and [26] can save energy by preventing collisions in data transmission to some extent, they require machine nodes to be active for a longer time to get channel access, which is not energy efficient. A time-controlled cellular access framework satisfying the delay requirements of massive MTC scenarios is proposed in [28], where the authors propose to divide machine nodes into classes based on QoS requirements and fixed access grants are reserved to occur for each class. Power-efficient MAC protocols for machine devices with reliability constraints are investigated in [29]. While the above mentioned solutions can ease the communications of a massive number of machine devices over cellular networks, enabling massive MTC services over cellular networks requires a novel MAC protocol which tackles scalability and energy efficiency at the same time. The energy-efficient massive concurrent access control to the shared wireless medium is still an open problem.

1.2.1.2 MAC Design for Clustered MTC

Cooperation among machine devices and a cluster-based operation of them is a promising approach for offloading the BS [10]. In addition to BS offloading, it has been shown that network-assisted clustering can significantly prolong the network battery lifetime [30]. Cluster-based MTC systems can be categorized into 6 categories with regards to the type of data gathering nodes and radio resources used for intra-cluster communications. Responsible nodes for data gathering can be both either (i) MTC devices; or (ii) special UEs, called gateways, which have more advanced radio front-end, data transmission, and energy storage capabilities than MTC devices. Regarding to the radio resource used for intra-cluster communications, cluster members (CMs) can communicate with the cluster heads (CHs) either (i) over dedicated cellular radio resources; (ii) reused cellular radio resources; or (iii) over unlicensed radio resources, e.g. over WiFi links. Integrating MTC gateways in cellular network enables handling a massive number of concurrent access requests using capillary networking, and extends cellular coverage to remote areas [10, 15]. Capillary networking has been studied in [31], where the authors claim that it will become a key enabler of the future networked society. Addressing the massive concurrent access problem using MTC gateways for existing LTE networks is investigated in [32]. In [33], the authors show that LTE networks would benefit from already deployed WiFi gateways in the indoor environments by implementing a flexible capillary admission control.

An emerging communication paradigm in cellular networks is direct Device-to-Device (D2D) communications, which are generally non-transparent to the cellular network, and are done with reduced intervention from the BS. D2D communications can occur either on the cellular radio resources, or unlicensed spectrum [34]. Introduction of D2D communications to cellular networks motivates the idea to relay the MTC traffic gathered/generated by MTC devices through the D2D links [35]. Relaying MTC traffic over D2D links has been investigated in [36]. The same problem has been considered in [37], where a tradeoff between latency and the transmit power, which is needed to deliver the aggregate traffic, is presented. In [38], a multi-hop routing scheme for MTC traffic incorporating opportunistic D2D links is presented. Without a fixed a priory installed special gateway or an opportunistic UE which is accessible and secure to relay data, each MTC device could act as a CH [35]. In this case, the data gathering nodes are also energy-limited, and hence, the choice of cluster size, communications protocol inside clusters, and CH (re)selection scheme is of paramount importance. Benefits and challenges of cooperation in massive MTC scenarios are investigated in [39].

1.2.2 MTC Scheduling

Scheduling is the process performed by the BS to assign radio resources to user equipments. In cellular networks, scheduling is not part of the standardization work, and is usually left for vendor implementation. However, due to the fact that signaling is standardized, each implemented schedulers should comply with the control requirements specified in the standards. Broadly speaking, one may categorize the existing MTC scheduling algorithms into 4 main categories as follows [40]: (i) channel-based schedulers, which prioritize devices with the highest signal to noise ratio (SNR) in order to maximize the system throughput [41, 42, 43]; (ii) delay-based schedulers, in which the delay budget prioritizes devices for resource allocation [41, 44]; (iii) fairness-based schedulers, in which a fair distribution of radio resources among UEs is guaranteed [45]; and (iv) hybrid schedulers, which consider a combination of the aforementioned metrics as well as other metrics like buffer status and power consumption [46, 40]. One sees while providing scalable yet energy efficient connectivity is considered as the key requirement for successful deployment of MTC services over cellular networks [47, 24], energy efficient uplink

MTC scheduling is almost absent in literature. Literature study on MTC scheduling over LTE networks reveals that delay, throughput, transmit power, and impact on QoS of HoC have been the frequently used scheduling metrics in previous works [40]. In [15], major challenges in MTC scheduling over LTE networks have been studied. Energy efficiency of small data communications over LTE networks has been studied in [19], where it is shown that LTE physical layer is not energy efficient for small data communications. Design of power-efficient uplink scheduler for delay-sensitive traffic over LTE systems has been investigated in [48], where the considered delay models and traffic characteristics are not consistent with the ones of MTC [8]. Uplink scheduling with MTC/HoC coexistence over cellular networks has been investigated in [46]. In [46], a simple energy consumption model has been used for MTC, which takes only the transmit power for reliable data transmission into account, and hence, the other sources of energy consumptions, including energy consumptions in operation of electronic circuits and access reservation, which are comparable with the energy consumption for reliable data transmission [49, 50], have been neglected. To the best of our knowledge, accurate energy consumption modeling for machine-type communications, individual and network battery lifetime models, and corresponding scheduler design approaches are absent in literature.

1.2.3 Performance Tradeoffs in Serving MTC

A thorough survey on LTE scheduling algorithms for MTC traffic is presented in [40]. This survey indicates that the existing MTC scheduling algorithms are mainly focused on the time-domain scheduling, and vary in resource provisioning for MTC and defining the priority metric used to select the device to be scheduled [40]. Regarding the resource provisioning problem, the existing scheduling schemes are categorized as: (i) joint MTC/HoC scheduling, in which scheduler doesn't differentiate between MTC and HoC UEs in resource allocation [46]; (ii) opportunistic MTC scheduling, in which the remaining resource elements from HoC scheduling are allocated to machine devices [41], and (iii) separate MTC/HoC scheduling in which a set of resource elements are reserved for MTC [51]. While in [41, 51] the impact of resource allocation to MTC on QoS for HoC has been investigated using numerical analysis, to the best of our knowledge the tradeoffs between resource allocation to MTC/HoC, Battery lifetime of MTC devices, QoS of MTC/HoC, and spectral and energy efficiency of networks are absent in literature. Also, the impact of BS deployment and operation strategies, e.g. BS density and BS sleeping, on the battery lifetime of MTC devices, and hence the battery/cost tradeoffs, have not been investigated in literature.

1.3 Problem Statement

Based on the literature review, we see that the evaluation of individual/network battery lifetime, and the identification of lifetime-aware deployment and operation solutions are very important to enable massive MTC services over cellular networks. Another important aspect is the analysis of the impact of battery lifetime-aware solutions on the energy consumption and spectral efficiency of the network, as well as QoS of HoC traffic. Based on these aspects, the specific research problems addressed in this dissertation are defined in the sequel.

1.3.1 Battery lifetime Assessment

The initial problem faced in lifetime-aware cellular network design, and hence in answering to HQ1, is the lack of a widely-accepted methodology to evaluate the battery lifetime of a given network. Then, defining suitable battery lifetime metrics and characterizing the impact of system and traffic parameters on the battery lifetime are key aspects to determine optimized solutions. To the best of our knowledge, there is no model for individual and network battery lifetime of machine-type devices deployed over cellular networks. The lack of battery lifetime analysis in this field has resulted in proposing MTC enabling schemes which might be effective in handling a massive number of devices, but significantly shorten the network battery lifetime. While a few number of existing research works have tried to consider energy efficiency for MTC in their proposed solutions, their energy consumption models are overly simplified, e.g. the circuit energy consumption has been ignored. Therefore, in order to obtain consistent and realistic results concerning the network lifetime, which in turn leads us to lifetime-aware solutions, different sources of energy consumptions in MTC devices should be carefully considered. Therefore with regard to battery lifetime assessment, the following research question is addressed:

- RQ1: Which energy consumption sources should be considered in deriving individual and network battery lifetime models?
- RQ2: How to derive a low-complexity model of individual/network battery lifetime?

1.3.2 Battery Lifetime-Aware Solutions

Concerning the first high level research problem dealt in this dissertation, we seek solutions to maximize network battery lifetime. Towards this end, MAC design for direct and clustered MTC are to be investigated.

1.3.2.1 Battery Lifetime-Aware Solutions in MAC Design for Clustered MTC

Consider a massive machine deployment in a single cell scenario without strict delay requirement for MTC traffic, which in turn implies that the gathered data from MTC devices is transmitted based on the best effort. Our main aim here is to minimize the amount of consumed energy at terminals per bit of received data at the BS. A potential application with such characteristics is connected sensors deployed over cellular networks for data gathering, e.g. temperature/humidity/presence monitoring in an area. To address the massive concurrent access issue in case of triggering an event and save energy in data transmission, we consider cluster-based MTC and let machine devices organize themselves locally, create machine clusters, and communicate through the cluster-heads to the BS. The cluster-head can be a pre-installed gateway, a cellular user, or a machine-type device. In this chapter, we focus on the latter case, where the cluster-head is also energy limited and seeks long battery lifetime. With clustering, the number of concurrent channel access requests can be reduced and the lifetime of cluster members can be extended because of less collision and less transmission power. However, the expected battery lifetime of cluster heads will decrease due to the energy consumptions in listening to the channel and relaying packets from CMs to the BS. Therefore, it is necessary to develop a clustering scheme to improve the overall network lifetime by considering the energy consumptions in both CM and CH nodes. The clustering problem consists of finding the number of clusters and the cluster head in each cluster. Solving the joint problem is extremely complicated, if not impossible. Then, we can follow a decoupled approach, and derive the following research questions:

- RQ3: What is the optimal cluster-size?
- RQ4: What is the optimal cluster-head selection scheme?

Furthermore, in order to maximize the lifetime gain from cluster-based MTC, several other research problems must be addressed, as follows:

- RQ5: Where should clustering be used?
- RQ6: Which communications protocols must be used inside and outside the clusters?
- RQ7: What is the impact of underlying intra cluster communications on primary communications?

1.3.2.2 Battery Lifetime-Aware Solutions in MAC Design for Direct MTC

For direct communications of CHs and unclustered MTC devices with the BS, they should pass the RACH connection procedure, and send their scheduling requests to the BS through the physical uplink control channel (PUCCH). The BS performs the scheduling and sends back the scheduling grants through the corresponding physical downlink control channel (PDCCH) for each node. Scheduling is the process performed by the BS to assign time-frequency resources available in physical uplink shared channel (PUSCH) to user equipments. In general, scheduling is not part of the standardization work, and is left for vendor implementation. However, signaling is standardized, and hence, any scheduling scheme should comply with the control requirements in the standards. The access reservation, scheduling, and scheduled data transmission procedures in existing cellular networks have been designed and optimized for a totally different traffic pattern, i.e. human-oriented communications. Then, for battery limited MTC devices deployed in cellular networks, these two procedures are the bottlenecks. Regarding the access reservation procedure, the following research questions must be answered:

- RQ8: How does energy consumption of MTC devices in random access scale with the number of devices served per BS?
- RQ9: How can one enhance performance of the random access procedure with introducing as low as possible complexity and cost to devices and the access network?

Regarding the scheduling procedure, a thorough survey on scheduling algorithms for MTC traffic indicates that the existing MTC scheduling algorithms are mainly focused on the time-domain scheduling, and vary in defining the priority metric used to select the device to be scheduled [40]. Regarding the scheduling metric, the existing algorithms could be categorized into 4 main categories as follows: (i) channel-based schedulers; (ii) delay-based schedulers; (iii) fairness-based schedulers; and (iv) hybrid schedulers, which consider a combination of the aforementioned metrics as well as other metrics like power consumption, buffer status, and data arrival rates. While providing scalable yet energy efficient communications is considered as the key requirement for successful deployment of MTC over existing cellular networks [24, 47], one sees that energy efficient uplink MTC scheduling has been almost neglected in previous research works. Then, concerning the scheduling procedure and scheduled data transmission, the following research questions must be addressed:

- RQ10: Is battery lifetime extension achievable by lifetime-aware scheduling? What is the theoretical model describing the coupling between battery lifetime and scheduling?
- RQ11: How can we design a network lifetime-aware scheduler suitable for serving massive MTC devices?

1.3.3 Performance Tradeoff Analysis

To realize a long lasting cellular-based M2M network, different aspects of cellular networks must be optimized for machine-type communications. The telecommunications industry has invested a great deal of efforts and money investigating how to build high-capacity, in terms of traffic volume, high-throughput low-latency mobile radio systems but forgotten about low-power high-capacity, in terms of number of connected devices, systems- in which energy management is of paramount importance. Now, we need to look back at the performance tradeoffs to see how the implemented solutions, which have been designed to optimize QoS of human-oriented communications or reduce energy consumptions of the BSs, can affect the

QoS of machine-type communications, and vice versa. For example, one need to investigate the potential impact of energy saving for the access network, i.e. green cellular network design, on the battery lifetime of machine devices. In existing cellular networks, the user equipments are charged in a daily manner, and hence, reduction in the battery lifetime of a smart-phone has been feasible at the cost of energy saving for the BS. However, with the introduction of machine-type subscribers in cellular networks, this tradeoff is to be investigated as for most devices, their batteries will not be charged to the end of their lifetime which is expected to be over 10 years [47]. To the best of our knowledge, the energy/lifetime/delay analysis of a system which serves mixed machine- and human-oriented traffics is absent in literature. Towards this end, the following research questions are addressed in this dissertation:

- RQ12: What are the tradeoffs between green and lifetime-aware cellular network design in deployment and operation phases?
- RQ13: What are the consequences of allocating radio resources to massive MTC services on energy consumption of the BSs, network spectral efficiency, experienced delay of non-MTC traffic, and battery lifetime of machine-type devices?

1.4 Research Methodologies

This section presents a brief explanation of the research approaches followed in this dissertation in order to answer the above mentioned research questions.

RQ1 and RQ2 deal with definition and identification of battery lifetime metrics. In order to answer these questions, we first performed a qualitative literature review. The lifetime issue for MTC is similar to that in wireless sensor networks (WSN). Then, we explored state-of-the-art energy consumption evaluation in medium access control and clustering design for wireless sensor networks, as well as cellular networks. The obtained knowledge was employed subsequently in deriving realistic energy consumption and battery lifetime models for MTC considering different sources of energy consumptions.

RQ3-RQ7 deal with designing a clean state MAC protocol for cluster-based MTC in order to maximize the network battery lifetime. First, the optimal cluster size is required. To address this question, we derive the corresponding network battery lifetime model for cluster-based MTC networks in which, density of data gathering nodes plays a crucial role in determining the level of consumed energy per gathered bit of data data, and hence, battery lifetime. Using this model, we are able to analytically derive the average cluster size which maximizes network battery lifetime. Furthermore, we extend our model to incorporate the impact of CH selection on the network battery lifetime, and hence, derive the CH selection scheme which maximize network battery lifetime (RQ4). For RQ5, we follow the same trend to find the lifetime-aware solutions. RQ6 deals with designing communications protocols inside and outside clusters. To address this question, we follow a quantitative experimental methodology approach to investigate the system's performance using link level simulations and to establish cause and effect relationships. Then, we follow a deductive approach in order to draw conclusions, and propose a novel MAC protocol. RQ7 deals with the interference from from underlying MTC on the primary communications. To answer this question, we use statistical tools, and analytically model the power of received interferences from random points inside the cell to the BS, and derive how it affects outage performance of a primary user in uplink communications.

In RQ8 and RQ9, we seek for performance evaluation and improvement of the access reservation procedure of existing cellular networks. Towards this end, we analytically model the key performance indicators (KPIs) of interest, including access rate and energy consumption. With the help of the derived closed form expressions of the objective functions, we can analytically dimension random access resources for any traffic load regime. Furthermore, we validate our derived analytical models by comparing them against system level simulation results.

In RQ10 and RQ11, we seek for battery lifetime-aware scheduling and transmit power control solutions. Towards this end, we first analytically investigate impact of scheduling on individual and network battery lifetime. Then, we formulate an optimization problem to maximize the desired network battery lifetime metric subject to limited radio resources available at the BS and limited energy resources available at the UEs. Furthermore, we use convex optimization theory to implement low-complexity search algorithms that can efficiently solve the optimization problem, and find lifetime maximizing scheduling solutions.

Furthermore, RQ12 and RQ13 aim at figuring out performance tradeoffs that serving MTC over cellular networks introduces to the system. To address RQ12, we need to analytically model energy consumption of the access network, experienced delay by users, and battery lifetime of machine devices. Following an inductive approach, we figured out a retrial queuing model that its results match well with the simulations results in performance evaluation of a green cellular network serving MTC/HoC coexistence traffic. This model is used subsequently in deriving closed-form expressions for batter lifetime, BS energy consumption, and experienced delay by HoC traffic. based on this analytical framework, one can figure out the energy-delay, energy-lifetime, and energy-delay tradeoffs. To answer RQ 13, we introduce a Markov chain that models uplink transmission of a single cell with coexistence of MTC/HoC traffic. Investigating this Markov chain, we can derive closed form expressions for KPIs of interest, and figure out the existing tradeoffs among them. Finally, we use system level simulations for data analysis and validation, and the deductive approach for drawing conclusions. The parameter sets utilized in the simulations are mostly based on the 3GPP and METIS specifications.

The repeatability of the results are ensured by the detailed explanation of the simulation approaches and parameters in the thesis and appended papers.

1.5 Contributions

The main contributions of this dissertation are listed in the following subsections.

Battery Lifetime Analysis

In answering to RQ1-RQ2, we introduce accurate energy consumption, individual, and network lifetime models for machine-type devices deployed in cellular networks by taking all sources of energy consumptions into account. This problem is also studied in the following papers:

- **Paper 1 (Paper A in the appendix):** G. Miao; A. Azari; T. Hwang, "E^2-MAC: Energy Efficient Medium Access For Massive M2M Communications," in IEEE Transactions on Communications, vol. x, no.99, pp. x-x.

- **Paper 2:** A. Azari and G. Miao, "Lifetime-aware scheduling and power control for cellular-based M2M communications," 2015 IEEE Wireless Communications and Networking Conference (WCNC), New Orleans, LA, 2015, pp. 1171-1176.

The author of this thesis acknowledges X. Chen and P. Zhang's efforts on investigation of feasibility of the some parts of paper 1 under supervision of Professor Miao. The author of this thesis has contributed in writing paper 1 from the draft conference versions prepared by the first author of paper 1, extending the ideas, developing analytical models, and deriving numerical and simulation results. The author of this thesis acknowledges valuable editing efforts done by the first author.

Battery Lifetime-Aware Solutions

In answering to RQ3-RQ7, we explore the impact of clustering on network lifetime and find the cluster size to maximize network lifetime. Also, we present a distributed cluster-head (re-)selection scheme, and explore the feasibility of clustering from network-lifetime perspective as functions of system, traffic, and resource allocations parameters. Furthermore, we propose a load-adaptive multiple access scheme, called n-phase CSMA/CA, which provides a tunable tradeoff between energy efficiency and delay by choosing n properly. RQ3-RQ6 are studied in paper 1, and the following paper:

- **Paper 3:** A. Azari and G. Miao, "Energy efficient MAC for cellular-based M2M communications," Signal and Information Processing (GlobalSIP), 2014 IEEE Global Conference on, Atlanta, GA, 2014, pp. 128-132.

The author of this thesis acknowledges X. Chen and P. Zhang's efforts on investigation of feasibility of paper 3 under supervision of Professor Miao. RQ7 is studied in paper 8, to be discussed later.

In answering to RQ8-RQ9, we introduce a novel statistical model for coexistence of MTC and HoC traffic over cellular networks, and investigate access rate and energy efficiency performance of existing random access procedures of cellular networks in dealing with the coexistence traffic. Random access resource provisioning for MTC traffic achieving a certain level of QoS is carried out using the presented analytical model, and validated using the system level simulations. These problems are also studied in the following papers:

- **Paper 4:** Amin Azari, Mohammad Istiak Hossain, and Jan I Markendahl, "RACH Dimensioning for Reliable MTC over Cellular Networks," Submitted to 2017 IEEE VTC.

- **Paper 5:** Mohammad Istiak Hossein, Amin Azari, and Jens Zander, "DERA: Augmented Random Access for Cellular Networks with Dense H2H-MTC Mixed Traffic," 2016 IEEE Globecom Workshops, Washington DC, 2016, pp. 1-5.

- **Paper 6:** Mohammad Istiak Hossain, Amin Azari, Jan Markendahl, and Jens Zander, "Enhanced Random Access: Initial Access Load Balance in Highly Dense LTE-A Networks for Multiservice (H2H-MTC) Traffic," Submitted to 2017 IEEE ICC.

In paper 5 and 6, the author of this thesis has contributed in analytical analysis and proofreading.

In answering to RQ10-RQ11, we present a battery lifetime-aware MTC scheduling framework, and explore uplink MTC scheduling based on the Max-Min lifetime-fairness, and analyze its contribution in reducing maintenance costs of machine devices. Also, uplink scheduling solutions for machine-type communications over single-carrier frequency division multiple access (SC-FDMA[1]) systems is investigated, and *low-complexity* scheduling solutions with limited feedback requirement are proposed. These problems are also studied in paper 2, and the following papers:

- **Paper 7 (Paper B in the appendix):** A. Azari and G. Miao, "Network Lifetime Maximization for Cellular-Based M2M Networks," to IEEE Transactions on Wireless Communications.

- **Paper 8:** Amin Azari; "Energy-efficient scheduling and grouping for machine-type communications over cellular networks," in Ad Hoc Networks, vol. 43, June 2016, pp 16-29.

- **Paper 9:** A. Azari and G. Miao, "Lifetime-Aware Scheduling and Power Control for M2M Communications in LTE Networks," 2015 IEEE 81st Vehicular Technology Conference (VTC Spring), Glasgow, 2015, pp. 1-5.

[1]SC-FDMA is used in LTE, LTE-A, and LTE-M networks for uplink transmission [52].

Analysis of the Performance Tradeoffs

In answering to RQ12, we develop a tractable framework to model the operation of a green BS which serves mixed MTC and HoC traffic, and saves energy by going to the sleep mode. We derive closed-form expressions for energy consumption of the BS, experienced delay by users and machines, and expected battery lifetime of machine devices. Then, we introduce the fundamental tradeoffs, and explore the impact of system and traffic parameters on the introduced tradeoffs. These problems are also studied in the following paper:

- **Paper 10 (Paper C in the appendix):** A. Azari and G. Miao, "Battery Lifetime-Aware Base Station Sleeping Control with M2M/H2H Coexistence," 2016 IEEE Globecom, Washington DC, 2016, pp. 1-5.

Finally, in answering to RQ13 we investigate energy efficiency, spectral efficiency, and network battery lifetime tradeoffs in resource provisioning for MTC services over cellular networks. Using these tradeoffs, we figure out the ways in which spectral/energy efficiency for the access network and QoS for high-priority services could be traded to prolong battery lifetimes of connected devices by compromising on the level of provisioned radio resources. These problems are also studied in the following paper:

- **Paper 11 (Paper D in the appendix):** A. Azari and G. Miao, "Fundamental Tradeoffs in Resource Provisioning for IoT Services over Cellular Networks," Submitted to 2017 IEEE ICC.

1.6 Outline

The rest of this thesis is organized as follows. In the next chapter, we present the system model, and the theoretical models and the research methodologies employed to answer research questions throughout the thesis. The key results achieved in answer to the research questions are presented in chapter III. Concluding remarks and future research directions are given in chapter IV. Finally, the included papers are appended to the end of dissertation.

Chapter 2

Modeling and Methodology

In the following, we first present the system model which is used throughout the thesis. Then, further details about theoretical models and the methodologies followed in this thesis to derive lifetime-aware solutions are presented[1].

2.1 System Model

Consider a single cell with one base station at the center and a massive number of static machine-type devices which are randomly distributed according to a spatial Poisson point process of intensity υ. Machine nodes are battery driven and long battery lifetime is crucial for them. Packet generation at each node is considered as a Poisson process, which is suitable for data gathering and reporting MTC applications like smart metering [53].

2.2 Network Battery Lifetime Modeling Approaches

Throughout this dissertation, we are seeking for network battery lifetime-maximizing solutions. Then, it is important to describe the way in which we have modeled the network battery lifetime in each problem. Towards this end, a general lifetime model is presented in the following subsection, which is followed by respective lifetime models for specific research questions.

2.2.1 The General Model for Network Battery Lifetime

In the following, the general model of network battery lifetime, which is used in answering research question throughout the dissertation, is presented.

[1] Parts of material presented in this chapter are based on our work in appended paper A (©2016 IEEE), paper B, paper C (©2016 IEEE), and paper D, which have been published or submitted to IEEE. Material is reused with permission.

2.2.1.1 Individual Battery Lifetime Modeling

As illustrated in Fig. 2.1, a typical machine node may have different energy consumption levels in different activity modes, including data gathering, processing, synchronization, transmission, and sleeping. For most reporting MTC applications, the packet generation process at each machine device can be modeled as a Poisson process [53]. Then, energy consumption of each device can be seen as a semi-regenerative process where the regeneration point is at the end of each successful data transmission epoch. Let us denote the remaining energy of the ith device at time t_0 as $E_i(t_0)$, the average time between two data transmissions as T_i, and the average packet size as D_i. Also, power consumption of node i in the sleeping and transmission modes are denoted as P_s and $P_i + P_c$ respectively, in which P_c is the circuit power consumed by electronic circuits in the data transmission mode and P_i is the transmit power for reliable data transmission. Now, we define the expected lifetime for node i at the regeneration point as the product of reporting period and the ratio between remaining energy and the average energy consumption per reporting period, as follows:

$$L_i(t_0) = \frac{E_i(t_0)T_i}{E_s + P_s[T_i - \frac{D_i}{R_i} - n_a^i T_a^i] + n_a^i T_a^i P_a + \frac{D_i}{R_i}[P_c + \eta P_i]}, \tag{2.1}$$

where R_i is the average expected data rate for node i, η is the inverse of power amplifier efficiency, and E_s is the average static energy consumption in each reporting period for synchronization, admission control, and etc. Also, P_a is the power consumption in the active mode for data gathering, T_a^i is the active mode duration, and n_a^i is the number of active modes per reporting period. Let us denote the average energy consumption in transmission-independent and -dependent modes by $\mathcal{E}_s^i$ and $\mathcal{E}_d^i$ respectively. Then, the lifetime expression in (2.1) could be rewritten as:

$$L_i(t_0) = \frac{E_i(t_0)}{\mathcal{E}_s^i + \mathcal{E}_d^i} T_i, \tag{2.2}$$

where

$$\mathcal{E}_s^i = E_s + P_s[T_i - n_a^i T_a^i - \frac{D_i}{R_i}] + n_a^i P_a T_a^i, \tag{2.3}$$

$$\mathcal{E}_d^i = \frac{D_i}{R_i}[P_c + \eta P_i].$$

2.2.1.2 Coupling between Energy Efficiency and Battery Lifetime

Energy efficiency of node i in the transmission mode is defined as [54]:

$$U_i(R_i) = \frac{R_i}{P_c + \eta P_i(R_i)}. \tag{2.4}$$

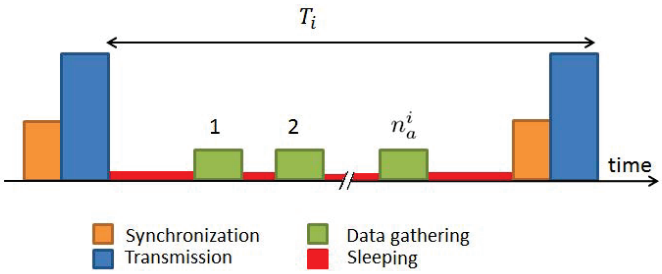


Figure 2.1: Power consumption profile for node i. Different modes consume different power levels.

It is shown that in case[2] $P_i(R_i)$ is strictly convex in R_i, $U_i(R_i)$ is strictly quasi-concave and we can find the optimal R_i which maximizes $U_i(R_i)$, i.e. the energy efficiency [54]. Now, we can rewrite the battery lifetime expression in (2.2) as a function of U_i, as follows:

$$L_i(t_0) = \frac{E_i(t_0)T_i}{\mathcal{E}_s^i + D_i \frac{P_c+\eta P_i}{R_i}} = \frac{E_i(t_0)T_i}{D_i} \frac{R_i}{P_c + \eta[P_i + \mathcal{E}_s^i \frac{R_i}{D_i \eta}]}. \tag{2.5}$$

Let us define $\tilde{P}_i(R_i)$ as $P_i(R_i)+\frac{\mathcal{E}_s^i}{\eta D_i}R_i$ which is again strictly convex in R_i if $P_i(R_i)$ is strictly convex. Then, we can rewrite (2.5) as:

$$L_i(t_0) = \frac{E_i(t_0)T_i}{D_i} \frac{R_i}{P_c + \tilde{P}_i} = \frac{E_i(t_0)T_i}{D_i} \tilde{U}_i(R_i). \tag{2.6}$$

The lifetime expression in (2.6) shows that battery lifetime is proportional to the energy efficiency $\tilde{U}_i(R_i)$. Then, maximizing the lifetime is equivalent to maximizing energy efficiency. The link-level energy efficiency can be maximized using the techniques in [54]. In this dissertation we mainly focus on the network-level energy efficiency.

2.2.1.3 Network Battery Lifetime as a Function of Individual Battery Lifetimes

Network lifetime is the time span between network deployment and when it is considered to be nonfunctional. The instant at which a MTC network is considered nonfunctional is application-specific. For example, in safety-critical MTC applications where loosing even one machine node degrades applications' performance or coverage, the shortest individual lifetime (SIL) may specify the network lifetime, i.e.

$$L_{\text{net}}^{\text{sil}}(t_0) = \min_i L_i(t_0).$$

[2] $F(x)$ represents F as a function of x.

SIL is also the case in sparse MTC deployment scenarios where the correlation between gathered data by different nodes is low. In other cases, e.g. where the correlation between gathered data by different nodes is high, the longest individual lifetime (LIL), or the average individual lifetime (AIL) may specify the network lifetime.

2.2.2 Network Battery Lifetime Modeling in Deriving Lifetime-Aware Solutions for Clustered MTC

In order to investigate if machine devices can benefit from clustering for prolonging battery lifetime, and develop a clustering scheme to improve the overall network lifetime, here we investigate our network battery lifetime model for clustered MTC. To keep the analysis tractable and obtain closed-form expressions, we consider a homogeneous MTC deployment in which machine nodes have similar communications characteristics, i.e. packet lengths and packet generation frequencies. Also, we consider the cluster-forming problem at the reference time where $E_i(t_0) = E_0, \ \forall i$, and E_0 is the full battery capacity. In order to achieve the highest SIL lifetime in each cluster, i.e. to maximize the time at which first energy drain happens in each cluster, machine nodes need to be in the CH mode in turns. This will avoid that a single node has its battery drained. Denote the average cluster size in the network as z. The transmit power levels in the CM and CH modes are denoted by P_t^m and P_t^h respectively. In each reporting period of each cluster activity, a node may be in the CH mode with probability $\frac{1}{z}$ and in the CM mode with probability $1 - \frac{1}{z}$. Then, the expected battery lifetime of each node in a cluster which is centered at distance d_h from the BS is expressed as the product of the cluster reporting period and the ratio between the remaining energy and the average energy consumption in each reporting period, as follows:

$$L_c(d_h, z) = \frac{E_0}{\frac{1}{z}\mathcal{E}_h + (1 - \frac{1}{z})\mathcal{E}_m} T_c. \tag{2.7}$$

The energy consumptions of each node in the CM and CH modes are written as:

$$\begin{aligned} \mathcal{E}_m &= E_s + \tilde{D}\frac{P_c + \eta P_t^m}{R_m}, \qquad (2.8)\\ \mathcal{E}_h &= E_s^h + \frac{(z-1)\tilde{D}}{R_m} P_l + [1 + \lambda_{comp}(z-1)]\tilde{D}\frac{P_c + \eta P_t^h}{R_h}, \end{aligned}$$

respectively. In this expression, λ_{comp} is the packet-length compression coefficient at the CHs, $\tilde{D}$ the average packet size, T_c the cluster reporting period, P_l the power consumption in the listening mode, $\frac{(z-1)\tilde{D}}{R_m} P_l$ models energy consumption of CH in receiving packets from the CMs, and E_s^h is the average static energy consumption in the CH mode. Also, given multiple access scheme for intra- and inter-cluster communications, R_m and R_h are derived, as described in appended paper A. With the help of the derived closed form of the objective function, we can analytically derive the lifetime maximizing cluster size, CH selection scheme, and communications protocol within the clusters. The key results indicating battery lifetime improvement from lifetime-aware MAC and clustering design are presented in appended paper A, and are summarized in chapter 3.

2.2.3 Network Battery Lifetime Modeling in Deriving Lifetime-Aware Solutions for Direct MTC

For optimizing direct communications of CHs and unclustered MTC devices with the BS, we need to derive a network lifetime model as a function of scheduled data transmission parameters. Let us focus on the air interface of 3GPP LTE Release 13 [55]. In this standard, radio resources are distributed in time and frequency domains. In the time domain, data transmissions are structured in frames where each frame consists of 10 subframes each with 1 ms length. In the frequency domain, the available bandwidth is divided into a number of subcarriers each with 15 KHz bandwidth. The minimum resource element in a frame that can be allocated to a user is a physical resource block pair (PRBP) which consists of 12 subcarriers spanning over one transmission time interval (TTI) [55]. Each TTI consists of two slots in the time domain and includes 12 (or 14) OFDM symbols if long (or short) cyclic prefix is utilized. Based on the LTE open-loop uplink power control mechanism [55], the uplink transmit power of each node is derived using the downlink pathloss estimation as:

$$PowC(c_i, \delta_i) = c_i P_0 \beta_i \theta_i [2^{\frac{1.25\text{TBS }(c_i,\ \delta_i)}{c_i N_s N_{sc}}} - 1], \tag{2.9}$$

in which, the number of assigned PRBPs to node i is denoted by c_i, the compensation factor by β_i, the estimated downlink pathloss by θ_i, the number of symbols in a PRBP by N_s, and the number of subcarriers in a PRBP by N_{sc}. Furthermore, the transport block size (TBS) can be found in Table 7.1.7.2.1-1 of [55] as a function of c_i and TBS index. The TBS index, denoted by $\delta_i \in \{0, \cdots, 33\}$, is a function of modulation and coding scheme (MCS) as depicted in Table 8.6.1-1 of [55]. Also, P_0 is a user specific value, which is set based on the required SNR level at the receiver as:

$$P_0 = \beta_i [\text{SNR}_{\text{target}} + P_n] + [1 - \beta_i] P_{max},$$

in which P_{max} is the maximum allowed transmit power, and $P_n = -209.26$ dB is the noise power in each resource block. Based on these definitions, the expected battery lifetime for node i is found as:

$$L_i(t) = \frac{E_i(t) T_i}{E_s^i + TTI[P_c + \eta PowC(c_i, \delta_i)]}. \tag{2.10}$$

Now, one can formulate the uplink scheduling and transmit power control problem as:

$$
\begin{aligned}
& \max_{c_i,\delta_i} \; L_{\text{net}}^{sil} && (2.11)\\
\text{subject to: } & \text{C.1: } \sum\nolimits_{i\in\mathcal{A}} c_i \le |C|, \\
& \text{C.2: } \bar{D}_i \le \text{TBS}(c_i,\delta_i), \quad \forall i\in\mathcal{A}, \\
& \text{C.3: } PowC(c_i,\delta_i) \le P_{max}, \quad \forall i\in\mathcal{A}, \\
& \text{C.4: } \delta_i \in \{0,\cdots,33\}; c_i \in \{1,\cdots,|C|\}, \quad \forall i\in\mathcal{A},
\end{aligned}
$$

where $\mathcal{A}$ is the set of devices to be scheduled, L_{net}^{sil} is given as a function of individual lifetimes in section 2.2.1.3, $\bar{D}_i = D_i + D_{oh}$, and D_{oh} is the size of overhead information for User Datagram Protocol (UDP), Internet Protocol (IP), and etc. Also, $|C|$ is the total number of available PRBPs, e.g. 110 in the case of LTE system with 20 MHz bandwidth. Low-complexity solutions for solving this problem have been proposed in the appended paper B. The key results indicating battery lifetime improvement from lifetime-aware uplink scheduling and transmit power control are presented in the appended paper B, and are summarized in chapter 3.

2.2.4 Network Battery Lifetime Modeling in Analysis of the Performance Tradeoffs

Serving MTC services over cellular networks introduces new performance tradeoffs which have not been met before in dealing with HoC traffic. This is due to the fact that QoS requirements and communications characteristics of MTC services are way different from HoC traffic. Let us investigate the potential impact of energy saving for the access network, i.e. green cellular network design, on the battery lifetime of machine devices. BS sleeping is a well-known technique for adapting to the traffic load, and saving energy for the access network. To investigate the potential impacts of BS sleeping on battery lifetime of machine devices we need to find a battery lifetime model as a function of BS sleeping parameters. In the sequel, we present a summary of methodology and theoretical model used in the appended paper C for solving the problem. The same methodology is used for battery lifetime modeling in resource provisioning for MTC services, which we skip it here. Consider a single cell with one base station at the center, and a massive number of user equipments (UEs) which are uniformly distributed in the cell. We assume that the links between UEs and the BS experience Rayleigh fading. The UEs are assumed to be heterogeneous in nature such that can be mainly categorized into two different categories: (i) the high-priority devices, called $\mathcal{P}_1$; and (ii) the low-priority devices, called $\mathcal{P}_2$, which are usually battery-limited machine-type devices with strict constraints on transmit power, and battery lifetime; however, they can tolerate delay in data transmission to some extent. The arrival of $\left\{\mathcal{P}_n\right\}_{n=1,2}$ devices is modeled as a Poisson process with rate of λ_n. Also, we assume that $\mathcal{P}_1$ devices have non-preemptive priority over $\mathcal{P}_2$ devices. After serving UEs which have been queued to be served, BS waits for new requests, where the listening window is exponentially distributed with rate μ. If no UE arrives in the listening window, BS goes to the sleep mode, where the sleep window follows a general distribution with cumulative distribution function (CDF), probability distribution function (PDF), Laplace-Stieltjes transform, and first and second moments as $V(x)$, $v(x)$, $v^*(s)$, $\bar{v}$, and $\hat{v}$, respectively. After sleeping window expiration, the BS wakes up and starts listening for serving potential arriving UEs. The number of connected $\mathcal{P}_2$ devices in future cellular networks is expected to become much higher than the existing user equipments [8]. Then, in order to prevent network congestion due to the huge number of $\mathcal{P}_2$ devices, we utilize an access class barring (ACB) scheme [26]. Based on this ACB scheme, when BS is asleep or busy, $\mathcal{P}_2$ devices will retry after a random backoff time. This backoff time is exponentially distributed with rate α. When BS is busy, $\mathcal{P}_1$ devices are queued to be served based on a processor sharing service discipline. When the BS is asleep, $\mathcal{P}_1$ devices keep listening to find the BS available. Let us denote noise power spectral density (PSD) at the receiver as N_0, and the channel gain between node i and the BS as $g_i = hr_i^{-\sigma}$ where $h \sim \exp(1)$, r_i is the communication distance, and σ is the pathloss exponent. Furthermore, in order to guarantee a constant received signal to interference and noise ratio (SINR), i.e. γ_n, we use the Shannon capacity formula and derive the required transmit power of node i as:

$$P_{t_{(i,n)}} = \gamma_n \Phi \gamma_0 r_i^{\sigma} / [\mathcal{G}h], \tag{2.12}$$

where γ_0 is the SINR gap between channel capacity and a practical coding and modulation scheme, $\mathcal{G}$ the product of transmit and receive antenna gains, $\Phi = (I + N_0)B$, I the upperbound on the PSD of out-of-cell interference, B the bandwidth, and $n \in \{1, 2\}$ denotes the priority category of the device. Regarding the uniform distribution of UEs in the cell, the PDF of the distance between a UE and the BS is written as $f(r) = \frac{2r}{R_c^2}$, where R_c is the cell radius and r is the communications distance. Then, we can derive the long-term average required transmit power for type n devices as:

$$\bar{P}_{t_n} = \int_0^{R_c} \frac{\gamma_n \Phi \gamma_0 r^{\sigma}}{\mathcal{G}} \frac{2r}{R_c^2} dr = \frac{2(\sigma+1) R_c^{\sigma-2} \gamma_n \Phi \gamma_0}{\mathcal{G}}. \tag{2.13}$$

Also, we can derive the unsuccessful transmission probability for node i as:

$$q_{(i,n)} = pr(P_{t_{(i,n)}} > P_{max}) \overset{(a)}{=} 1 - \exp\big(-\frac{r_i^{\sigma} \gamma_0 \Phi \gamma_n}{\mathcal{G} P_{max}}\big), \tag{2.14}$$

in which (a) is due to the fact that h is exponentially distributed. Then, regarding the uniform distribution of UEs in the cell, the average unsuccessful transmission probability can be derived as:

$$\bar{q}_n = \int_0^{R_c} \left[1 - \exp(-\frac{r^{\sigma} \gamma_0 \Phi \gamma_n}{\mathcal{G} P_{max}})\right] \frac{2r}{R_c^2} dr. \tag{2.15}$$

In order to derive the exact expression of $\bar{q}_n$, the integral tables in [56] can be used. For example, in case $\sigma = 4$ we have:

$$\bar{q}_n = 1 - \sqrt{\pi}\text{Erfc}(AR_c^2)/2AR_c^2,$$

in which $A = \sqrt{\frac{\gamma_0 \Phi \gamma_n}{\mathcal{G} P_{max}}}$, and Erfc$(\cdot)$ is the error function [56]. We assume that the uplink service requirement of each type n device is exponentially distributed with average length of $\bar{d}_n$ bits. Then, the average service time for a successful transmission is derived as:

$$u_n = \bar{d}_n [B \log_2(1 + \gamma_n)]^{-1}, \quad n \in \{1,2\}. \tag{2.16}$$

By considering the retransmission policy, the average service time can be rewritten as:

$$\bar{u}_n = \sum_{k=1}^{\infty} \bar{q}_n^{[k-1]} [1 - \bar{q}_n] k \; u_n = \frac{1}{1 - \bar{q}_n} \frac{\bar{d}_n}{B \log_2(1 + \gamma_n)}. \tag{2.17}$$

Let us denote the CDF of the service time by U_n. By considering the retransmissions, the service time can be well approximated by an exponential distribution with mean of $\bar{u}_n$. Denote by ξ the state of the BS, where $\xi \in \{0, 1, 2, 3\}$ refer to the sleep, busy with a $\mathcal{P}_1$ device, busy with a $\mathcal{P}_2$ device, and listen states, respectively. Using theorem 5.2 in [57], one can derive the stability condition for this queuing system as $\rho = \rho_1 + \rho_2 < 1$, where ρ_n is defined as $\bar{u}_n \lambda_n$. The detailed steady state analysis of the system is given in the appended paper C. The average energy consumption from data generation to successful data transmission for machine devices can be modeled as:

$$E_n^{ac} + \left[P_c + \eta \bar{P}_{t_n}\right] \bar{u}_n, \tag{2.18}$$

where P_c is the average circuit power consumption, and η is the inverse of power amplifier efficiency. Also, E_n^{ac} is the average energy consumption from data generation until successful transmission, and is formulated as $P_c W_n \alpha \tau$, where τ is the average spent time in each trial, and W_n is the average delay from data generation until successful transmission, derived from a retrial queuing model with server breakdown as a function of BS sleeping parameters in the appended paper C. Then, the expected battery lifetime for a $\mathcal{P}_n$ device is found as:

$$\mathcal{L}_n = \frac{E_0}{E_s + E_n^{ac} + \left[P_c + \eta \bar{P}_{t_n}\right] \bar{u}_n} T, \tag{2.19}$$

where T is the reporting period, and E_s is the average static energy consumption in each reporting period for data gathering, synchronization, and etc. Using this closed-form expression for battery lifetime as a function of traffic, uplink radio resources, and BS sleeping parameters, we are able to analytically investigate the how BS sleeping can affect battery lifetime of machine-type devices. The key results indicating the performance tradeoffs in serving MTC services are presented in paper C and D, and are summarized in chapter 3.

Chapter 3

Key Results

In this chapter, we present the key results obtained in this thesis[1].

3.1 Network Battery Lifetime Analysis

In paper A and B, reprinted in the appendix, accurate energy consumption models for machine-type communications are derived by taking all sources of energy consumptions into account. These models are employed subsequently in order to derive individual and network battery lifetime models. We have used the network battery lifetime metric as an objective function in a variety of network design problems, e.g. designing access reservation and scheduled data transmission procedures, in order to derive lifetime-aware solutions. Our analytical and simulations results, to be presented in the following sections, indicate that incorporating battery lifetime-awareness to the design of future cellular networks enables them to serve long-lasting massive MTC services.

3.2 Lifetime-Aware MAC Design for Clustered MTC

In paper A, reprinted in the appendix, we propose a clean state MAC for clustered MTC over cellular networks. Towards this end, given MTC traffic and characteristics, we investigate the cluster-size and CH selection scheme which maximize the network lifetime. Furthermore, we investigate the choice of communications protocols within the clusters. In clustered MTC, communications consist of two phases: (i) intra-cluster communications from CMs to CHs and (ii) inter-cluster communications from CHs and non-clustered nodes to the BS. The two phases use orthogonal resources. Fig. 3.1 illustrates a potential frame structure for lifetime-aware data gathering from MTC devices over LTE systems. In the first phase, all

[1]Parts of material presented in this chapter are based on our work in appended paper A (©2016 IEEE), paper B, paper C (©2016 IEEE), and paper D, which have been published or submitted to IEEE. Material is reused with permission.

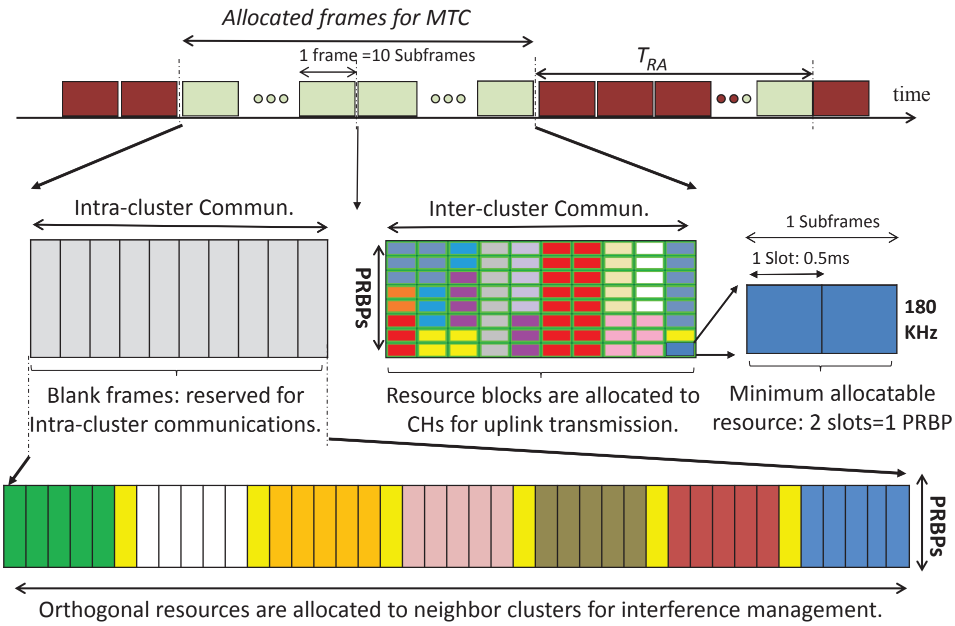


Figure 3.1: The proposed E^2-MAC for clustered MTC over LTE systems. More details on E^2-MAC can be found in appended paper A.

cluster members send data to their cluster heads. Then, the CHs will forward the data to the BS in the second phase.

While inter-cluster communications benefit from scheduled data transmission supported by BSs, for intra-cluster communications we propose to benefit from contention-based approaches, e.g. CSMA/CA technique. As energy efficiency of contention-based multiple access techniques decreases by increasing the traffic load, we present a flexible and load-adaptive multiple access protocol, called n-phase CSMA/CA, which divides each contention interval into n phases. In each phase, only a portion of the nodes can compete for data transmission. Before the assigned phase starts, machine nodes keep sleeping instead of listening and newly arrived packets are buffered.

Fig. 3.2 depicts the tradeoffs between energy efficiency, spectral efficiency, and experienced delay with different numbers of phases. By increasing n, i.e. number of phases, probability of successful transmission increases, which in turn results in higher energy efficiency due to a less number of retransmissions and shorter time spending in idle-listening mode. On the other hand, one sees that the average packet delay increases in the number of phases because of packet buffering until the assigned contention slot starts. The novel contention-division concept in the n-phase CSMA/CA can be applied in other contention-based protocols, e.g. ALOHA

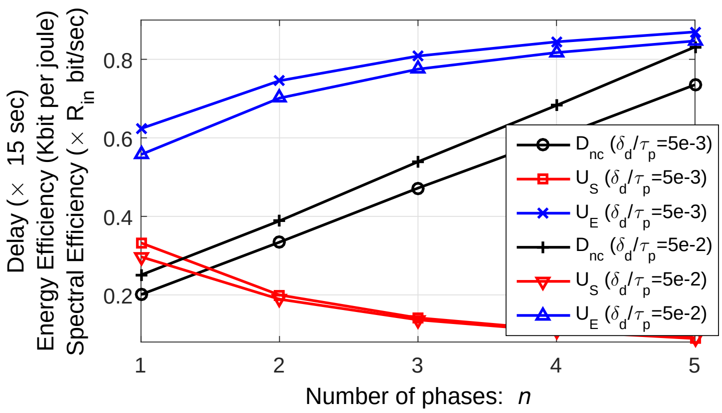


Figure 3.2: Energy efficiency, delay, and spectral efficiency for the n-phase CSMA/CA. The parameters for simulation can be found in appended paper A. (©2016 IEEE. Reused with permission.)

and 802.11, to improve their energy efficiency.

Fig. 3.3 compares lifetime performance of the E^2-MAC with the lifetime-maximizing cluster-size, i.e. $z = 100$, against the E^2-MAC with non-optimal cluster size; the pure contention-based MAC through RACH called cMAC; E^2-MACn in which there is no CH reselection; and analytical results derived from the developed model for battery lifetime of clustered MTC. In this figure, $(x,y)E^2$-MAC refers to the E^2-MAC where x is the number of phases for the n-phase CSMA/CA and y is the average cluster size.

From Fig. 3.3, one sees that using cMAC, a great number of nodes die very early because of energy wastage in collisions and idle listening, while the remaining nodes last for a longer time because of reduced contention in channel access. Furthermore, we see that using the E^2-MACn, the respective CDF curve has a mild slope because the batteries of the first set of CHs drain out of energy very soon and the ones of the last set of CHs last for a long time. Also, using $(1,100)E^2$-MAC, where 100 is the lifetime maximizing cluster-size, as derived in paper A, one sees the CDF curve has a steeper slope which means almost all machine nodes die in a limited time-window indicating replacement of their batteries can be done all at once. This in turn reduces the maintenance cost. The semi-vertical curves in this figure present the expected CDF of individual lifetimes. One sees that the derived CDF curves from the simulation results are centered on their expected lifetime from the analytical results. Furthermore, it is evident that the battery lifetime can be further improved by increasing the number of phases for the n-phase CSMA/CA, e.g. by using $(3,100)E^2$-MAC instead of $(1,100)E^2$-MAC. One sees that the SIL lifetime of the $(1,100)E^2$-MAC is 20 times higher than the cMAC because the former benefits from clustering to overcome the massive concurrent access problem, is 53 times higher

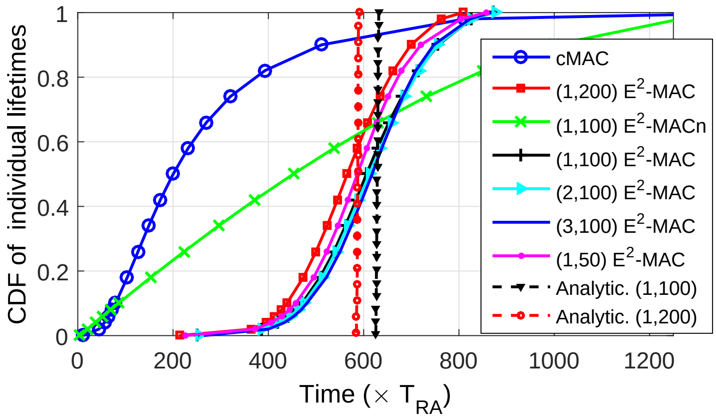


Figure 3.3: Battery lifetime analysis. The parameters for simulation can be found in appended paper A. (©2016 IEEE. Reused with permission.)

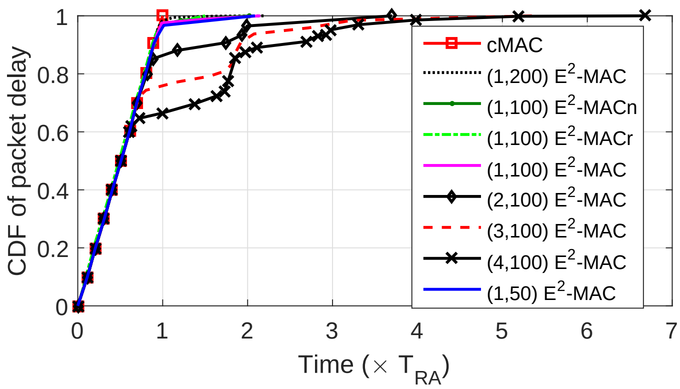


Figure 3.4: Experienced delay analysis. The parameters for simulation can be found in appended paper A. (©2016 IEEE. Reused with permission.)

than the (1,100)E^2-MACn because the (1,100)E^2-MAC distributes load of being CH among all cluster members, is 0.11 and 0.18 times higher than the (1,50) and (1,200)E^2-MAC respectively because the (1,100)E^2-MAC benefits from the lifetime maximizing cluster-size, and finally is 6% less than the (3,100)E^2-MAC because the latter benefits from the 3-phase CSMA/CA in intra-cluster communications to increase energy efficiency. Fig. 3.4 indicates that the lifetime improvement from n-phase CSMA/CA comes at the cost of introducing delay in data transmission, as we discussed in the interpretation of Fig. 3.2.

3.3 Lifetime-Aware Scheduler Design for MTC

In paper B, reprinted in the appendix, we analytically explore uplink MTC scheduling based on the Max-Min lifetime-fairness, and analyze its contribution in reducing the maintenance costs of energy-limited machine devices. Then, an uplink scheduler for MTC over single-carrier frequency division multiple access (SC-FDMA[2]) systems is presented. Also, we investigate low-complexity scheduling solutions with limited feedback requirement. We extend the proposed scheduling solutions for existing 3GPP LTE-Advanced networks in order to provide lifetime-improvement evidence in the context of LTE.

To provide lifetime improvement evidence, we apply our proposed scheduling algorithms to a 3GPP LTE-A system and present the simulation results. The deployment of machine devices and their traffic model follow the proposed models in [58, annex A] for smart metering applications, and are reflected in appended paper B.

Fig. 3.5 represents the PDF of battery lifetimes of machine devices using the following scheduling schemes: Scheme 1, in which time- and frequency-domain schedulers aim at maximizing the SIL network lifetime; Scheme 2, in which a round robin (RR) scheduler is used for time-domain scheduling, and low-complexity SIL-aware scheduler is used for frequency-domain scheduling; Scheme 3, in which aims at maximizing the LIL network lifetime; Scheme 4 which consists of two round robin schedulers for time- and frequency-domain scheduling; and Scheme 5 which consists of a channel-aware scheduler for time-domain scheduling, a round robin scheduler for frequency domain scheduling, and represents the proposed channel-aware MTC scheduling schemes in literature [41, 42, 43]. The x-axis in Fig. 3.5 has been depicted in log-scale to highlight the differences in PDFs when the initial battery drains happen, which is our concern in case of SIL network lifetime maximization. One sees that the first energy drain using scheme 1, which aim at maximizing SIL network battery lifetime, happens much later than the first energy drain using the benchmarks, i.e. scheme 4 and 5. Also, we see that the last energy drain using scheme 3, which aims at maximizing the LIL network lifetime, happens much later the benchmarks. Furthermore, we see that the PDF of scheme 1 has a compact shape, which shows that the individual lifetimes of machine devices are distributed

[2]SC-FDMA is used in LTE, LTE-A, and LTE-M networks for uplink transmission [52].

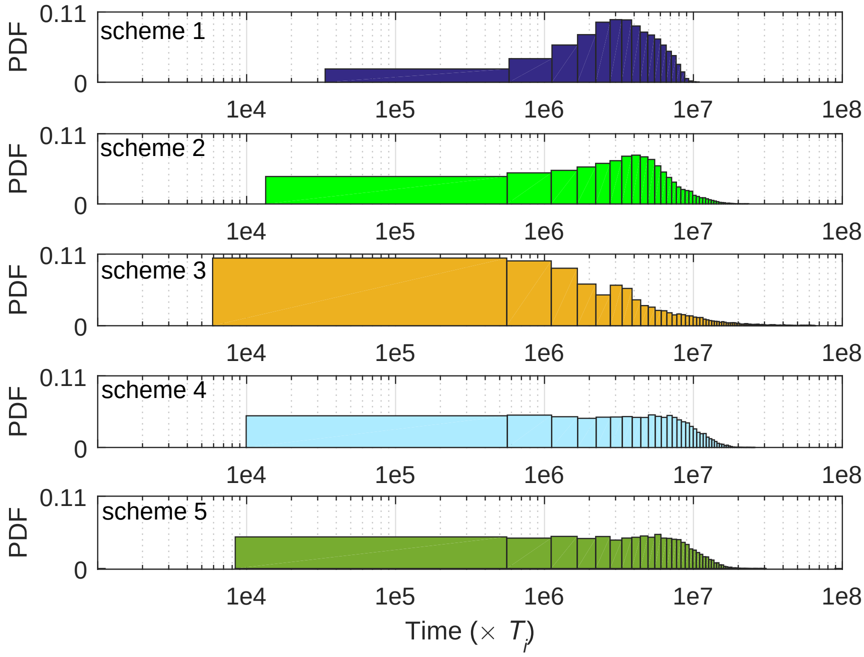


Figure 3.5: Empirical PDF of individual lifetimes using different scheduling schemes. (Reprinted from paper B, submitted for peer review to IEEE)

in a limited time interval. This in turn implies that the maintenance costs for MTC networks can be reduced using SIL-aware scheduling, as their batteries can be replaced almost at the same time.

Important insights to lifetime-aware network design can be drawn from the presented simulation results in paper B, which are summarized here. First, the size of data to be transmitted, i.e. actual data plus overhead information, greatly affects the network battery lifetimes. There are many MTC applications in which, only a few bits of data are needed to be transmitted; however, the physical layer of existing LTE networks enforces adding extra overhead information to this data. Then, the existing LTE infrastructure is not efficient for small data communications and this problem must be addressed in the next generations of cellular networks. Second, the extra required signal-to-noise ratio (SNR) at the BS can significantly shorten the network battery lifetime. Then, dense deployment of the BSs, small cells technology, and introduction of LTE category M and narrow-band LTE-M with improved link budget for the cell edge devices [59], can contribute significantly in building long lasting MTC networks.

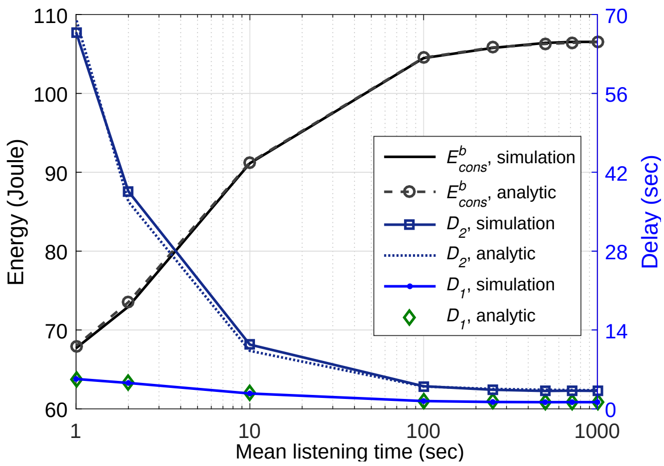
3.4 Lifetime-Aware Green Cellular Network Design

In paper C, reprinted in the appendix, we develop a tractable framework to model operation of a green BS which serves mixed MTC and HoC traffic. We derive closed-form expressions for energy consumption of the BS, delay, and battery lifetime. Then, we introduce a strong tradeoff between energy saving for the BS and battery lifetime for machine devices, and explore the impact of system and traffic parameters on the introduced tradeoff. We also extend our solutions to the multi-cell scenario, derive closed-form expressions for the energy-lifetime tradeoff, and investigate the performance impact of the control parameters.

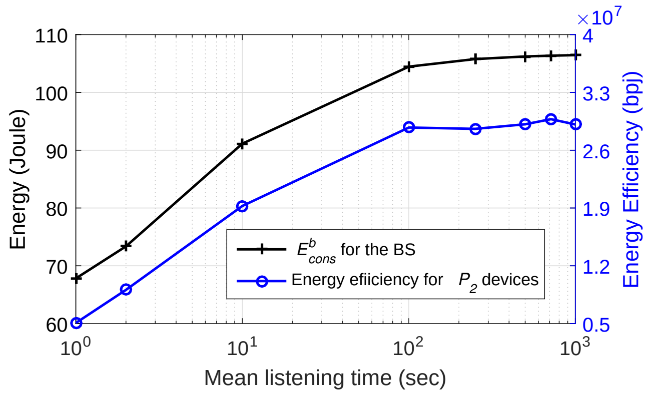
Fig. 3.6a illustrates the tradeoff between energy saving for the BS and the experienced delay in data transmission. The x-axis in this figure represents the listening time, which indicates the amount of time BS spends in idle listening before going to the sleep mode. Note that if any service request arrives during the listening time, BS serves it immediately and the listening counter is restarted. The solid lines are the numerical results, while the dashed lines have been derived from the analytical expressions derived in paper C. One sees while the operational costs of the BS increase in the listening time, the experienced delay by users decreases in the listening time. Fig. 3.6b presents the tradeoff between energy saving for the BS and energy efficiency in data transmission for machine devices. We see that both the average energy consumption of the BS and the energy efficiency of machine devices in data transmission, which in turn determines the battery lifetime, increase in the idle listening time. Fig. 3.6c illustrates the empirical CDF of individual lifetimes versus idle listening time. By comparing Fig. 3.6b with Fig. 3.6c, it is evident that a higher level of energy efficiency in data transmission provides a higher level of battery lifetime, while both of them are achieved at the cost of increasing the operational costs of the access network.

3.5 Lifetime-Aware Resource Allocation to MTC

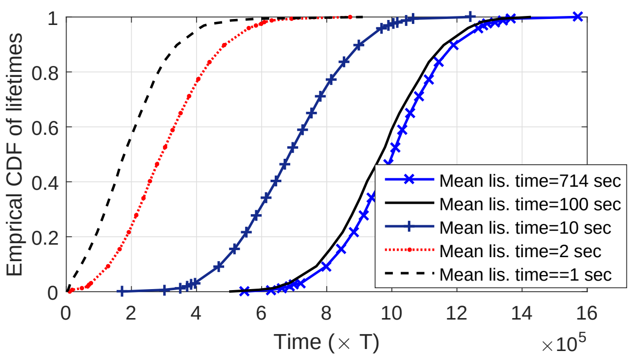
In paper D, reprinted in the appendix, we develop a tractable framework to model energy consumptions of MTC devices deployed over cellular networks, experienced delay and spectral efficiency of MTC/HoC traffic in uplink transmissions, and energy consumption of the BSs as a function of levels of allocated RACH and PUSCH resources to MTC and HoC. Our analyses figure out the ways in which spectral/energy efficiency for the access network and QoS for non-MTC services could be traded to extend battery lifetimes of connected devices by compromising on the level of provisioned radio resources. For a full list of insightful observations derived from performance tradeoff analysis in serving MTC/HoC coexistence, one may refer to the appended paper D. In the following, one of the key finding of this paper is discussed. To ease understanding of coupling among battery lifetime for MTC, experienced delay by HoC, consumed energy by the BS, and uplink spectral efficiency, in Fig. 3.7 the optimized operation points have been depicted together. The background of this figure is a 2D view of energy consumption of the BS as a function of the ratio of allocated RACH and PUSCH resources to MTC. In the background, different energy consumption levels are depicted by different colors, ranged from yellow to dark blue, indicating high to low energy consumption regimes respectively. Considering the minimized BS energy consumption operation point as a reference, one sees in Fig. 3.7 that the average energy consumption of the BS, energy efficiency of machine-type communications, and experienced delay by HoC increase when extra radio resources are allocated to MTC traffic. The increase in energy efficiency of MTC communications is due to the fact that the access probability over RACH and success probability over PUSCH increase in the amount of allocated resources, which in turn results in decreasing QoS for HoC and in increasing energy consumption of the access network. Then, one sees that improper resource provisioning for MTC traffic not only degrades QoS of non-MTC services and decreases battery lifetime of MTC devices, but also increases energy consumption of the access network.



(a) Tradeoffs between the average energy consumption of the BS per unit time (left y-axis), and the experienced delay (right y-axis)



(b) Tradeoff between average energy consumption of the BS per unit time, and energy efficiency for $\mathcal{P}_2$ devices (in bit-per-joule)



(c) Empirical CDF of lifetimes for different mean listening time values

Figure 3.6: Performance tradeoffs. Simulations parameters can be found in appended paper C. (©2016 IEEE. Reused with permission.)

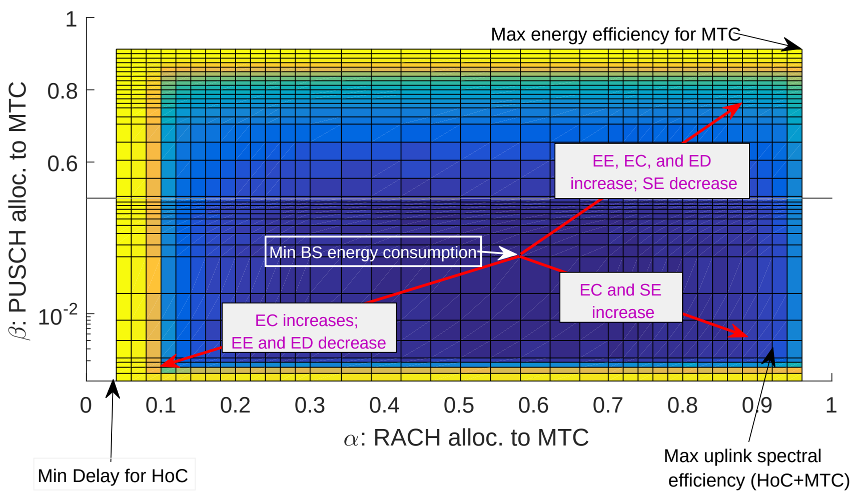


Figure 3.7: Performance tradeoffs (Reprinted from paper D, submitted for peer review to IEEE ICC 2017)

Chapter 4

Conclusions and Future Work

4.1 Concluding Remarks

Providing energy-efficient small data communications is the key requirement for enabling massive MTC services over cellular networks. In this dissertation, we first present accurate energy consumption models for cellular-based machine-type communications. The derived expressions are subsequently employed in defining individual and network battery lifetime models. Based on these models, we introduce a novel system design framework called *battery lifetime-aware network design*, and present battery lifetime-aware solutions for the clustering, scheduling, power control, BS sleeping control, and resource allocation problems. More specifically, we derive the cluster-size, cluster-head selection scheme, communications protocols within the clusters, scheduling scheme, BS sleeping techniques, and resource allocation schemes that aim at maximizing the network battery lifetime. Furthermore, we figure out the fundamental tradeoffs between energy saving for the access network, energy saving for machine-type devices, and spectral efficiency of the system, which in turn indicate the impact of prolonging network lifetime on other performance metrics. We also extend the derived solutions to existing cellular networks, and present simulation results in the context of LTE. Analysis of our developed analytical models figures out the ways in which scarce radio and energy resources for the access network and QoS for human-oriented communications could be preserved while coping with the ever increasing number of energy-limited machine-type devices in cellular networks. The simulation results show that the network lifetime is significantly extended under proposed lifetime-aware solutions. It is also shown that lifetime-aware uplink MTC scheduling lets machine nodes to last for a long time and die approximately at the same time, which in turn contributes significantly in network's maintenance costs reduction. In general, the results of this dissertation can be used to analyze and optimize the battery lifetime performance of deployed machine-type devices over cellular networks.

4.2 Future Works

The present dissertation includes our answers to RQ1-RQ13. These solutions have been also published in [30, 60, 61, 62, 63, 64, 65]. We are still working on RQ8 and RQ9. Our preliminary results on these research problems can be found in [66, 67, 68, 69]. Here, we have incorporated lifetime-awareness to a limited set of system design problems. In order to realize a battery-friendly cellular network, the proposed battery lifetime-aware framework should be extended to several other system design problems, including security and authentication, automatic retransmissions, density of BS deployment, handover, and etc. The derived lifetime-aware solutions from these problems would pave the way to integrate massive MTC services in cellular networks.

Included Papers

Bibliography

[1] L. Srivastava, T. Kelly, and et al., "The internet of things," International Telecommunication Union, Tech. Rep. 7, Nov 2005.

[2] M. Alam, R. Nielsen, and N. Prasad, "The evolution of M2M into IoT," in *First International Black Sea Conference on Communications and Networking*, July 2013, pp. 112–115.

[3] G. Lawton, "Machine-to-machine technology gears up for growth," *Computer*, vol. 37, no. 9, pp. 12–15, Sept. 2004.

[4] J. Kim, J. Lee, J. Kim, and J. Yun, "M2M service platforms: Survey, issues, and enabling technologies," *IEEE Communications Surveys Tutorials*, vol. 16, no. 1, pp. 61–76, Jan. 2014.

[5] S. E. C. DMC R&D center, "Cisco visual networking index: Global mobile data traffic forecast update, 20142019," Tech. Rep., March 2015.

[6] Ericsson, Huawei, NSN, and et al., "A choice of future M2M access technologies for mobile network operators," Tech. Rep., March 2014.

[7] H. Blaser and C. Fenger, "M2M service providers and LTE: new revenue opportunities," Tech. Rep., 2013.

[8] 3GPP TS 22.368 V13.1.0, "Service requirements for machine-type communications," Tech. Rep., 2014. [Online]. Available: http://www.3gpp.org/ftp/Specs/archive/22_series/22.368/22368-d10.zip

[9] H. Cho, "Machine to machine (M2M) communications technical report," 2011, IEEE 802.16 Broadband Wireless Access Working Group.

[10] F. Ghavimi and H. H. Chen, "M2M communications in 3GPP LTE/LTE-A networks: Architectures, service requirements, challenges, and applications," *IEEE Communications Surveys Tutorials*, vol. 17, no. 2, pp. 525–549, 2015.

[11] M. Beale, "Future challenges in efficiently supporting M2M in the LTE standards," in *IEEE Wireless Communications and Networking Conference Workshops*, April 2012, pp. 186–190.

[12] K. Zheng, F. Hu, W. Wang, W. Xiang, and M. Dohler, “Radio resource allocation in LTE-advanced cellular networks with M2M communications,” *IEEE Communications Magazine*, vol. 50, no. 7, pp. 184–192, July 2012.

[13] R. Ratasuk, A. Prasad, Z. Li, A. Ghosh, and M. Uusitalo, “Recent advancements in M2M communications in 4G networks and evolution towards 5G,” in *IEEE International Conference on Intelligence in Next Generation Networks*, 2015, pp. 52–57.

[14] M. Islam, A.-e. Taha, and S. Akl, “A survey of access management techniques in machine type communications,” *IEEE Communications Magazine*, vol. 52, no. 4, pp. 74–81, 2014.

[15] A. Gotsis, A. Lioumpas, and A. Alexiou, “M2M scheduling over LTE: Challenges and new perspectives,” *IEEE Vehicular Technology Magazine*, vol. 7, no. 3, pp. 34–39, Sept. 2012.

[16] K. Zheng, F. Hu, W. Wang, W. Xiang, and M. Dohler, “Radio resource allocation in LTE-advanced cellular networks with M2M communications,” *IEEE Communications Magazine*, vol. 50, no. 7, pp. 184–192, 2012.

[17] M. Gerasimenko, V. Petrov, O. Galinina, S. Andreev, and Y. Koucheryavy, “Impact of machine-type communications on energy and delay performance of random access channel in LTE-advanced,” *Transactions on Emerging Telecommunications Technologies*, vol. 24, no. 4, pp. 366–377, 2013.

[18] 3GPP TR 37.868 V11.0.0, “Study on RAN improvements for machine-type communications,” Tech. Rep., Sep 2011.

[19] K. Wang, J. Alonso-Zarate, and M. Dohler, “Energy-efficiency of LTE for small data machine-to-machine communications,” in *IEEE International Conference on Communications*, 2013, pp. 4120–4124.

[20] Nokia Networks, “5g use cases and requirements,” Tech. Rep., 2014.

[21] P. Park, P. Di Marco, P. Soldati, C. Fischione, and K. Johansson, “A generalized markov chain model for effective analysis of slotted IEEE 802.15.4,” in *IEEE International Conference on Mobile Adhoc and Sensor Systems*, Oct 2009, pp. 130–139.

[22] Ericsson, “Ericsson mobility report, on the pulse of the networked society,” Tech. Rep., 2015.

[23] ETSI TS 102 690 V1.1.1, “Machine-to-machine communications (M2M); functional architecture,” International Telecommunication Union, Tech. Rep., October 2011.

[24] A. Laya, L. Alonso, and J. Alonso-Zarate, "Is the random access channel of LTE and LTE-A suitable for M2M communications? a survey of alternatives," *IEEE Communications Surveys Tutorials*, vol. 16, no. 1, pp. 4–16, First 2014.

[25] 3GPP TS 22.368 v11.2.0, "Service requirements for machine-type communications; stage 1," Tech. Rep., Jun 2011.

[26] S. Duan, V. Shah-Mansouri, and V. W. Wong, "Dynamic access class barring for M2M communications in LTE networks," in *IEEE Global Communications Conference*, 2013, pp. 4747–4752.

[27] Y. Liu, C. Yuen, X. Cao, N. Ul Hassan, and J. Chen, "Design of a scalable hybrid MAC protocol for heterogeneous M2M networks," *IEEE Internet of Things Journal*, vol. PP, no. 99, 2014.

[28] S. Y. Lien and K. C. Chen, "Massive access management for QoS guarantees in 3GPP machine-to-machine communications," *IEEE Communications Letters*, vol. 15, no. 3, pp. 311–313, March 2011.

[29] H. S. Dhillon, H. C. Huang, H. Viswanathan, and R. A. Valenzuela, "Power-efficient system design for cellular-based machine-to-machine communications," *IEEE Transactions on Wireless Communications*, vol. 12, no. 11, pp. 5740–5753, November 2013.

[30] A. Azari and G. Miao, "Energy efficient MAC for cellular-based M2M communications," in *2nd IEEE Global Conference on Signal and Information Processing*, 2014.

[31] J. Sachs, N. Beijar, P. Elmdahl, J. Melen, F. Militano, and P. Salmela, "Capillary networks–a smart way to get things connected," *Ericsson Review, September*, vol. 9, 2014.

[32] V. Mii, J. Mii, X. Lin, and D. Nerandzic, in *Ad-hoc, Mobile, and Wireless Networks*. Springer Berlin Heidelberg, 2012, vol. 7363, pp. 413–423.

[33] M. I. Hossain, A. Laya, F. Militano, S. Iraji, and J. Markendahl, "Reducing signaling overload: Flexible capillary admission control for dense MTC over LTE networks," in *IEEE 26th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications*, 2015, pp. 1305–1310.

[34] A. Asadi, Q. Wang, and V. Mancuso, "A survey on device-to-device communication in cellular networks," *IEEE Communications Surveys Tutorials*, vol. 16, no. 4, pp. 1801–1819, Fourthquarter 2014.

[35] A. Laya, L. Alonso, J. Alonso-Zarate, and M. Dohler, "Green MTC, M2M, internet of things," *Green Communications: Principles, Concepts and Practice*, pp. 217–236, 2015.

[36] N. Pratas and P. Popovski, "Zero-outage cellular downlink with fixed-rate D2D underlay," *IEEE Transactions on Wireless Communications*, vol. 14, no. 7, pp. 3533–3543, July 2015.

[37] G. Rigazzi, N. Pratas, P. Popovski, and R. Fantacci, "Aggregation and trunking of M2M traffic via D2D connections," in *IEEE International Conference on Communications*, June 2015, pp. 2973–2978.

[38] G. Rigazzi, F. Chiti, R. Fantacci, and C. Carlini, "Multi-hop D2D networking and resource management scheme for M2M communications over LTE-A systems," in *International Wireless Communications and Mobile Computing Conference*, 2014, pp. 973–978.

[39] N. AbuAli, "Enabling technologies of energy efficient cooperative M2M networks: Benefits and challenges," in *IEEE 38th Conference on Local Computer Networks Workshops*, 2013, pp. 987–993.

[40] M. Mehaseb, Y. Gadallah, A. Elhamy, and H. El-Hennawy, "Classification of LTE uplink scheduling techniques: An M2M perspective," *IEEE Communications Surveys Tutorials*, no. 99, 2015.

[41] A. S. Lioumpas and A. Alexiou, "Uplink scheduling for machine-to-machine communications in LTE-based cellular systems," in *IEEE GLOBECOM Workshops*, 2011, pp. 353–357.

[42] S. Zhenqi, Y. Haifeng, C. Xuefen, and L. Hongxia, "Research on uplink scheduling algorithm of massive M2M and H2H services in LTE," in *IET international conference on Information and communications technologies*, 2013, pp. 365–369.

[43] M. K. Giluka *et al.*, "Class based dynamic priority scheduling for uplink to support M2M communications in LTE," in *IEEE World Forum on Internet of Things*, 2014, pp. 313–317.

[44] N. Afrin, J. Brown, and J. Y. Khan, "Performance analysis of an enhanced delay sensitive LTE uplink scheduler for M2M traffic," in *Australasian Telecommunication Networks and Applications Conference*, 2013, pp. 154–159.

[45] S. A. Mahmud *et al.*, "Fairness evaluation of scheduling algorithms for dense M2M implementations," in *IEEE Wireless Communications and Networking Conference Workshops*, April 2014, pp. 134–139.

[46] A. Aijaz, M. Tshangini, M. Nakhai, X. Chu, and A. Aghvami, "Energy-efficient uplink resource allocation in LTE networks with M2M/H2H co-existence under statistical QoS guarantees," *IEEE Transactions on Communications*, vol. 62, no. 7, pp. 2353–2365, July 2014.

[47] N. Networks, "Looking ahead to 5G: Building a virtual zero latency gigabit experience," Tech. Rep., 2014.

[48] M. Kalil *et al.*, "Low-complexity power-efficient schedulers for LTE uplink with delay-sensitive traffic," *IEEE Transactions on Vehicular Technology,*, vol. 64, no. 10, pp. 4551–4564, 2015.

[49] G. Miao, N. Himayat, G. Li, and S. Talwar, "Low-complexity energy-efficient scheduling for Uplink OFDMA," *IEEE Transactions on Communications*, vol. 60, no. 1, pp. 112–120, Jan. 2012.

[50] G. Miao and G. Song, *Energy and Spectrum Efficient Wireless Network Design*. Cambridge University Press, 2014.

[51] I. Abdalla and S. Venkatesan, "A QoE preserving M2M-aware hybrid scheduler for LTE uplink," in *IEEE International Conference on Selected Topics in Mobile and Wireless Networking*, 2013, pp. 127–132.

[52] N. Abu-Ali *et al.*, "Uplink scheduling in LTE and LTE-advanced: Tutorial, survey and evaluation framework," *IEEE Communications Surveys Tutorials*, vol. 16, no. 3, pp. 1239–1265, Third 2014.

[53] 3GPP, "USF capacity evaluation for MTC," Tech. Rep., 2010, TSG GERAN 46 GP-100894.

[54] G. Miao, N. Himayat, and G. Y. Li, "Energy-efficient link adaptation in frequency-selective channels," *IEEE Transactions on Communications*, vol. 58, no. 2, pp. 545–554, 2010.

[55] 3GPP TS 36.213, "Evolved universal terrestrial radio access (E-UTRA), physical layer procedures," Tech. Rep., (Release 13).

[56] D. Zwillinger, *Table of integrals, series, and products*. Elsevier, 2014.

[57] I. Dimitriou, "Analysis of a priority retrial queue with dependent vacation scheme and application to power saving in wireless communication systems," *The Computer Journal*, pp. 1363–1380, 2012.

[58] 3GPP, "Study on provision of low-cost machine-type communications (MTC) user equipments (UEs)," *3GPP TR 36.888 V12.0.0*, 2013.

[59] Nokia Networks, "LTE-M – optimizing LTE for the internet of things," Tech. Rep., 2015.

[60] A. Azari and G. Miao, "Lifetime-aware scheduling and power control for M2M communications over LTE networks," in *IEEE Wireless Communications and Networking Conference*, 2015.

[61] A. Azari, G. Miao, and T. Hwang, "E^2-MAC: Energy efficient medium access for massive M2M communications," *To appear in the IEEE Transactions on Communications*, 2016.

[62] A. Azari and G. Miao, "Lifetime-aware scheduling and power control for cellular-based M2M communications," in *IEEE Wireless Communications and Networking Conference (WCNC)*, 2015.

[63] ——, "Battery lifetime-aware base station sleeping control with M2M/H2H coexistence," accepted in IEEE Globecom, 2016.

[64] A. Azari, "Energy-efficient scheduling and grouping for machine-type communications over cellular networks," *Ad Hoc Networks*, vol. 43, pp. 16 – 29, 2016.

[65] A. Azari and G. Miao, "Network lifetime maximization for cellular-based M2M networks," *Submitted to IEEE Transactions on Wireless Communications*, 2016.

[66] A. Azari, M. I. Hossain, and J. Markendahl, "RACH dimensioning for reliable MTC over cellular networks," submitted to Proc. IEEE VTC, 2017.

[67] A. Azari and G. Miao, "Fundamental tradeoffs in resource provisioning for IoT services over cellular networks," submitted to Proc. IEEE ICC, 2017.

[68] M. I. Hossain, A. Azari, J. Markendahl, and J. Zander, "Enhanced random access: Initial access load balance in highly dense LTE-A networks for multiservice (H2H-MTC) traffic," submitted to Proc. IEEE ICC, 2017.

[69] M. I. Hossain, A. Azari, and J. Zander, "DERA: Augmented random access for cellular networks with dense H2H-MTC mixed traffic," accepted in IEEE Globecom Workshops, 2016.