Type-Based Multiple Access for massive connectivity: design and implementation aspects

Master thesis

Michel Massanet Ginard

AALBORG UNIVERSITY
STUDENT REPORT

Aalborg University
Electronics and IT

Here you can write something about which tools and software you have used for typesetting the document, running simulations and creating figures. If you do not know what to write, either leave this page blank or have a look at the colophon in some of your books.

AALBORG UNIVERSITY
STUDENT REPORT

Electronics and IT
Aalborg University
http://www.aau.dk

Title:
Type-Based Multiple Access for massive connectivity: design and implementation aspects

Theme:
Wireless Communication Systems

Project Period:
Master thesis

Project Group:
WCS10-1050

Participant(s):
Michel Massanet Ginard

Supervisor(s):
Petar Popovski (AAU)
Johannes Dommel (Fraunhofer HHI)
Zoran Utkovski (Fraunhofer HHI)

Copies: 1

Page Numbers: 50

Date of Completion:
June 4, 2020

Abstract:

The increased number of Internet-of-Things devices comes with the emergence of massive connectivity wireless systems. On these scenarios, a typical application is the estimation of a field parameter or Quantity of Interest (QoI). Type-Based Multiple Access (TBMA) is a communication protocol in charge of estimating a given QoI based on the histogram of the received measurements from devices. On this thesis, we investigate the performance of TBMA when it is implemented on a wireless technology. Narrowband Internet of Things (NB-IoT) is elected as the baseline technology due to its reduced signal bandwidth and its low-power characteristics. Three transmission schemes are proposed and their estimation performance is evaluated under different scenarios. Error probability results show the limitations of TBMA in terms of minimum number of preambles and maximum states of the QoI. Furthermore, we analyse the implementation aspects when incorporating the TBMA protocol to a real wireless system by focusing on NB-IoT and propose alternatives upon system limitations.

Contents

Preface

This work is the outcome of the Master Thesis conducted during the 10th semester of the Master's programme in Wireless Communication Systems at Aalborg University. This master thesis has been conducted in collaboration with the Fraunhofer Heinrich-Hertz-Institute (HHI). Citations follow the IEEE referencing method and Units follow the SI system.

As author of this thesis, I would specially like to acknowledge to Fraunhofer HHI for supporting me during this thesis and for letting me participate on such innovative project. I would also like to acknowledge to following persons: my supervisors (both HHI and AAU) for tracking and advising me on the development of this topic; Dennis Wieruch for following my track during the implementation of the NB-IoT preambles on MATLAB; and Fernando Loro Velardo and Miguel Angel Gutiérrez Estévez for providing me with a wide vision of future machine learning enhancements.

Aalborg University, June 4, 2020

Author 1

<mmassa18@student.aau.dk>

List of Figures

List of Tables

Abbreviations

3GPP 3rd Generation Partnership Project

AWGN Additive White Gaussian Noise

BS Base Station

CE Coverage Enhancement

CP Cyclic Prefix

DCI Downlink Control Indicator

DL Downlink

eMBB enhanced Mobile Broadband

eNB evolved Node B

EPA Extended Pedestrian A

ETU Extended Typical Urban

GPRS General Packet Radio Service

GSM Global System for Mobile Communications

HTC Human Type Communication

IoT Internet-of-Things

ITU-R International Telecommunication Union - Radiocomunication Sector

LLR Log-Likelihood Ratio

LOS Line-Of-Sight

LPWA Low Power Wide Area

LTE Long Term Evolution

LTE-M Long Term Evolution MTC

mMTC massive-MTC

MTC Machine Type Communication

NB-IoT Narrowband Internet of Things

NOMA Non-Orthogonal Multiple Access

NPBCH Narrowband Physical Broadcast Channel

NPDCCH Narrowband Physical Downlink Control Channel

NPDSCH Narrowband Physical Downlink Shared Channel

NPRACH Narrowband Physical Random Access Channel

NPUSCH Narrowband Physical Uplink Shared Channel

NRS Narrowband Reference Signals

PMF Probability Mass Function

QoI Quantity of Interest

RA Random Access

SG Symbol Group

SIB System Information Block

SNR Signal to Noise Ratio

TA Timing Advance

TBMA Type-Based Multiple Access

TDD Time Division Duplex

TDL Tapped Delay Line

TV Total Variation

UE User Equipment

UL Uplink

URLLC Ultra Reliable Low Latency Communications

WSN Wireless Sensor Networks

1

Introduction

1.1 Massive connectivity

The new era of telecommunications is enabled by the arrival of 5G to our societies. 5G technologies consider novel use cases that previous mobile network generations were not able to fulfill. In addition to the well-known Human Type Communication (HTC), the applications addressed by 5G include the new paradigm of Machine Type Communication (MTC) which represents most of the challenges on the development of new wireless solutions. 5G applications can be classified into three main categories [1]: enhanced Mobile Broadband (eMBB), Ultra Reliable Low Latency Communications (URLLC) and massive-MTC (mMTC). See Figure 1.1 for some application examples. On eMBB, main requirements are generally given by HTC demands with scenarios of typically low quantity of devices and high data transmissions. The other two, i.e. URLLC and mMTC, are encompassed on the Internet-of-Things (IoT) framework, where generally the type of communication refers to MTC. On URLLC, higher priority is given to devices that require robust connectivity and very low latency whereas mMTC addresses massive connectivity and comprise a range of applications with a massive amount of devices transmitting small data packets with generally some degree of tolerance on reliability and latency.

The number of connected devices in wireless networks is expected to grow with the future enabling 5G technologies. IoT smart objects are expected to reach 212 billion entities deployed globally by the end of 2020 [3]. Typical mMTC applications are smart or remote metering, smart lights, smart cities or smart agriculture. For instance, an exemplary smart metering scenario would be thousands of temperature meters deployed over a large industrial plant reporting their measured values. As for this exemplary applications, and generally many mMTC applications, the device's density might reach up to 30.000 devices in a single cell [4].

The characteristics of the massive connectivity paradigm are distinct from the well-known HTC applications. Most common and relevant mMTC characteristics are [5], [6]:

- Small data packets

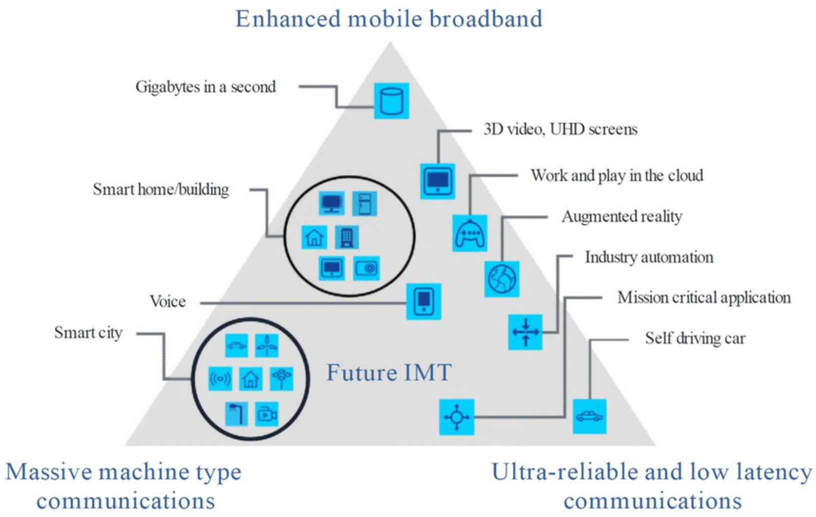


Figure 1.1: 5G usage scenarios defined by the ITU-R [2]

- Traffic mostly in Uplink (UL)
- Large number of device (up to 30.000 devices per cell)
- Energy-efficient transmissions
- Sporadic device activity: periodic or event-driven transmissions

With regard to the last bullet point, depending on the communication protocol the devices can transmit either periodically (semi-persistent scheduling) or event-driven (i.e. based on an event, for example a power outage). Under periodic transmissions, the large number of devices can be split on transmission groups and scheduled by the Base Station (BS) at different time windows. However, with thousands of devices and continuous periodic updates the BS requires novel techniques to adapt the large quantity of transmissions. Under event-driven transmissions, the number of active devices is given by the ones affected by the event. In this case, the device activity is sporadic and generally the radio access is uncoordinated. One fact to highlight here is that the data generated by the devices could be correlated. Taking up the previous example, in case of a power outage, several devices would detect the same event and therefore the information transmitted by most of the devices would be the same. In all cases the main challenge is the following: *how to adapt or coordinate the high number of devices transmitting to the same access point?*

Conventional multiple access technologies are not scalable for the future of massive connectivity. Technologies like Long Term Evolution (LTE) include large overhead on the radio access and high redundancy coding which is highly inefficient for MTC. For example, the redundant information involved on channel estimation and coding procedures is potentially higher than the payload information. On current standardisation processes an evolution of LTE is being carried out in order to satisfy the most common and priority MTC requirements. The most sounded technologies are Narrowband Internet of Things (NB-IoT), a Low Power Wide Area (LPWA) technology with a signal bandwidth of 180 kHz and low-cost radio chips, and a variant of LTE for Machine Type Communication (MTC) called Long Term Evolution MTC (LTE-M), which meets higher data rates applications [5]. Despite both technologies are called to be used by MTC applications, NB-IoT is the greatest candidate for massive connectivity due to its low power consumption, reduced signal bandwidth and its coverage enhancements [7]. Nonetheless, current NB-IoT features are still not capable of handling massive number of devices transmitting simultaneously, so additional wireless transmission schemes must be investigated to meet the strict demands of some massive connectivity use cases.

On the current literature, novel approaches address the massive connectivity challenge [5], [8]. One research branch deals with Non-Orthogonal Multiple Access (NOMA) techniques. In NOMA, multiple devices share the same physical resource grid in a non-orthogonal fashion. The outcome is a considerable reduction of needed physical resources to multiplex all devices. Other research directions argue with grant-free transmissions. The motivation behind this approach is to face the overhead generated when devices access to the radio resources. For example, on most uncoordinated multiple access schemes, the number of collisions considerable increases with high number of devices (e.g. ALOHA) or the overhead generated for collision avoidance can overwhelm the system (e.g. collision avoidance-ALOHA) [8]. Grant-free schemes emerge as a potential solution where the devices transmit data without the necessity of performing a Random Access (RA) procedure.

One typical function of Wireless Sensor Networks (WSN) is the sensing of a state of the nature or Quantity of Interest (QoI). Sensors (or generally devices) collect information of the QoI and forward it to the fusion-center (a central processing entity) which is partially or fully in charge of the detection or estimation task. This scenario is usually referred as distributed estimation/detection due to the spatially distributed measurements taken by devices [9]. Research work has addressed the detection and estimation problems of WSN taken into account efficient transmissions with reduced bandwidth and energy consumption. On this context, two main strategies are found: local and centralized estimation [9]–[11]. One remark to be introduced is the possible nomenclature confusion between centralized and decentralized estimation. Local and centralized estimation are strategies or configurations of distributed estimation framework. Distributed estimation is also named by literature as decentralized estimation. To avoid confusion, we refer as distributed estimation to the estimation of a field parameter using measurements taken by physically distributed sensors. Local estimation is based on *intelligent devices* that have knowledge of the QoI statistics and therefore can make local decisions based on the measured values. On this case, devices transmit hard of soft information regarding their local estimates to the fusion-center which is in charge of making the final decision. Centralized estimation includes *dumb devices* which do not have knowledge of the state of the nature statistics and therefore cannot make local decisions. On this case, devices take measurements and forward them to the fusion-center which conducts the full estimation task based on all gathered measurements. Note that due to the limited amount of

UL resources, on the local estimation approach soft-information has to be quantized just as the reported measurements on centralized estimation.

One of the most well-known centralized estimation approaches is Type-Based Multiple Access (TBMA). TBMA was proposed by Mergen in [12] as a solution for parameter estimation in large number of devices scenarios with limited transmission energy. TBMA is communication protocol in which sensors transmit orthogonal waveforms encoding the devices' measurements and the fusion-center infers the QoI based on the histogram of the received measurements from the devices.

1.2 Type-Based Multiple Access

At this work we focus on TBMA as a communication protocol to enable distributed estimation. TBMA is categorised as an information-centric protocol because the inference of the QoI relies only on the estimated histogram of the devices' observations. This approach differs from classic device-centric protocols where individual observations are generally requested. Note that device-centric protocols (as they exist on current technologies like NB-IoT or LTE-M) are generally inefficient under massive amount of active devices. The fundamental pillar of TBMA is that by assigning an orthogonal codeword to each of the possible measured values by the devices, the receiver is able to estimate the histogram of the observations based on the received signals and infer from it the value of the QoI. Once the information about the QoI is obtained, device-centric approaches might be considered as an extension to TBMA in order to acquire further information of particular devices measuring specific values (e.g. where the device is located, other parameters of interest, etc.).

Some standpoints to consider TBMA as a potential approach on mMTC use cases are now mentioned. TBMA takes advantage of one of mMTC characteristics: the data correlation. The correlation of the measured data would cause that the observed values by devices have high probability to be repeated among several devices (e.g. high temperatures on a room where a fire has accidentally started). If observed values are mapped to orthogonal preambles, the devices with same observed value would overlap on a non-orthogonal fashion. This fact can be taken as an advantage because in order to estimate the histogram of the observations, the computation of how many devices transmitted each preamble would be sufficient. However, note that the possible observed values are constrained by the quantification levels defined by the system (usually limited to the number of preambles). TBMA can be considered a special form of NOMA due to devices transmit non-orthogonal signals if same measured value is repeated among devices. Moreover, TBMA includes grant-free transmissions because devices already know which physical resource to use, so no RA request is needed. In this sense, TBMA reduces the limited number of UL resources and reduces the overhead and collisions generated on grant-based communications.

TBMA research directions have focused on information-theoretic standpoint. Some approaches analyse the use of TBMA for centralized detection versus local detection [9]. Others showed the necessity of channel normalisation under zero-mean channel fading [12] and analysed the effect of the transmission rate on the estimation performance [13]. Other papers proposed alternative implementations of TBMA [14] through a variant based on non-orthogonal codewords with the objective of improving the spectral efficiency. The TBMA setup was extended to a multi-cell scenario on [15] discussing where the centralised process should be located: at the cloud or at the edge. Most existing work claims for the potential of TBMA in terms of savings of power and bandwidth specially on the regime of large number of devices.

So far, the advantages of TBMA for massive connectivity have been extensively discussed, but the implementation of TBMA to a real wireless system has not been addressed yet. Several approaches are already available on the literature and the analysis of how feasible is to include it on existing or future wireless technologies is of great interest.

1.3 Project scope

Then main characteristics and challenges of massive connectivity have been introduced on the previous section. This projects focuses on TBMA as a potential enabler of information-centric uses cases under massive connectivity scenarios. TBMA research directions have focused on information-theoretic standpoint and its possible implementation challenges and adaptation have not been investigated yet.

The scope of this project is to design realistic transmission schemes including the TBMA principle and to study the implementation aspects that might bring into practice on a real wireless system. Due to its physical layer design, NB-IoT will be taken as the reference technology to be used for TBMA. For the evaluation of the transmission schemes, different scenarios will be recreated and the influence of most relevant TBMA parameters will be discussed. Considering the obtained results, the implementation aspects of each transmission scheme and of the general TBMA system will be analysed and modifications of current or future standards will be proposed.

The remainder of this document is organised as follows. Chapter 2 introduces the NB-IoT technology and the RA preambles. Chapter 3 describes the system model considered for this work. Chapter 4 shows the performance of the transmission schemes under two scenarios, discusses the results and proposes enhancements. Chapter 5 evaluates the implementation aspects involved on the establishment of a TBMA-based protocol on a real wireless system and Chapter 6 concludes this master thesis reflecting the main outcomes of this work.

2

Narrowband Internet of Things

Narrowband Internet of Things (NB-IoT) is a cellular technology standardised by the 3rd Generation Partnership Project (3GPP) and introduced from Release 13. It was designed to fulfill some MTC requirements with focus on LPWA device characteristics, battery powered operation up to 10 years and radio module cost under 5$ [16]. Taking the heritage from LTE, it was designed as a suitable precursor of 5G mMTC technologies [7]. The NB-IoT specifications on this thesis have been taken from 3GPP Release 14.

2.1 Physical layer

NB-IoT signals occupy a bandwidth of 180 kHz in both Downlink (DL) and UL. In DL, Orthogonal Frequency Division Multiplexing is applied with a subcarrier spacing of 15 KHz. In the UL, multi-tone and single-tone transmissions are supported based on Single Carrier Frequency Division Multiple Access. The multi-tone mode uses 15 KHz subcarrier spacing and the single-tone mode uses either 3.75 KHz or 15 KHz subcarrier spacing. One of the most relevant features from the NB-IoT physical layer is the inclusion of repetitions on the transmissions (i.e. the signal is transmitted a number of repeated times according to the device's channel quality). Due to devices might experience different radio conditions, NB-IoT defines up to three Coverage Enhancement (CE) levels. The CE levels represent different groups or clusters of devices. The decision of how many CE levels are defined on each cell is a network operator design parameter. Among other features, each CE level defines the number of repetitions that a device or the BS has to incorporate during a transmission. Thanks to the time diversity implied by the repetitions and its narrowband signals, NB-IoT has a coverage improvement of 20 dB with respect to Global System for Mobile Communications (GSM) and General Packet Radio Service (GPRS) [17].

In DL three physical channels are defined [18]: the Narrowband Physical Broadcast Channel (NPBCH), which carries the information that is broadcasted to all the users, the Narrowband Physical Downlink Control Channel (NPDCCH), which indicates to User Equipment (UE)s information about their DL data and UL grants or UL scheduling information, and the

Narrowband Physical Downlink Shared Channel (NPDSCH), which contains traffic data for specific UEs. In UL, only two physical channel are defined:

- Narrowband Physical Uplink Shared Channel (NPUSCH): it carrier the UL data packet transmissions.
- Narrowband Physical Random Access Channel (NPRACH): it is used for performing initial access to the network, to request UL resources or to reconnect to the BS after a link failure.

2.2 NPRACH preambles

Random Access (RA) procedure is usually conducted for initial UE UL synchronization when UL or DL data arrive and the UE is not synchronised. Additionally, it is also executed for initial access when establishing a new radio link. On NB-IoT, RA is conducted by means of the NPRACH. The RA procedure is based on four-stage handshake message with the following procedure: (1) the UE randomly selects and transmit one of the available NPRACH preambles, also kwown as RA preambles. From the received signal, the evolved Node B (eNB) estimates the UE Timing Advance (TA). (2) If the eNB successfully detects the NPRACH preamble, it responds with a message including a temporary identification for the UE and the TA estimate. The (3) and (4) messages initialise and conclude the contention-resolution procedure to detect possible collisions not detected during the first two steps.

With relation to the massive connectivity challenge, NB-IoT RA procedure expects that only one UE attempts to access during the same RA window using the same NPRACH preamble. In the hypothetical case that two (or more) UEs select the same preamble, a collision may occur (note that this is more likely for larger number of devices). The RA procedure under massive number of devices has been already addressed by [19]. For the RA attempts, devices make use of NPRACH preambles defined on [20]. On predecessor technologies like LTE, the RA preambles are made of Zadoff-Chu sequences which occupy a bandwidth of 1.08 MHz. Additionally, its peak-to-average power rate requires higher power amplifier back-off [21]. Both facts make unachievable the adaptation of the LTE preambles to NB-IoT channel bandwidth and low-power-based transmissions. Therefore, a new design was required for the NPRACH preambles.

The NPRACH has a total channel bandwidth of 180 KHz and contains the RA preambles. The NPRACH transmissions use a single-tone 3.75 KHz subcarrier spacing where a total of 48 subcarriers are found on the NPRACH bandwidth. Hence, 48 preambles are defined on NB-IoT [17]. One NPRACH preamble is composed of $N_{SG} = 4$ Symbol Group (SG)s, where each symbol group is made of one Cyclic Prefix (CP) and five symbols. Two SG formats are stated and define the duration of the CP. Format 0 has a CP duration of $T_{CP} = 66.7\ \mu s$ whereas format 1 has $T_{CP} = 267\ \mu s$. The use of format 0 allows a maximum cell range of 8 km while with format 1 up to 35 km can be covered [21]. Both formats have a five symbols duration of $T_{SEQ} = 1.333\ ms$ [20]. The duration of a symbol group can be defined by $T_{SG} = T_{CP} + T_{SEQ}$.

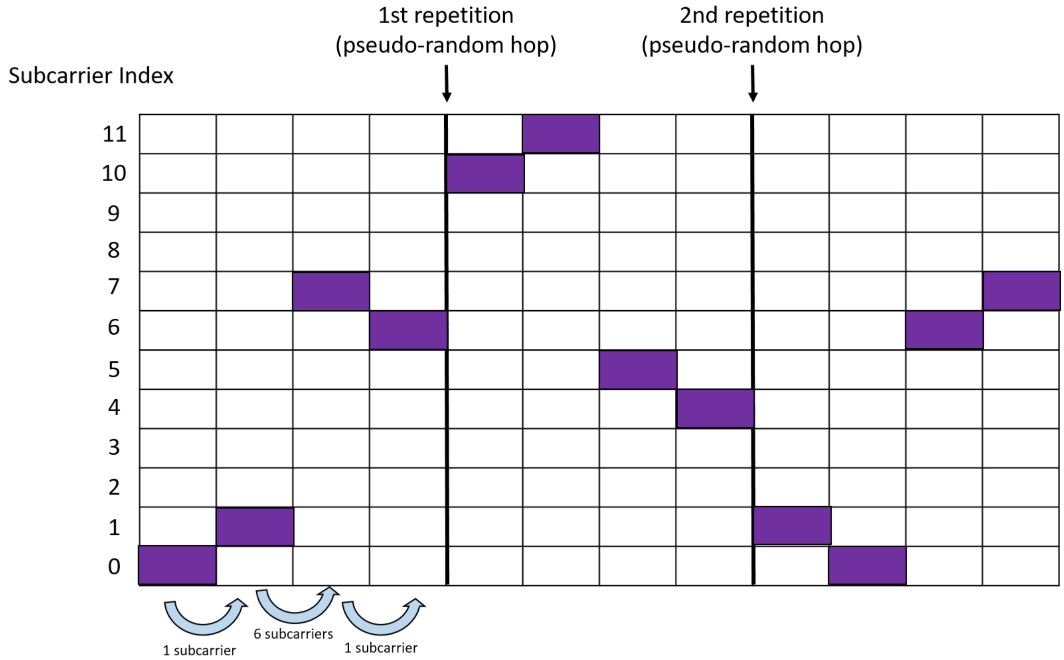


Figure 2.1: NPRACH preambles illustration.

Preamble repetitions can be applied depending on the CE level experienced by the UE with a number of repetitions $N_{rep} \in \{1, 2, 4, 8, 16, 32, 64, 128\}$. More number of repetitions have been included from Release 15. In case repetitions are included, the NPRACH preamble are composed of $N_{SG} = 4N_{rep}$ symbol groups. Each CE group, is assigned a set of 12, 24, 36 or 48 preambles, that is, the all available preambles are distributed among the CE groups, so devices on different groups do not share same NPRACH preambles. The symbol groups are transmitted following a particular frequency hopping pattern defined in [20] and illustrated in Figure 2.1. On the one side, there is a 4-symbol groups frequency hopping fixed pattern: a one subcarrier hopping distance for the 1^{st} and 3^{rd} hop, and a 6-subcarriers hopping distance for the 2^{nd} hop. On the other side, for every repetition, an outer layer pseudo-random hopping is applied which depends on the cell identity that the UE is connected to [21]. This configurations helps reducing the interference with RA attempts from neighbouring cells. Note that the symbol groups can only hop among a set of 12 concatenated subcarriers, which means a maximum extension of the preamble on 45 KHz. Hence, the frequency hopping pattern is described by the following rule

$$\tilde{n}_{SC}^{RA}(i) = \begin{cases} (\tilde{n}_{SC}^{RA}(0) + f(i/4))\, mod\, N_{SC}^{RA} & i\, mod\, 4 = 0\, and\, i > 0 \\ \tilde{n}_{SC}^{RA}(i-1) + 1 & i\, mod\, 4 = 1, 3\ and\ \tilde{n}_{SC}^{RA}(i-1) mod\, 2 = 0 \\ \tilde{n}_{SC}^{RA}(i-1) - 1 & i\, mod\, 4 = 1, 3\ and\ \tilde{n}_{SC}^{RA}(i-1) mod\, 2 = 1 \\ \tilde{n}_{SC}^{RA}(i-1) + 6 & i\, mod\, 4 = 2\ and\ \tilde{n}_{SC}^{RA}(i-1) < 6 \\ \tilde{n}_{SC}^{RA}(i-1) - 6 & i\, mod\, 4 = 2\ and\ \tilde{n}_{SC}^{RA}(i-1) \geq 6 \end{cases} \tag{2.1}$$

where $n_{SC}^{RA}(i) = n_{start} + \tilde{n}_{SC}^{RA}(i)$ is the subcarrier index of the i-th symbol group, $N_{SC}^{RA} = 12$, $\tilde{n}_{SC}^{RA}(0) \in \{1, 2, ..., 48\}$ and $f(t)$ is the pseudo-random function that defines the initial preamble for a new repetition (see [20] for the definition of $f(t)$). Note that $\tilde{n}_{SC}^{RA}(0)$ is the selected initial subcarrier which by default is selected randomly and identifies the selected preamble. The parameter n_{start} is given by the chosen initial subcarrier index and the NPRACH configuration included on the System Information Block (SIB) broadcasted by means of NPBCH.

Following with the frequency hopping rule, the transmitted signal at symbol group $i \in \{0, 1, 2, ..., N_{SG} - 1\}$ is given by

$$x_i(t) = \beta_{NPRACH} e^{j2\pi(n_{SC}^{RA}(i) - N_{SC}^{UL} 0.5 k_0 + 0.5)\Delta f_{RA}(t - T_{CP})}, \tag{2.2}$$

where $0 \leq t \leq T_{SG}$, β_{NPRACH} is a scaling factor according to the transmitted power [22], $\Delta f_{RA} = 3.75$ KHz, k_0 accounts for the difference in subcarrier spacing between NPRACH and NPUSCH, and N_{SC}^{UL} is the number of consecutive subcarriers in an UL resource unit.

3

System model

3.1 Scenario

We consider a wireless system between a fusion-center (also known as BS) and a large number of devices with limited power consumption. The BS has fixed location and the devices are considered to be static (e.g. embedded on machinery, walls or underground). The propagation delay of each device is determined by the physical distance between device and BS. The wireless system follows a narrowband transmission scheme, i.e. the transmitted signals have a bandwidth much smaller than the channel bandwidth. The system has been operating during a long-time period allowing a learning-process of some communication parameters.

The BS has N attached devices of which K are active being $0 \leq K \leq N$. The devices are considered to be active in a given a transmission interval when they have UL data to transmit during this period. The periodicity of the transmission interval is not relevant. The transmission interval defines the time instance where all active devices transmit simultaneously. The devices are denoted by $n \in \mathcal{N} = \{1, 2, ..., N\}$ and the active devices are referred as $k \in \mathcal{K} = \{\varnothing, 1, 2, ..., K\}$ where $\mathcal{K} \subset \mathcal{N}$. Both n and k will be used to refer to general and active devices respectively. We focus on the UL communication direction, that is, the transmitted information from the active devices to the fusion-center. The number of antennas at devices and BS is $N_{TX} = 1$ and $N_{RX} = 1$ respectively. The extension of the system model to $N_{RX} = 2$ is trivial but comes at cost of more bulky notation. Nonetheless on this project only one antenna at the BS is assumed. Figure 3.1 shows the system model for K active devices. The included notation will be specified during the following sections.

The devices observe a Quantity of Interest (QoI) which is a physical parameter (or state of nature) such as temperature, pressure or pollution level. The QoI is described by the discrete random variable θ which can take Q values/states (i.e. $\theta \in \{\theta_1, \theta_2, ..., \theta_Q\}$) and has a Probability Mass Function (PMF) $Pr[\theta = \theta_q] = p_q$ where $q \in \mathcal{Q} = \{1, 2, ..., Q\}$. Despite all N devices of the system take observations of the QoI, we center our interests on the measurements taken by the K active devices. Each device k takes a measurement/observation X_k which can take values on the alphabet $\mathcal{M} = \{1, 2, ..., M\}$ (i.e. $X_k = m$ where $m \in \mathcal{M}$). Usually the

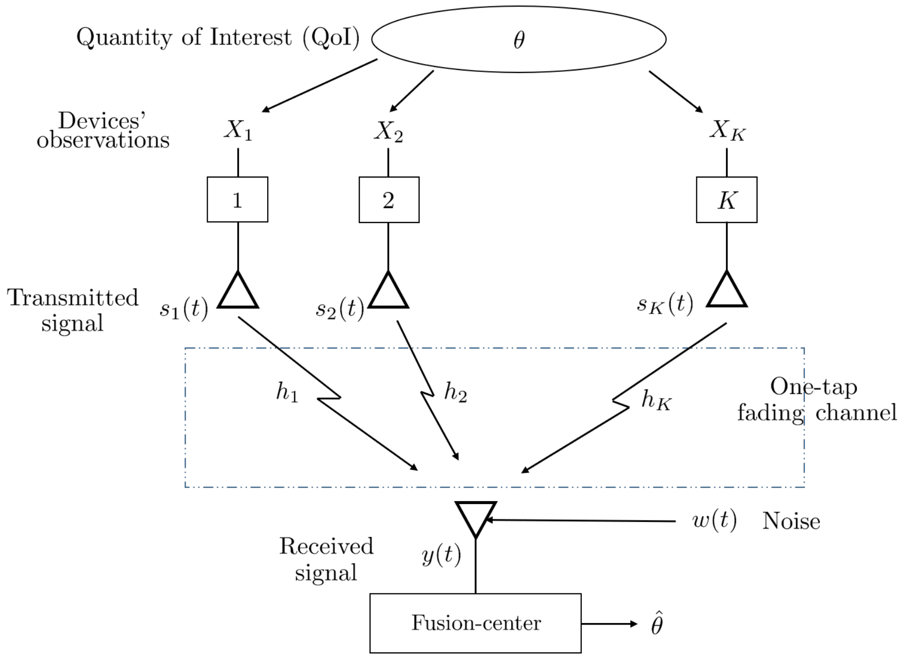


Figure 3.1: System model illustration.

observations X_k are real-valued measurements which are later quantized into M levels resulting into a discrete random variable. Due to TBMA makes use of the quantized observations, we directly assume that X_k can only take discrete values. The problem of optimized quantization from analogue to discrete observations is out of the scope of this project. The distribution of observations X_k is conditioned by θ, begin its conditional PMF $Pr[X_k = m|\theta = \theta_q] = p_{\theta_q}(m)$. The family of distributions p_{θ_q} is known at the fusion-center. The distribution of devices observations' conditioned by θ is generally referred as p_θ.

3.2 Signal model

All devices on the system share a codebook $\boldsymbol{\phi} = [\phi^1(t), \phi^2(t), ..., \phi^M(t)]$ where $0 \leq t \leq T$ being T the duration of the codeword $\phi^m(t)$, and $\{\phi^i(t), \phi^j(t)\}$ are orthogonal signals with unit energy $\forall ij, i \neq j$. On this work, preamble and signature are used as synonyms of codeword. Active devices map their observation X_k to a preamble from the codebook $\boldsymbol{\phi}$. The chosen preamble by device k follows a TBMA protocol being $\phi^m(t)$ when the measured value is $X_k = m$. Therefore, each preamble $\phi^m(t)$ represents a possible observed value m. We denote $s_k(t)$ to the transmitted signal by device k. Note that this only applies to active devices. The device k observing $X_k = m$ has a transmitted signal given by $s_k(t) = \alpha_k \phi^m(t + D_k)$ where α_k is the complex compensation factor for device k and D_k is the transmission delay introduced by device k.

All devices' signals are allocated on the same system bandwidth B and transmitted through a wireless fading channel. Doubly block-fading is assumed due to (*i*) the system is narrow-banded, and (*ii*) both communication sides are static (no movement). In other words, the channel is invariant in time and frequency during the transmission interval. Hence, the channel affecting device k during a transmission period can be modelled as a one-tap channel $h_k(t) = h_k\delta(t)$ where $h_k \sim \mathcal{CN}(0,1)$ is Rayleigh fading i.i.d. across all active devices. The channel coefficient h_k is also expressed as $h_k = |h_k|e^{j\phi_k}$ for later convenience. The received signal at the fusion-center is given by

$$y(t) = z(t) + w(t) = \sum_{k\in\mathcal{K}} h_k s_k(t - t_{TA_k}) + w(t) = \sum_{k\in\mathcal{K}} h_k \alpha_k \phi^m(t + D_k - t_{TA_k}) + w(t), \quad (3.1)$$

where $z(t)$ is the sum of the K received signals, t_{TA_k} is the propagation delay of device k, and $w(t) \sim \mathcal{CN}(0, \sigma_w^2)$ is i.i.d complex Gaussian noise with zero mean and variance σ_w^2. The noise variance is given by the system Signal to Noise Ratio (SNR), which is defined as

$$SNR = \frac{P_{z(t)}}{\sigma_w^2}, \quad (3.2)$$

where $P_{z(t)}$ is the average power of $z(t)$ on interval $0 \leq t \leq T$.

3.3 Communication protocol

The wireless system follows an information-centric TBMA-based protocol. Unlike traditional device-centric protocols where the objective is to obtain devices' individual payload, on TBMA the objective is to obtain information about the QoI. Therefore, devices do not include an identifier on the transmitted signal since the objective is to obtain information of the QoI regardless the identity of the device.

As previously stated, the devices transmit orthogonal preambles which encode the observed value $m \in \{1, 2, ..., M\}$. Devices observing value m will transmit preamble $\phi^m(t)$. Therefore, preambles are used in a non-orthogonal fashion by devices observing the same value. With the received signals, the fusion-center estimates the value of the QoI, i.e. θ, which conditioned the measurements made by the devices. The transmitted information about the QoI is given by the empirical measure, that is, the histogram of the measurements taken by the K devices

$$\tilde{\boldsymbol{p}} = \frac{1}{K}(K_1, K_2, ..., K_M), \quad (3.3)$$

where $K_m = \sum_{k\in\mathcal{K}} 1(X_k = m)$ is the number of devices that observe m. At the fusion-center the estimated empirical measure is denoted by $\hat{\tilde{\boldsymbol{p}}}$ and is computed as

$$\hat{\tilde{\boldsymbol{p}}} = \frac{1}{\hat{K}}(R_{\phi^1}, R_{\phi^2}, ..., R_{\phi^M}), \quad (3.4)$$

where R_{ϕ^m} is the correlation of preamble $\phi^m(t)$ with the received signal and $\hat{K} = \sum_{m\in\mathcal{M}} R_{\phi^m}$. The correlation of each preamble with the received signal is computed as

$$R_{\phi^m} = \int_{t=0}^{T} \phi^m(t) \cdot y(t)^* dt, \tag{3.5}$$

where $y(t)^*$ represents the complex conjugate of the received signal. Note that if $y(t) = \phi^m(t)$, then $R_{\phi^m} = 1$. On the first TBMA approaches [12], [13], the fundamental TBMA model considered UL time synchronization ($D_k = t_{TA_k} = 0$) and channel coefficients equal to one or equalised channel ($h_k\alpha_k = 1$). Under these assumptions, the received signal is expressed as

$$y(t) = \sum_{m\in\mathcal{M}} K_m\phi^m(t) + w(t), \tag{3.6}$$

which under $\phi^m(t)$ orthogonal preambles and absence of noise, it is found that $R_{\phi^m} = K_m$. In other words, the estimated empirical measure contains the number of devices that observed each m-th value. However, due to the nature of the wireless channels, the channel gains cannot be considered equal to one, devices have different propagation delays and noise is inevitable. Those factors can lead to an estimated empirical measure different from the original (i.e. $\tilde{\boldsymbol{p}} \neq \hat{\tilde{\boldsymbol{p}}}$). Therefore, two main factors affect on the estimation of the empirical measure:

- Non-coherent addition of signals due to the channel coefficients
- Time unsynchronization of received preambles due to different propagation delays (also leading to non-coherent superposition).

Estimation of θ

The family of distributions p_{θ_q} is known at the fusion-center. In order to infer the value of the parameter θ, the fusion-center computes the f-divergence between estimated empirical measure and original observations' distribution. Numerable f-divergence metrics exists on statistics analysis [23]. Most well known f-divergences are the Kullback–Leibler divergence or the ξ^2-divergence among others. However, it must be taken into account that the empirical measure might take probability zero for some preambles. Hence, the most appropriate f-divergence is the Total Variation (TV). The total variation between two probability distributions a and b is defined as

$$D_{TV}(a, b) = 0.5 \sum_{m\in\mathcal{M}} |a_m - b_m|,$$

where a_m and b_m are respectively the m-th element of the distributions a and b. Note that $0 \leq D_{TV} \leq 1$, being $D_{TV} = 0$ when a and b are the same distribution. In order to infer the parameter θ, the fusion-center compares the estimated empirical measure $\hat{\tilde{\boldsymbol{p}}}$ with each of the distributions p_{θ_q}. Hence, by means of the total variation between the estimated empirical measure $\hat{\boldsymbol{p}}$ and each distribution p_{θ_q}, we infer the field parameter as

$$\hat{\theta} = \arg\min_{\theta_q}\{D_{TV}(p_{\theta_q}, \hat{\boldsymbol{p}})\},$$

where $\hat{\theta}$ is the estimate of θ. The decision rule takes into account all states θ_q have the same probability of occurrence p_q.

3.4 Performance metrics

The system performance will be evaluated by means of the error probability P_e defined by

$$P_e = Pr[\hat{\theta} \neq \theta], \tag{3.7}$$

where θ is the parameter that describes the QoI and $\hat{\theta}$ is its estimated value at the fusion-center. The computation of $\hat{\theta}$ (explained on previous section) is influenced by the number of active devices K on the system and the estimated empirical measure $\hat{\boldsymbol{p}}$. The number of active devices has a direct impact on the quantity of information reported about the QoI. Larger number of active devices facilitates the reconstruction of p_{θ_q}. For example, if $K = 1$ the empirical measure would be just one value with barely information about the QoI whereas with $K = 300$ the distribution has more samples (in fact, 300) to represent the original p_{θ_q}. The other parameter of influence, i.e. the estimated empirical measure, is affected to several parameters from the system: the SNR, the number of preambles M, the number of states Q of the QoI and the channel propagation (h_k and t_{TA_k}). Nevertheless, note that the effect of the wrong estimation of the empirical measure on the error probability is more severe when K takes short values.

3.5 Transmission schemes

Three transmissions schemes are designed in order to give a variety of TBMA-based implementations. Each transmission scheme (shortened as TS) incorporates modifications of parameters of the signal model that lead to different error probability performances. Table 3.1 summarizes each of the schemes and their implications on the system model. The advantages and disadvantages of each transmission scheme together with their implementation aspects are discussed on Chapter 5.

3.5.1 TS1 - Channel and TA compensation

The channel and TA compensation scheme, shortened as TS1, is determined by two facts

- All devices compensate the phase and amplitude of their channel h_k. For compensating the channel, the compensation factor of device k is set to $\alpha_k = \frac{1}{h_k} = \frac{1}{|h_k|e^{j\phi_k}}$.

Table 3.1: Summary of transmission schemes

TS abbreviation	**TS fundamentals**	**Channel compen-sated?**	α_k	**TA compen-sated?**	D_k	**Received signal**
TS1	Channel and TA compensation	Yes	$\frac{1}{\lvert h_k \rvert e^{j\phi_k}}$	Yes	t_{TA_k}	$y(t) = \sum_{m \in \mathcal{M}} K_m \phi^m(t) + w(t)$
TS2	Channel phase and TA compensation	Only phase	$\frac{1}{e^{j\phi_k}}$	Yes	t_{TA_k}	$y(t) = \sum_{k \in \mathcal{K}} \lvert h_k \rvert \phi^m(t) + w(t)$
TS3	Uncontrolled transmission	No	1	No	0	$y(t) = \sum_{k \in \mathcal{K}} h_k \phi^m(t - t_{TA_k}) + w(t)$

- All devices modify their transmission delay D_k to compensate their TA. To do so, it is set $D_k = t_{TA_k}$. In this way, the propagation delay is counteracted and all transmitted signals are UL timely synchronized.

Incorporating this configuration to the system model achieves integer forcing of the transmitted preambles at the receiver. This means that the received signal is an integer combination of codewords $\phi^m(t)$

$$y(t) = \sum_{k \in \mathcal{K}} h_k \frac{1}{h_k} \phi^m(t + t_{TA_k} - t_{TA_k}) + w(t) = \sum_{m \in \mathcal{M}} K_m \phi^m(t) + w(t), \tag{3.8}$$

where K_m is the number of devices observing m. As observed on Equation 3.8, the received signal is a integer combination of preambles with noise addition. This transmission scheme can be seen as an equivalent to Additive White Gaussian Noise (AWGN) channel with all devices having same TA. Note that in absence of noise and having orthogonal preambles it is obtained $R_{\phi^m} = K_m$.

3.5.2 TS2 - Channel phase and TA compensation

The channel phase and TA compensation scheme, shortened as TS2, is similar to TS1 but with a slight difference. In TS2 only the phase of the channel is compensated in comparison to TS1 where both amplitude and phase are compensated. This transmission scheme is relevant for devices observing deep fading events since the channel amplitude compensation requires high power consumption. TS2 includes the two following facts:

- All devices compensate the phase of their channel h_k. For compensating only the phase, the compensation factor of device k is set to $\alpha_k = \frac{1}{e^{j\phi_k}}$.
- All devices modify their transmission delay D_k to compensate their TA. To do so, it is set $D_k = t_{TA_k}$. In this way, the propagation delay is counteracted and all transmitted signals are UL timely synchronized.

By equalising the phase of the channel, coherent superposition of received preambles is achieved. However, fading is not compensated and therefore transmitted preambles are scaled by the amplitude of the channel gain $|h_k|$. Equation 3.9 shows the received signal when TS2 is applied.

$$y(t) = \sum_{k \in \mathcal{K}} |h_k| e^{j\phi_k} \frac{1}{e^{j\phi_k}} \phi^m(t + t_{TA_k} - t_{TA_k}) + w(t) = \sum_{k \in \mathcal{K}} |h_k| \phi^m(t) + w(t), \tag{3.9}$$

3.5.3 TS3 - Uncontrolled transmission

Shortened as TS3, this scheme includes devices which do not compensate neither channel nor TA. It is referred as an uncontrolled transmission because the devices do not modify the amplitude or delay of the signal, therefore not controlling the effects of radio propagation.

- The channel is not compensated, therefore being $\alpha_k = 1$ for all devices.
- The transmission delay is set to $D_k = 0$, therefore signals do not achieve time alignment at the receiver and the arrival depends on the propagation delay of each device.

By the addition of these parameters, the received signal turns into

$$y(t) = \sum_{k \in \mathcal{K}} h_k \phi^m(t - t_{TA_k}) + w(t), \tag{3.10}$$

It can be demonstrated that if the channel is not equalised but devices counteract their TA, the system performance is the same because the superposition of the signals is non-coherent in both ways whether or not received preambles achieve UL time synchronization.

3.6 Adaptation to Narrowband Internet of Things

In order to see how close the system model can approximate to existing or future radio technologies, hereby we adapt the system model to a real wireless technology. Note that the adaptation focuses on the physical layer structure, although the use of some features from upper layers might be also considered (e.g. clustering). The election of the technology is justified by the following scenario characteristics: (*i*) large number of devices, (*ii*) with energy-constraints, and (*iii*) narrowband signals. Therefore, from the physical layer point of view, a very appropriate technology is NB-IoT.

NB-IoT has a signal bandwidth of $B = 180$ KHz. On UL, two physical channel are found: NPUSCH and NPRACH. We propose the use of NPRACH and specially of the NPRACH preambles. On the normal functioning of this physical channel, devices start RA procedures with the transmission of preambles on a given format specified by the NPRACH configuration. If the NPRACH preambles are used with a TBMA purpose, it only implicates another interpretation of the received preambles.

The NPRACH preambles are usually organised on sets of $M = 12, 24, 36$ or 48. For this work, we consider any M integer value up to 48 preambles, i.e. $M \in \{1, 2, ..., 48\}$. Taken the NPRACH preambles structure described on Section 2.2, the system codebook is given by a set of M orthogonal codewords $\boldsymbol{\phi} = [\phi^1(t), \phi^2(t), ..., \phi^M(t)]$ shared by the N devices of the system. Each m-th codeword (or preamble) has the following expression

$$\phi^m(t) = \frac{1}{\sqrt{T}} \sum_{i=0}^{N_{SG}-1} \phi_i^m(t_i), \tag{3.11}$$

which is the sum of N_{SG} symbol groups where $0 \leq t \leq T$ being $T = N_{SG} T_{SG}$, and $\phi_i^m(t_i)$ is the i-th symbol group from m-th preamble which is defined as

$$\phi_i^m(t_i) = e^{j2\pi(n_{SC}^{RA}(i) - N_{SC}^{UL} 0.5k_0 + 0.5)\Delta f_{RA}(t_i - T_{CP})}, \tag{3.12}$$

where $T_{SG}\, i \leq t_i \leq T_{SG}(i+1)$ determines the time instances where the i-th symbol group is defined, $T_{SG} = T_{CP} + T_{SEQ}$ is the duration a symbol group, and $n_{SC}^{RA}(0) = m$ is the initial subcarrier index which also affects subcarriers $n_{SC}^{RA}(i)$ for $i > 0$ according to the rule explained on Section 2.2. Rest of the parameters definition can be found in Section 2.2.

Even though the system model does not include clustering of devices, the NB-IoT CE-based grouping is considered and its implications on the results and on the TBMA protocol will be further discussed on Chapter 5.

4

Performance evaluation

On this Chapter, we discuss the performance of the different TBMA transmission schemes introduced on the previous chapter. We analyse the influence of TBMA parameters such as number of preambles or number of states. The results are generated through Monte Carlo simulations with 10000 independent realizations. For the regime of the results $P_e < 10^{-2}$, we consider that the accuracy of the results is sufficient. The simulations have been conducted with MATLAB R2019b. The performance will be given in function of the number of active devices K.

4.1 Considered scenarios

The simulation parameters are summarized on Table 4.1. The system SNR is set to 6 dB. All devices have the same NPRACH preambles configuration: $N_{rep} = 8$ and preambles with format 1, i.e. $T_{CP} = 267\ \mu s$ and $T_{SEQ} = 1.333\ ms$. The use of preambles with format 1 allows devices located up to 32 km far away from the BS. Devices are physically uniformly distributed on a cell radius of 32 km. On an actual NB-IoT system, this NPRACH configuration would be equivalent to devices having same CE level. The implication of having simultaneously different NPRACH configurations will be further discussed on Section 4.3. By analysing different system parameters, we aim at studying what are the minimum system requirements to achieve a target probability of error. For this work, we will consider a error probability objective of $P_e = 10^{-2}$.

4.1.1 Scenario A - Binary-state QoI

The following described scenario is labelled as Scenario A. We consider a binary state of nature, i.e. the QoI can take two values/states $\theta \in \{\theta_1, \theta_2\}$. Both states have equal probability of occurrence $Pr[\theta = \theta_1] = Pr[\theta = \theta_2] = 0.5$. The observations X_k can take M values and have a conditional probability distribution given by the distributions:

Table 4.1: Simulation parameters

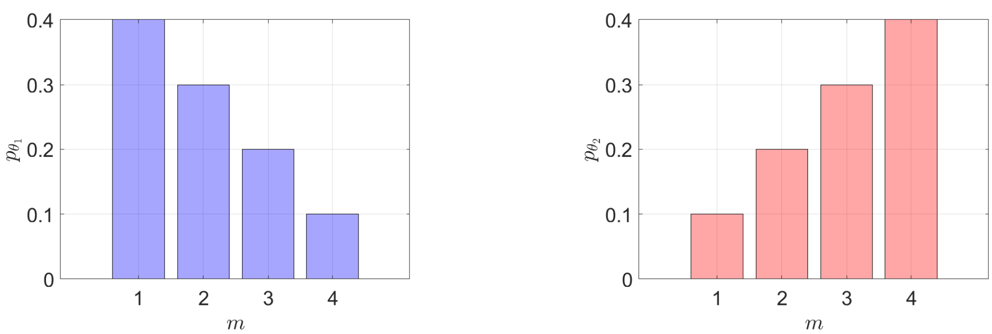
Parameter	Value
SNR	6 dB
N_{rep}	8
NPRACH preamble format	1
T_{CP}	267 μs
T_{SEQ}	1.333 ms
Cell range	32 km

$$p_{\theta_1}(m) = p_M(M - (m-1)) \qquad \text{if } \theta = \theta_1 \tag{4.1}$$

$$p_{\theta_2}(m) = p_M\, m \qquad \text{if } \theta = \theta_2 \tag{4.2}$$

where $p_M = \frac{1}{\sum_{m \in \mathcal{M}} m}$. This generates two distributions of descendant and ascendant probabilities. Note that the distributions corresponds to the quantized observations and could represent a low temperature or pollution state for $\theta = \theta_1$ (shorter m are more likely) and a high temperature or pollution state for $\theta = \theta_2$ (larger m are more likely). To provide a more intuitive understanding, Figure 4.1 illustrates the two observations' distributions on Scenario A for $M = 4$. Appendix A contains more illustrative distributions for other M values. In this case, θ can represent the mean of the observed values. Then for $M = 4$, the field parameter could take two values $\theta \in \{2, 3\}$. Another interpretation is just considering two states θ_1 and θ_2 as the tendency to measure low or high values of a physical parameter.

The value of M has a direct impact on the distributions' observations. Large M values increase the number of quantization levels so that the discrete m values are a closer approximation of the original real-valued measurements. The consequence of increasing M is an improvement of the P_e due to the increase of transmitted information. This turns M as a



(a) $\theta = \theta_1$ **(b)** $\theta = \theta_2$

Figure 4.1: Scenario A observations' distributions for $M = 4$

relevant system parameter which can be tuned to achieve a certain P_e application-objective within a certain number of devices. However, there is a trade-off between the limited number the preambles available on the system and the target P_e. On an optimal implementation, we want to minimize the number of resources (preambles) and achieve the lowest error probability.

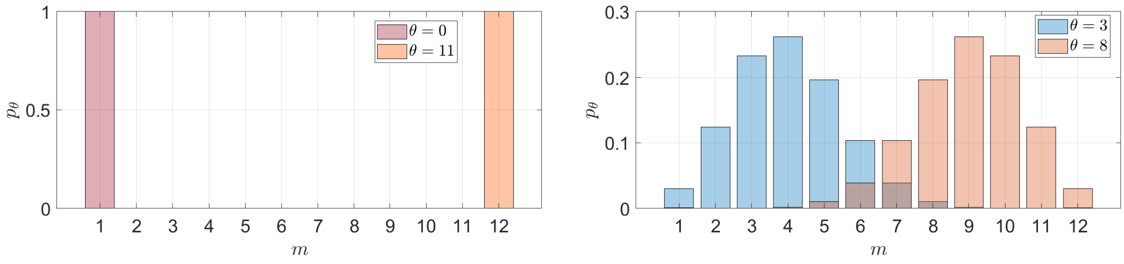
4.1.2 Scenario B - Multi-state QoI

The following described scenario is labelled as Scenario B. The QoI can take Q states, i.e. $\theta \in \{\theta_1, \theta_2, ..., \theta_Q\}$. All states have the same probability of occurrence $Pr[\theta = \theta_q] = \frac{1}{Q}$. The observations X_k made by the devices can take M possible values. For this scenario we consider $M = 12$ although it can be generalised for any M. The devices' observations are distributed according to

$$p_{\theta_q} \sim 1 + \mathcal{B}(M-1, \frac{\theta_q}{M-1}), \tag{4.3}$$

where $\mathcal{B}(n,p)$ is a binomial r.v. with number of trials $n = M-1$ and probability of success $p = \frac{\theta_q}{M-1}$. The addition of "1" to the binomial distribution is to make p_{θ_q} be defined on $m \in \{1, 2, ..., M\}$. Note that our parameter of interest $\theta_q = p(M-1)$ defines the mean of the binomial distribution and can take any real value on the interval $[0, M-1]$ (although we only consider Q possible values). The selection of values θ_q is an interesting point and might lead to an extent discussion. The following lines include a brief discussion of how to choose θ_q in order to fairly evaluate the error probability performance in function of Q. For a more intuitive understanding, let us interpret the QoI as the ambient temperature. In case $Q = 2$ states are considered, one possibility is to take the most extreme states, i.e. $\theta \in \{0, M-1\}$ but this would correspond with states of very low or very high measured values with generally low probability of occurrence. Instead, for $Q = 2$ medium states are the most reasonable choice to transmit relevant information. For instance, an acceptable choice would be $\theta \in \{3, 8\}$. The two probability distributions of the discussed $Q = 2$ cases are illustrated on Figure 4.2. As observed, by selecting inter-medium states all measured values are covered an relevant information about the event will be acquired. Note also that with $\theta \in \{0, M-1\}$ there should not be misdetection between states since each observed value can only be mapped to one possible state.

The distributions p_{θ_q} correspond to the discrete measurements X_k. Be aware that the original measurements are usually real-valued and have a different probability distribution. On distributed estimation, it is common to model the original measurements as Gaussian distributed with certain mean and variance, where the mean is usually the parameter of interest (see [11] among others). The quantization of the sensor's measurements and its implication on the system performance is out of the scope of this project. In general, the Gaussian distribution is quantized and the resulting distribution of the quantized observations is a non-symmetric quantized version of the Gaussian distribution. We consider that the binomial distribution might approximate to the quantized observations. The advantage of the defined distribution



(a) $\theta \in \{0, 11\}$

(b) $\theta \in \{3, 8\}$

Figure 4.2: Two cases for Scenario B observations' distributions with $Q = 2$

on Equation 4.3 is the possibility to tune M and Q. Appendix A illustrates all the considered distributions for Scenario B.

One of the objective of Scenario B is to see the influence of the number of states Q on the system performance. This is also a trade-off between how accurate is the information of the QoI (the larger Q, the more precise information) and the required error probability. Think that including more states Q implicates that the estimated empirical measure at the receiver has to be accurate enough to differentiate between one state or another. Being $M = 12$, we consider the following Q cases

$$Q = 2,\ \theta \in \{3, 8\} \tag{4.4}$$
$$Q = 4,\ \theta \in \{1, 3, 8, 10\} \tag{4.5}$$
$$Q = 6,\ \theta \in \{1, 3, 5, 6, 8, 10\} \tag{4.6}$$
$$Q = 12,\ \theta \in \{0, 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11\} \tag{4.7}$$

4.2 Results

4.2.1 Scenario A

The results when utilising TS1 are shown in Figure 4.3. The evaluated configuration of preambles are $M = 2, 4, 12$ and 24. The largest number of preambles ($M = 24$) provides the best performance among the shown configurations. This proves that increasing the quantization levels, and therefore the number of preambles that encode the quantized values, provides more information about the QoI and improves the estimation performance. Nonetheless, on the regime of $P_e = 10^{-2}$ this improvement is more significant for short than for large number of preamble. For instance, with $K = 16$ the error probability for $M = 2$ and $M = 4$ is respectively 0.02 and 0.084, while for $M = 12$ is 0.097 showing a small improvement with respect to $M = 4$. Even $M = 24$ provides almost same performance compared to $M = 12$ at $K = 16$. The reason is that the improvement obtained by $M = 12$ and $M = 24$ is notorious for the regime $P_e < 10^{-2}$ showing apparent gains in terms of error probability on that regime. Hence,

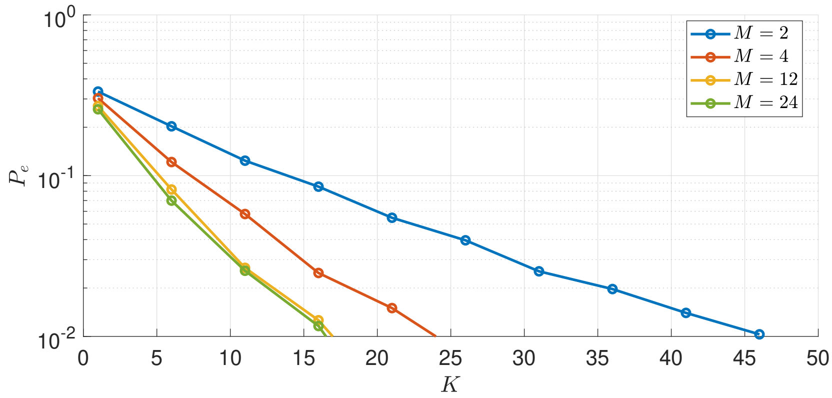


Figure 4.3: Error probability on Scenario A for TS1

short number of preambles are sufficient to achieve the objective error probability when the channel and TA are compensated.

Figure 4.4 shows the P_e performance for TS2. Similar trends to the performance of TS1 is observed although reaching same P_e values for larger K. As introduced on Section 3.5, TS2 has similar configuration to TS1 but without channel amplitude compensation. The empirical measure estimation is affected by the amplitude of the channel which increases the error probability with respect to TS1. Comparing with the results of TS1 (Figure 4.3), using $M = 2$ the $P_e = 10^{-2}$ is reached with a minimum of $K = 46$ for TS1 and $K = 58$ for TS2. The difference is not significant and supports the use of TS2 due to its advantages regarding energy savings. Furthermore, with $M = 12$ and $M = 24$ the worsening of TS2 with respect to TS1 is less significant on the regime of $P_e = 10^{-2}$.

With respect to the last of the transmission schemes, Figure 4.5 shows the estimation performance for TS3. On a first sight, readers might notice that only large number of preambles

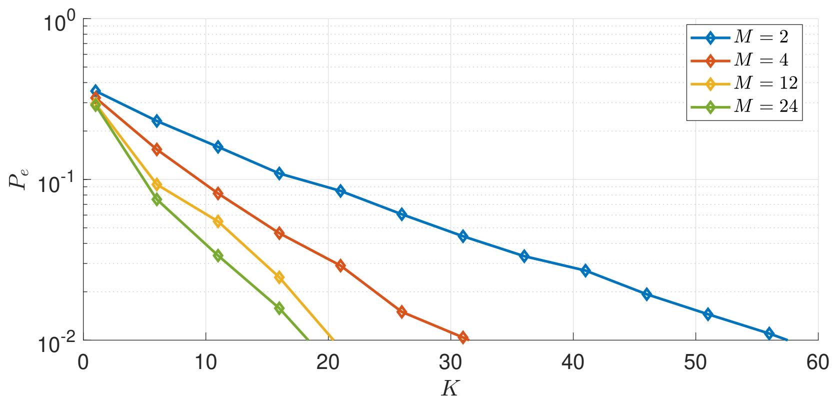


Figure 4.4: Error probability on Scenario A for TS2

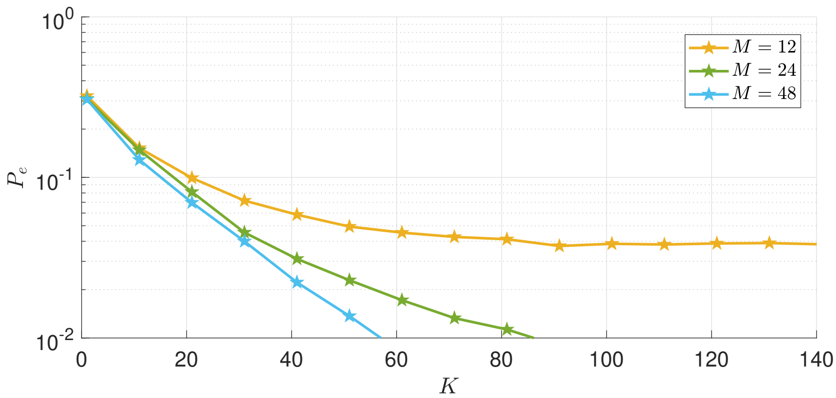


Figure 4.5: Error probability on Scenario A for TS3

($M = 12, 24$ and 48) are evaluated. The reason is that the channel and the propagation delays have a severe effect on the P_e performance therefore requiring large number of preambles to achieve proper results. The error probability saturates for $K > 70$ in the $M = 12$ case. The fact is that with uncontrolled transmissions, no matter how K increases, at some point the non-coherent addition of signals limits the improvement of the QoI estimation. This saturation happens in the same manner for $M = 24$ and $M = 48$ although on the lower regime of P_e.

To sum up, the main conclusions of the analysis of Scenario A are the followings. The use of TS1 and TS2 have provided the best performance showing the advantages of including channel and TA compensation strategies on the communication scheme. Even though TS2 does not incorporate the compensation of the amplitude of the channel, it has shown close approximation to the performance of TS1. TS1 and TS2 have shown a considerably good performance for low number of preambles and the use of large quantity of preambles is recommended to achieve very low error probabilities ($P_e < 10^{-2}$). TS3 has shown the severe effect of non-coherent superposition of signals requiring configurations with up to 48 preambles. Additionally on TS3, saturation of the error probability has been observed after a certain number of active devices.

4.2.2 Scenario B

The election of the number of states Q is a trade-off between increasing the information about the QoI at cost of worse P_e performance. Figure 4.6 presents the error probability performance with different Q configurations for TS1, TS2 and TS3.

Starting with TS1, on Figure 4.6 we observe that $Q = 6$ outperforms $Q = 12$ because larger number of states implicates a decrease of the estimation performance. Note that comparing with Scenario A ($M = 12$ on Figure 4.3), $P_e = 10^{-2}$ is reached with only $K = 17$. For that case, a binary QoI ($Q = 2$) was taken into account. By increasing the number of states to $Q = 6$ or $Q = 12$, the minimum required K to achieve the P_e objective is $K = 53$ and $K = 74$ respectively. Using TS2, the performance has similar trends to TS1 but with a slightly decrease of performance caused by the non-compensated channel fading. Between $Q = 6$ and $Q = 12$ curves, TS1 has an offset of around 20 devices, while for TS2 curves differ in almost 30 devices. For instance, at $K = 50$ and $Q = 6$, TS1 achieves $P_e = 0.0115$ and TS2 achieves $P_e = 0.0172$ which is an insignificant difference. Similarly happens also for $Q = 12$. Hence, due to the implementation advantages of TS2, we support the use of TS2 over TS1.

The uncontrolled transmissions of TS3 comes at cost of worsening of the P_e performance as observed on Figure 4.6. This is already observed looking at the Q values evaluated in comparison with TS1 and TS2. While on the other transmission schemes acceptable performance is obtained with larger number of states, TS3 rapidly decreases the performance with short number of states. With $Q = 2$, the target P_e is achieved with only 14 devices. However, with the increase to $Q = 4$, the estimation performance saturates around $P_e = 0.04$ not showing any improvement with the increase of K. The reason is that with $Q = 2$, the measured values m are quite differentiated for each state while $Q = 4$ increases the difficulty on the estimation task (compare the probability distributions of Figure A.3a and Figure A.3b). An interesting appreciation is the difference on error performance for $Q = 2$ between Scenario A and Scenario B. In case of $Q = 2$ and $M = 12$, both scenarios are at the same comparison level. We have only evaluated this case for TS3. Under these circumstances, we observe a different performance comparing $M = 12$ case in Figure 4.5 (remember that Scenario A has fixed $Q = 2$) and TS3 $Q = 2$ case in Figure 4.6 (remember that Scenario B has fixed $M = 12$). While Scenario B reaches the $P_e = 10^{-2}$ with short K, Scenario A saturates around $P_e = 0.04$ and does not reach the target error probability. Here, the differentiation factor is clearly the definition of the probability distributions. The two compared cases are illustrated by probability distributions from Figure A.1c (Scenario A) and Figure A.3a (Scenario B). As observed on Scenario A, all m values are covered by significant probability for both states whereas on Scenario B both states are quite differentiated by the m values of each state.

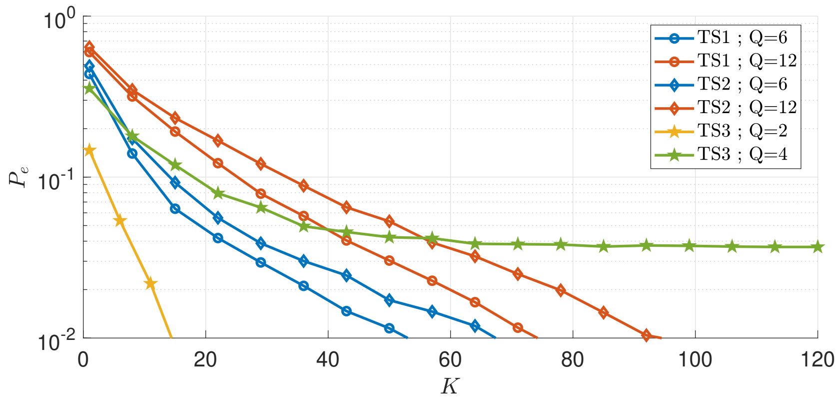


Figure 4.6: Error probability on Scenario B for TS1, TS2 and TS3

4.3 Discussion and enhancements

The results from above show the positive impact on the error probability of increasing M and decreasing Q. Moreover, the increase of K also yields to an improvement on the error performance, so the results show the minimum required K to achieve a given error probability. The transmission schemes have shown distinctive estimation performance being TS1 the best and TS3 the worst in terms of error probability. TS2 provides results on the same order of magnitude than TS1. In cases where devices are located in small areas close to each other the performance of TS2 might approximate more to TS1. In those cases, the channel amplitude $|h_k|$ could be correlated among close devices. If all devices have similar $|h_k|$, all received preambles would be affected by the same factor. Due to the normalization factor $\hat{K}$ on Equation (3.4), if all devices are affected by same $|h_k|$, the channel coefficient is normalised and the estimated empirical measure on TS2 approximates to TS1. The main advantage of using TS2 is the reduction of the power consumption by not compensating the channel amplitude. Moreover, the full channel amplitude compensation (as TS1 does) is not always feasible due to power constraints. A further discussion about the implications of each transmission schemes can be found on Chapter 5.

Readers might hesitate if the results presented on this chapter correspond to realistic massive connectivity scenarios. The point here is how many devices are required to consider an scenario as massive. With this regard we point to 3GPP TR 37.868 [4] where mMTC traffic models are proposed to determine typical values of active devices on MTC scenarios. According to this document, typical number of connected devices are $N \in \{1000, 3000, 5000, 10000, 30000\}$. From the N connected devices, the number of K active devices is determined by the traffic model. Two traffic models are proposed. Both consider that in a given period T_a all devices are active (have UL data to transmit). For the activity pattern, Model 1 makes use of an uniform distribution with $T_a = 60\ s$ and Model 2 utilises a beta distribution with $T_a = 10\ s$. Model 1 represents scenarios where devices are active uniformly over a period of time, i.e. in a non-synchronized manner. Model 2 represents an extreme scenario where large amount of MTC devices access the network in a highly synchronized manner, e.g. after a power outage [4]. As an exemplary scenario, we consider the periodicity of transmissions (time between transmission intervals, i.e. time in which the devices can become active) to be every 1280 ms. This is a typical value for the NPRACH periodicity [18]. For all the N considered above, the number of active devices is $K \in \{21, 64, 106, 213, 640\}$ for Model 1 and $K \in \{31, 93, 155, 310, 931\}$ for Model 2. Note that results depend on the periodicity of transmission and T_a. The error probability results shown in Section 4.2 comprise K values which might be below some of the K typical values determined above. Nevertheless, the error probability decreases or remains the same for larger K under fixed SNR. Hence, the minimum error probability performance can be extrapolated for larger K.

The results from Section 4.2 have been generated under a fixed SNR of 6 dB. Other SNR results have been dismissed due to the low influence on the P_e performance. The low impact of the SNR on the results has two prevalent reasons. Firstly, the good orthogonality properties of the preambles provokes that at the output of the correlator an SNR gain is obtained, therefore increasing the noise robustness of the system. Secondly, the SNR affects to the empirical measure estimation which has only a greater influence on the short regime of K. Nonetheless, the SNR has a significant impact for $SNR < -20$ dB, provoking a notorious degradation on the estimation of θ. This robustness is another advantage to consider TBMA-based protocols on massive connectivity scenarios because the low impact of the SNR on the system performance could be used to reduce the transmitted power from the devices. This fact is a great advantage for mMTC scenarios where devices are commonly energy-constrained.

Another point to think over is the effect of CE-based clustering on the system performance. Take into account that the set-up scenario includes devices transmitting with same N_{rep}, so equivalently, all devices belong to same CE-based group (see Section 4.1). Following the preambles structure from NB-IoT (see Section 2.2), N_{rep} modifies the length of the preamble. In order to make possible the use of preambles with different length, the set or codebook of 48 NPRACH preambles has to be partitioned on orthogonal codebooks for each CE-group (i.e. for each N_{rep}). The N_{rep} value for each group is normally set up by the BS and its main purpose is to enhance the coverage of devices on a group. If each group is assigned a different set of preambles, then having more than one group has no effect on the performance metrics.

In general, TS1 and TS2 provide better performance than TS3 under the same M and Q configurations. On the contrary, as discussed on the next Chapter, TS3 is simpler scheme because it does not have power constraints or does not require the devices' knowledge of their TA. Motivated by the practical advantages of TS3, alternative receiver configurations should be taken into account to improve the performance when using TS3. With such purpose, we propose the use of machine learning algorithms. The application of machine learning techniques is specially interesting for configurations like TS3 where the channel is not equalised and/or the TA is not compensated. Machine learning can be used either to estimate the empirical measure or to directly infer the parameter θ. The uncertainty variables are the channel coefficients of each device (h_k) and their TA (t_{TA_k}). The training of the machine learning algorithm implies a long-term learning process where prior information is gathered in order to help on the estimation process:

- Likely number of active devices
- Likely TA of active devices
- Likely observed values.
- Likely channel coefficients for active devices

For instance, if devices are measuring the temperature at night on winter period, the fusion-center does not expect to receive measures of high temperatures. In this way, a previous learning process could help on the estimation process. Additionally, the learning process is facilitated by the static conditions of the devices, therefore persevering the same propagation characteristics. If the fusion-center can compute sufficient statistics of devices' transmissions, machine learning could improve the estimation performance. Hence, we consider machine learning-assisted solutions as a potential future research line on TBMA systems.

4.3.1 TBMA vs local estimation approaches

TBMA has been studied as a communication protocol to enable centralized estimation of a field parameter. On the distributed estimation paradigm, local estimation is also a possible solution for the parameter estimation problem. Although existing work has considered TBMA as a communication protocol to encode information about local decisions (see [24]), we consider TBMA to be used on estimation approaches involving the transmission of quantized observations (centralized estimation-based).

Local estimation solutions using NB-IoT preambles raise as an interesting alternative to TBMA. The physical layer could be used on the same way but with different mapping from observed values to transmitted preambles and different interpretation of received preambles at the fusion-center. While in TBMA the preambles encode the measured values, in a local estimation approach the preambles would encode the hypothesis that is locally decided or soft-information about the local estimate. Note that the soft-information must be quantized into M quantization levels, just us that observed values on TBMA. The quantization problem for either observed values or soft-information of the local estimate has a relevant impact on the transmission of information and should be considered under a implementation point of view. The choice of local or centralized strategies is up to debate. Local estimation might be a good choice in scenarios with a binary hypothesis QoI. However, devices need to be *intelligent* so that they have knowledge of the observations distributions to be able to conduct local decisions. In that case, only two preambles are needed to encode the local decision. In case additional information aims at be transmitted (for instance quantized Log-Likelihood Ratio (LLR)), more preambles would be required. Furthermore, when more than two states are defined ($Q > 2$), transmitting local estimation information would require from more preambles.

Authors in [9] compared centralized and local detection over multiple-access channels assuming that the received preambles are added in a coherent fashion (AWGN channel). Under individual power constraints and discrete observed data, histogram fusion (a.k.a TBMA) and soft-information showed better detection performance than hard-information (one bit local decision quantization). Transmission of hard-information provided good results only when the SNR is sufficiently low. However, TBMA and soft-information strategies dominates over hard-information for number of devices sufficiently large. On the context of Cloud-Radio Access Networks (C-RAN) the use of local and centralized detection has been evaluated in [10]. On this work both approaches are assessed motivated by the optimization of resources on fronthaul links. However, the estimation problem is evaluated at the fronthaul link being the sensors' observed values directly transmitted through non-orthogonal signatures without quantization at devices. The effect of the quantization of measured values or local estimates is not evaluated over the radio interface which is the one that would carry the NB-IoT preambles that encode the measured values or local estimates.

As a final remark we want to emphasize that election of one strategy or another is not a trivial decision. The optimal choice is directly influenced by the quantization function of either the observed values or the soft-information of the local estimate. Therefore, future work should address this optimization problem considering the use of NB-IoT physical layer.

5

Implementation aspects

A TBMA-based protocol has been proposed based on the system model described in Chapter 3. Its performance has been discussed on Chapter 4 and the influence of different system parameters have been analysed. On this work, we also analyse how the defined model and the taken assumptions come into truth on a real wireless system and to which extent the proposed approach can be really implemented.

An appropriate evaluation direction is to choose an already existing technologies and analyse what changes or implications have to be considered to develop the proposed TBMA protocol. Even though current technologies might not completely fulfill the functioning of a TBMA-based system, the efforts involved on modifying or including enhancements on next standard releases are more feasible than developing a completely new technology for TBMA. Nonetheless, the implementation aspects should be considered in the same manner if TBMA aims at being implemented on any other technology.

NB-IoT is found as the most suitable candidate not only due to its characteristics for massive connectivity but also due to the standardisation efforts as a 5G technology [7]. Hence, the following analysis will focus on adapting TBMA to future releases of NB-IoT and to suggest physical and access layers features of TBMA-based systems for massive connectivity.

5.1 Scenario and wireless channel

As part of the massive connectivity paradigm, this work targets applications which involve a large number of devices on a wireless system. On massive connectivity scenarios, it is common to have devices with null or reduced mobility (devices are usually embedded on physical surfaces, e.g. temperature meters on pipelines) [6]. Due to sporadic transmission patterns, mMTC communication protocols are usually designed to fulfill the requirements under the condition of static devices, not including mobility features when the devices are active (e.g. NB-IoT only conducts handovers between BSs if the device is not transmitting, i.e. IDLE state). With these arguments, we consider a fair assumption the fact that the devices are static while being active.

Table 5.1: Implementation aspects of the transmission schemes considered on this work

TS abbreviation	**TS fundamental**	**TDD system and UL reciprocity required**	**Sensitive to high power consumption**	**Previous TA learning process required**
TS1	Channel and TA compensation	✓	✓	✓
TS2	Channel phase and TA compensation	✓	✗	✓
TS3	Uncontrolled transmission	✗	✗	✗

The system model assumes the use of narrowband signals (i.e. the signal bandwidth is much larger than the coherence bandwidth of the channel). This is achieved by NB-IoT with only 180 KHz bandwidth. Furthermore, NPRACH signals extent only to 45 KHz (see Section 2.2). Considering a frequency correlation coefficient of 0.5, the coherence bandwidth of a channel is usually approximated as $B_c \approx \frac{1}{5\sigma_{RMS}}$ where σ_{RMS} is the Root Mean Square delay spread of the channel [25]. Under commonly used Tapped Delay Line (TDL) channel models like Extended Pedestrian A (EPA) or Extended Typical Urban (ETU) [26], the coherence bandwidth is respectively $B_c \approx 4.66$ MHz for EPA and $B_c \approx 200$ kHz for ETU which are larger than 45 KHz. In order to verify this approximation, simulations with multi-path channel models have been also conducted. We compared the one-tap channel model with EPA and ETU models. The power delay profile of EPA and ETU models can be found in [26]. TS2 and TS3 were evaluated on this comparison. In case of TS2, the phase is considered to be a constant value and it is equalised in the same way for both channel models (see Section 3.5.2). The simulation results did not show significant difference on the performance and therefore we consider that one-tap channel model is a fair assumption for our system model.

To conclude with the scenario considerations, it must be taken into account the reduction of the system complexity to one-single cell scenario. Multi-cell TBMA is a potential line of future work and has already been studied in [15]. One relevant constraint is the inter-cell interference due to the non-orthogonal frequency reuse among cells. When deploying a cellular network, the effect of interference on the performance metrics must be analysed. This effect is less severe in campus network deployments like indoor industrial plants. Nonetheless, NB-IoT NPRACH preambles are already aware of the harmful effects of inter-cell interference and add a pseudo-random hop frequency pattern to decrease this effect (see Section 2.2).

5.2 Transmission schemes

Table 5.1 summarizes the system implications of each of the transmission schemes described on Section 3.5. The main conditional factors are the channel and TA compensation.

Channel compensation

The channel is compensated at the transmitter side. Previous to compensation, the UL channel must be estimated. Two UL channel estimation strategies are proposed: (*i*) individual UL channel estimation at the BS for all devices using UL independent reference signals and feedback from the BS to devices about their individual channels or (*ii*) individual DL channel estimation at the devices using DL reference signals and consider UL reciprocity. On the one hand, the main drawback of (*i*) is that on presence of large number of devices, the overhead generated for this process is very high (each device needs independent resources on UL and DL). On the other hand, on (*ii*) the process only requires one set of DL reference signals common for all devices and the individual channel estimation at each device. Nevertheless, (*ii*) can only be assumed on behalf of TDD systems since UL reciprocity is required. Some cellular system usually include a Time Division Duplex (TDD) mode. Since Release 15, NB-IoT supports TDD mode [27]. On this way by using the Narrowband Reference Signals (NRS), DL channel estimation would be enabled with no modification of the UL or DL NB-IoT frame structure. The reciprocity between DL and UL channels is also supported by the fact that the channel time-variation is diminished due to the null mobility of the devices. However, up to now, NB-IoT does not include mechanisms to equalise the UL channel at the transmitter. As mentioned, by using the existing physical layer, DL channel estimation and UL reciprocity can be conducted but modifications on the transmitted signal are required to equalise the channel at devices.

Full channel equalisation is constrained by the limited power consumption of the devices. This limitation was already taken into account by existing TBMA work [9], [12], [28]. On presence of deep fades, the required power to compensate such fades is potentially higher than the maximum allowed by the battery of the device. Additionally, the system might also limit the transmitted power to avoid interference. Assuming that the channel amplitude can be fully compensated is not realistic. Since the power-issues of channel compensation are mostly related to TS1, three possible alternatives are proposed to original the TS1 [29]:

a) <u>Channel compensation with saturation after certain power value</u>.

   Devices compensate the channel up to a certain channel power γ. When the required power compensation exceeds γ, the compensation saturates to that value. The compensation factor is then expressed by the following expression

$$\alpha_k = \begin{cases} \frac{1}{|h_k| e^{j\phi_k}} & \text{if } |h_k|^2 > \gamma \\ \frac{1}{\sqrt{\gamma} e^{j\phi_k}} & \text{if } |h_k|^2 \leq \gamma \end{cases}$$

   The value of the parameter γ should be determined by the power limitation of the device and the maximum UL transmitted power defined by the regulation standard. The disadvantage of this strategy is that if the channel cannot be completely compensated, it might lead to a wrong empirical measure estimation. Note that this modification could be incorporated to TS1 and P_e performance is expected to be be laying between original TS1 and TS2 cases.

b) Switching transmission off in presence of deep fades.

Devices compensate the channel up to a certain channel power γ. If the required power compensation exceeds γ, the device do not transmit. The channel compensation factor is therefore

$$\alpha_k = \begin{cases} \frac{1}{|h_k|e^{j\phi_k}} & \text{if } |h_k|^2 > \gamma \\ 0 & \text{if } |h_k|^2 \leq \gamma \end{cases}$$

If the number of active devices is sufficiently high, this might be a good choice to save power. The impact on the P_e performance is given by the reduction of active devices. For instance, if $K = 100$ but 10 devices are observing deep fades and switch off, then only 90 devices are effectively active.

c) Only channel phase compensation

Devices only compensate the phase of the channel, therefore not compensating the channel fading (amplitude). Unlike TS1 a) or TS1 b), the transmitted power does not depend on how deep is the fading event. Additionally, note that this strategy is followed by TS2, and even though the fading is not compensated, the P_e performance is comparable to TS1.

Considering that $\gamma = 0.25$, Figure 5.1 compares on Scenario A and $M = 2$ the original TS1 with alternatives a) and b) for TS1, and TS2. A power constraint of $\gamma = 0.25$ implies that devices compensate up to 6 dB of channel fading. As observed, TS1 provides the best performance while TS2 the worst. The performance of alternatives proposed for TS1 lay in between TS1 and TS2. For TS1 b) and due to channel coefficients are Rayleigh faded, the power constraint causes that 11.75% of devices switch off. In the same way, for TS1 a) only 11.75% of devices are not able to fully compensate the channel amplitude. Since the limitation only affects to a few devices, it is observed close performance between the original TS1 and its alternatives. Higher γ would provoke the approximation of TS1 a) and TS1 b) to the performance of TS2.

The choice of one strategy or another is a trade-off between the power consumption limitation of devices and achieving sufficient P_e performance for the use case. For instance, on some applications the devices might be embedded on locations with access to constant power supply and TS1 strategies like a) or b) are feasible. However, on typical mMTC uses cases like smart agriculture or remote metering, devices have limited power consumption and therefore reduced energy-transmission turn into a priority choice to consider (e.g. TS2 or TS3). Even in very extreme energy-limited systems, TS2 should be avoided because the compensation of the channel phase implicates the constant or periodic process of channel estimation which also consumes battery.

TA compensation

On TS1 and TS2, devices set their transmission delay to $D_k = t_{TA_k}$ in order to compensate their TA. Be aware that devices need to know how far from the BS they are in terms of distance

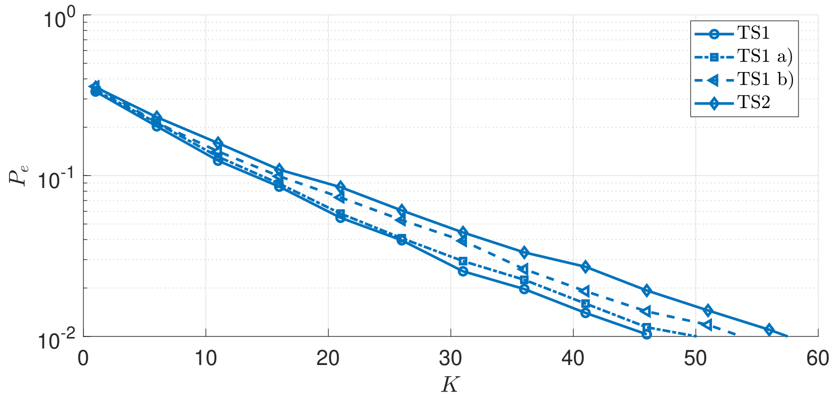


Figure 5.1: Error probability of power-constrained schemes on Scenario A and $M = 2$ with power constraint $\gamma = 0.25$

or propagation delay. To that end, the system should have been functioning for a long time and devices should already know their TA (either they computed it or the BS computed it and communicated to them). NB-IoT provides already methods to determine the TA. On the normal functioning of the NPRACH, preambles are utilised to start RA procedures. On this procedure, the BS computes the TA of the device and communicates back to it. If the devices are static (assumption discussed on Section 5.1), the TA does not change and therefore devices can keep the TA communicated on their first RA attempt. In case mobility exists, the device should communicate to the BS about the position modification and start a TA update procedure.

TA compensation could be avoided if devices are grouped in a small area with similar distance to BS. In those cases, the TA is approximately the same among devices, therefore the signals would arrive aligned in time at the receiver without necessity of TA compensation.

5.3 Physical layer

As proposed on Section 3.6, the physical layer of NB-IoT can be used with very few modifications. In particular, the TBMA implementation only affects to UL channels, and for that, the NPRACH could be utilised because it contains transmitted signals based on orthogonal preambles. The advantages of using the NPRACH preambles are mainly two: they are narrowband signals and have low power consumption. These characteristics oppose to RA preambles from other technologies like LTE where the preambles are based on Zadoff-Chu (ZC) sequences [20] which have much larger bandwidth (1.08 MHz) and require higher power amplifier back-off, reducing the efficiency of the power amplifier and out-consuming the device battery life time [21].

Channel estimation does not require from modification of the physical layer structure. As previously mentioned, NRS can be used to estimate the channel on DL, and the TDD mode of NB-IoT enables UL reciprocity. NRS are included on broadcast and dedicated DL transmissions (NPBCH, NPDCCH and NPDSCH). However, channel compensation at transmitter side is not configured by default on NB-IoT. Looking at the signal model described on Section 3.2, the most relevant parameters introduced by the system model are α_k and D_k which are device-specific. According to the NB-IoT standard [20], preambles are scaled by the parameter β_{NPRACH} according to the transmitted power (see Equation (2.2)). This parameter is device-specific and could be equivalently used as α_k. On its normal functioning, β_{NPRACH} is used by the power control mechanism of NPRACH, similarly to TS1 a) scheme, to compensate the channel power up to certain maximum transmission power based on the DL channel estimation from NRS [22]. This mechanism can be reused for TS1 a) and modifications for the switching off of devices could be included in case TS1 b) is used. Additionally, TS1 strategies and TS2 compensate the phase of the channel coefficient at each device. To that end, if the parameter β_{NPRACH} is considered as a complex coefficient, the phase of the channel could be compensated without addition of new paremeters. Hence, our parameter α_k could be substituted by β_{NPRACH} on NB-IoT standard.

The other parameter, transmission delay D_k, is just a delay introduced by device k to counteract (or not) its propagation delay (i.e. the TA). The only implication on the transmitted signal is a modification of the transmission time. TA compensation strategies address the achievement of UL time synchronization. By default, in the UL shared channel NPUSCH, signals are required to achieve UL time synchronization to be transmitted in a orthogonal fashion on the grid of physical resources. On this way, devices modify the transmission time to make UL signals arrive at the receiver at specific time. In case TA compensation is enabled, NPRACH signals would incorporate the UL time synchronization characteristic of NPUSCH.

Another fact to highlight is the non-modification of the BS receiver processing. At the BS, the usual functioning on RA protocol incorporates a bank of correlation operations between the NPRACH preambles and the received signal. The proposed TBMA receiver keeps same correlation-based processing and the only implication is a different interpretation of the output of the correlator.

5.4 Clustering of devices

It is common that in cellular protocols devices are clustered (or grouped) according to specific guidelines. Depending on each cluster, devices can adapt some parameters of the physical layer. Clustering can be taken as an advantage on the communication protocol to improve the radio link, include explicit information or make special requests for a specific groups of devices (e.g. reduce transmit power). TBMA-based protocols might also take advantage of clustering either by utilising existing schemes or by introducing an adapted one.

5.4.1 Existing cellular clustering

On common cellular networks like LTE or NB-IoT, clustering of devices is usually determined by the quality of the wireless channel, i.e. the coverage level. On NB-IoT, up to 3 CE levels are defined. Depending on CE level, the devices usually have different N_{rep} and different codebook of preambles (see Chapter 2). This existing clustering strategy can be used and taken as an advantage for TBMA. For instance, devices with high CE level (bad channel quality and/or far from BS) might consume more power due to the higher number of repetitions. For that cluster, low-power consumption schemes might be a good option (e.g. TS2 or TS3). Additionally, the TS1 alternatives a) and b) proposed on Section 5.2 should be also considered. On the other side, devices with low CE level (good channel quality and/or close to BS) are prone to have more available power and therefore take advantage of its better P_e performance by means of TS1 alternatives.

Other advantage of the CE-based clustering is the implicit information contained on the fact of belonging to a certain cluster. For example, on open environments with high presence of Line-Of-Sight (LOS) (e.g. rural area on smart agriculture applications), the CE-level is highly correlated with remoteness of the device. On those cases, and based on that each cluster utilises a different set of preambles, the fusion-center could estimate the value of the QoI at different remoteness from the BS.

Although the system model of this thesis does not consider more than one cluster, be aware that the NB-IoT clustering technique does not affect to the P_e performance. As explained on Section 2.2, each cluster is assigned a different set of preambles or codebook. The different groups can be treated as independent systems with different configurations (transmission scheme, preambles, number of states, etc.). Hence, several clusters can work in parallel and the QoI can be evaluated for each of them separately.

5.4.2 TBMA-based clustering

The definition of specific clustering strategies for TBMA systems may result in an improvement of the distributed estimation. Instead of using the CE-based clustering, a special TBMA grouping could be used instead. Different group-rules can be defined. For instance, the cluster criterion can be the distance to BS (or TA). By this means, each cluster represents a group of devices with different remoteness to the BS. This strategy has some correlation with the CE-based clustering due to, in general terms, the further from the BS, the worse coverage level. Nonetheless, by fixing the distance to BS as the group-rule we ensure to group the devices according to their remoteness. The distance-based clustering provides additional information about the QoI. Since the fusion-center monitors the QoI for each cluster, information about the QoI with respect to the distance is also acquired. On Figure 5.2 we depict a scenario illustration where devices monitor a parameter related to the rain and devices are clustered on two groups depending on the distance to BS (close or far from BS). As observed, the remote group observes $\theta = \theta_2$ (i.e. it is raining) and the closer cluster observes $\theta = \theta_1$ (i.e. it is not raining). Thanks to the distance-based clustering, the fusion-center obtains information of the

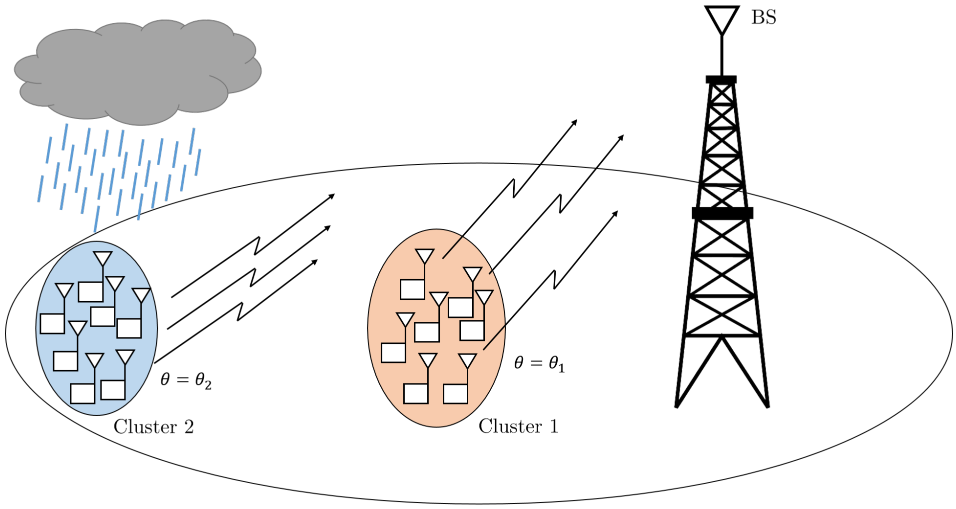


Figure 5.2: Illustration of distance-based clustering on the monitoring of a rain-dependent parameter.

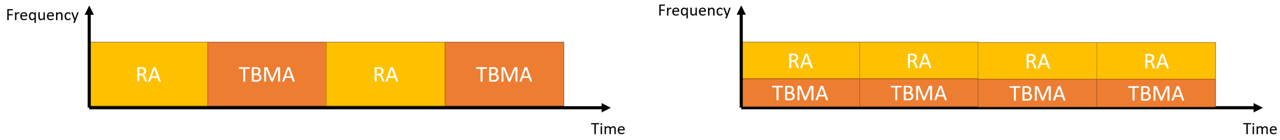
rain parameter at each cluster, therefore having implicit knowledge of the distance.

Another alternative would be to make groups measure a different QoI. For instance, a system with two groups, one could monitor temperature and the other pollution level. This strategy allows having two TBMA protocols in parallel. Many other rules can be defined to either acquire additional information about the QoI or to use efficiently the available resources. In fact, TBMA-based clusters might reuse the physical layer implications of cellular grouping like NB-IoT. The only required modification would be a different interpretation of the cluster rule (e.g. distance-based instead of CE-based). The rest of implications can be reused, i.e. each cluster has a different set of preambles and each group can set a different N_{rep}.

5.5 Extension to device-centric protocol

TBMA is generally used as an information-centric protocol, where instead of interested on the individual payload, the system is interested on the information of the QoI. Nevertheless, after obtaining general information of the QoI, depending on the acquired information, additional information to specific devices could be requested such as location of specific devices, real-valued measurements of specific devices, constant monitoring of the parameters of interest, other QoI, etc.

Therefore, we also consider extensions to device-centric approaches. The use of a cellular protocol is an interesting choice due to existing communication protocols could be reused. On NB-IoT, before sending any requested information, devices need to obtain UL resources through a RA procedure. Be aware that the RA procedure on NB-IoT is conducted through the NPRACH which we propose to be also used for TBMA. Nonetheless, both RA and TBMA protocols can coexist on the same physical channel. Figure 5.3 illustrates two possible configurations to enable RA and TBMA over the NPRACH. One way to achieve this is to separate



(a) Time-alternate RA and TBMA

(b) Parallel RA and TBMA

Figure 5.3: Illustration of coexistence between RA and TBMA on the NPRACH.

both protocols at different time windows, that is, both procedures are alternated on time. Another way is to execute both in parallel by splitting the preambles for both processes. In both cases, the fusion-center must be aware of the different interpretation of the received preambles at each time windows and/or set of preambles. Devices could be informed of the RA configuration through the SIB on the NPBCH (as usual) which could also contain information about the TBMA configuration. In that case, modification on the NB-IoT SIB by including TBMA-specific fields should be considered on future standardisation processes.

The greatest inconvenient of using RA procedures to acquire further information is that under high number of devices and the reduced amount of RA preambles, simultaneous RA attempts might provoke high number of collisions. Hence, we propose two strategies:

I. Acquisition of device-specific information from devices observing m.

The fusion-center might be interest on getting further information from devices observing a value m or set of m values (e.g. large values of m values which, for example, mean that the temperature is very high). On those cases, acquiring further information from the devices (e.g. their location) might be critical on some applications.

The UL data request should be started by the fusion-center by means of a DL message, for instance, stated on the Downlink Control Indicator (DCI), and requesting that all devices which observed m should start a RA procedure. The configuration parameters of the RA could be also embedded on the DCI, defining when and which preambles can be used. If devices success on the RA attempt, they will get individual UL resources in which they can transmit device-specific information (e.g. location or another measurement). Note that this strategy is feasible if the number of devices K_m which observed m is relatively low. Authors studied on [19] the RA success probability for different number of devices, number of attempts and number of preambles. For instance, with for 3 attempts, 12 preambles and 30 devices on the RA procedure, the success probability is 0.2. When assigning the resources for RA, these results should be taken into account. In case K_m is considerable large, an attempt probability could be stated on the DCI defining the probability of a starting a RA attempt of devices observing m. With this probability, we reduce the amount of devices starting the RA procedure and therefore reducing the number of collisions.

II. Acquisition of device-specific information from devices on a given area

If the value of the QoI discloses an exceptional or important situation, the fusion-center can start a process to get further information from the devices on a specific location or area. Through a DL message (for instance, on a DCI), the fusion-center must state the value of the parameters that define the location or the area (e.g. coordinates, distance to BS or antenna through which the devices are connected). Then, the devices located on each defined area start RA procedures in order to get UL resources and provide device-specific information. Note that the devices must be aware of their location or the area. This strategy is specially relevant if the devices are already clustered according to their remoteness or location.

In any of the cases I. or II., the request of device-specific information can be done through the DCI. In case of utilising the NB-IoT technology, the DCI is included on the NPDCCH and has reserved fields (on the DCI format N1) to make individual RA requests initiated by a "PDCCH order" [22]. These orders are usually for specific UEs and specify the preamble they have to use during RA procedure (no collisions happen, also known as non-contention based). This available design can be reused without modification and with a different interpretation. For instance and regarding strategy I., the field which normally states the identifier of the device requested on the "PDCCH order" could encode the device-specific information that is requested. The field which specifies the allocated subcarrier/preamble for a specific device (I_{SC}) can be interpreted as the value that the devices had to measure ($m = I_{SC}$) to start the request of UL resources. In such way, that if $I_{SC} = m$, all devices measuring m should start a RA procedure.

6

Conclusion

This master thesis has studied the use of TBMA for parameter estimation on massive connectivity scenarios and considering system limitations. TBMA has been proposed as a solution to achieve centralized estimation of a QoI together with savings of UL power and bandwidth. NB-IoT has been chosen as the most suitable technology for TBMA due to its physical layer design for massive connectivity. Motivated by these facts, the reuse of most NB-IoT features for TBMA systems has been considered.

Three transmission schemes have been proposed with the objective of comparing the system performance under different system configurations. The proposed schemes include alternatives to counteract or not the non-coherent superposition of signals caused by the wireless channel and TA of each device. The system model has incorporated the use of the NB-IoT UL physical layer, with special stress on the preambles from the NPRACH. These preambles have been used to encode the devices' observations following the TBMA principle. Two scenario have been evaluated. On the first, a binary QoI case has been analysed under the dependency of the number of preambles encoding the devices' observations. Results showed that short number of preambles ($M \leq 4$) are sufficient to achieve a error probability objective of 10^{-2} when coherent addition of signals is achieved. However, when coherent superposition is not achieved by the communication scheme, configurations with up to 48 preambles are required to obtain error probabilities on the 10^{-2} regime. On the second scenario, a multi-state QoI was studied with the objective of analysing the influence of the number of states into the estimation performance. Results showed that increasing the number of states on the parameter of interest provokes an increase of the error probability. In both scenarios, the best performance belonged to schemes with channel and TA compensation strategies. Nonetheless, on scenarios where devices are located close to each other, their TA can be approximated as the same and TA compensation is not required to achieve UL time synchronization.

An analysis of the practicalities of TBMA systems has been carried out. Main conclusions are that TBMA can be deployed using the NPRACH of NB-IoT, only requiring a different interpretation of the transmitted and received preambles. Additionally, the existing protocol on the NPRACH (i.e. Random Access) could coexist with TBMA over the NPRACH by either alternating them at different time-windows or splitting the available preambles for both processes. Each transmission scheme implication has been evaluated. Main outcome is that devices' power limitation caused by the compensation of channel amplitude can be handled by solutions like only compensating the channel phase or limiting the maximum compensable power. In those cases the estimation performance approximates to the full channel compensation performance. Finally, other appreciated advantages of using cellular technologies like NB-IoT for TBMA are the reuse of clustering and the possible extension to device-centric protocols to either include implicit information or transmit device-specific information on request.

6.1 Future work

The use of transmission schemes with reduced complexity has great advantages from the implementation point of view. However, this type of schemes suffer from the severe effects of radio propagation resulting into a deterioration of the estimation performance. Hence, enhancements focused on the improvement of these configurations are of high interest on the extension of this work. As discussed on Section 4.3, we consider as interesting future line the use of machine learning-assisted solutions to improve the performance estimation of TBMA systems. Furthermore, local estimation strategies could outperform TBMA-based systems on certain scenarios and its comparison must be analysed under a NB-IoT implementation perspective. Another interesting direction is the extension of the one-cell TBMA system presented on this thesis to a the multi-cell set up. Continuing with the analysis studied on [15], considering a multi-cell NB-IoT system that incorporates the TBMA protocol could be a very relevant research line. Here, we find specially interesting the implication of the frequency hopping pattern of NB-IoT preambles which includes a pseudo-random hop pattern for repetitions in order to reduce inter-cell interference. Its effect on the estimation performance of a multi-cell scenario must be further investigated.

Bibliography

[1] S. Mattisson, “Overview of 5G requirements and future wireless networks”, in *ESSCIRC 2017 - 43rd IEEE European Solid State Circuits Conference*, 2017, pp. 1–6.

[2] ITU-R. (2015). IMT vision—framework and overall objectives of the future deployment of IMT for 2020 and beyond, [Online]. Available: `https://www.itu.int/dms_pubrec/itu-r/rec/m/R-REC-M.2083-0-201509-I!!PDF-E.pdf`.

[3] J. Gantz and D. Reinsel, “The Digital Universe in 2020: Big Data, Bigger Digital Shadows, and Biggest Growth in the Far East”, 2012.

[4] 3GPP, “Study on RAN Improvements for Machine-Type-Communications”, TR 37.868, 2011, Version 11.0.

[5] M. B. Shahab, R. Abbas, M. Shirvanimoghaddam, and S. J. Johnson, “Grant-free Non-orthogonal Multiple Access for IoT: A Survey”, *IEEE Communications Surveys Tutorials*, pp. 1–1, 2020.

[6] 3GPP, “Service requirements for Machine-Type Communications”, Tech. Rep. 22.368, 2019, Version 15.0.

[7] S. Narayanan, D. Tsolkas, N. Passas, and L. Merakos, “NB-IoT: A Candidate Technology for Massive IoT in the 5G Era”, in *2018 IEEE 23rd International Workshop on Computer Aided Modeling and Design of Communication Links and Networks (CAMAD)*, 2018, pp. 1–6.

[8] Y. Wu, X. Gao, S. Zhou, W. Yang, Y. Polyanskiy, and G. Caire, *Massive Access for Future Wireless Communication Systems*, 2019. arXiv: `1910.12678`.

[9] K. Liu and A. M. Sayeed, “Type-Based Decentralized Detection in Wireless Sensor Networks”, *IEEE Transactions on Signal Processing*, vol. 55, no. 5, pp. 1899–1910, 2007.

[10] Z. Utkovski, O. Simeone, T. Dimitrova, and P. Popovski, “Random Access in C-RAN for User Activity Detection With Limited-Capacity Fronthaul”, *IEEE Signal Processing Letters*, vol. 24, no. 1, pp. 17–21, 2017.

[11] K. Liu, H. El Gamal, and A. Sayeed, “Decentralized Inference Over Multiple-Access Channels”, *IEEE Transactions on Signal Processing*, vol. 55, no. 7, pp. 3445–3455, 2007.

[12] G. Mergen and L. Tong, “Type based estimation over multiaccess channels”, *IEEE Transactions on Signal Processing*, vol. 54, no. 2, pp. 613–626, 2006.

[13] A. Anandkumar and L. Tong, “Type-Based Random Access for Distributed Detection Over Multiaccess Fading Channels”, *IEEE Transactions on Signal Processing*, vol. 55, no. 10, pp. 5032–5043, 2007.

[14] J. Dommel, Z. Utkovski, S. Stańczak, and O. Simeone, “Joint Source-Channel Coding and Bayesian Message Passing Detection for Grant-Free Radio Access in IoT”, in *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2020, pp. 8574–8578.

[15] R. Kassab, O. Simeone, and P. Popovski, *Information-Centric Grant-Free Access for IoT Fog Networks: Edge vs Cloud Detection and Learning*, 2019. arXiv: 1907.05182.

[16] D. Telekom, “White paper: Narrowband IoT, The Game Changer for the Internet of Things”, Tech. Rep., 2017.

[17] R. Schwarz, “White paper: Narrowband Internet of Things”, Tech. Rep., 2016, p. 42.

[18] L. Feltrin, G. Tsoukaneri, M. Condoluci, C. Buratti, T. Mahmoodi, M. Dohler, and R. Verdone, “Narrowband IoT: A Survey on Downlink and Uplink Perspectives”, *IEEE Wireless Communications*, vol. 26, no. 1, pp. 78–86, 2019.

[19] G. H. Baracat and J. M. C. Brito, “NB-IoT Random Access Procedure Analysis”, in *2018 IEEE 10th Latin-American Conference on Communications (LATINCOM)*, 2018, pp. 1–6.

[20] 3GPP, “Evolved Universal Terrestrial Radio Access (E-UTRA); Physical channels and modulation”, TS 36.211, 2019, Version 14.8.

[21] X. Lin, A. Adhikary, and Y. Eric Wang, “Random Access Preamble Design and Detection for 3GPP Narrowband IoT Systems”, *IEEE Wireless Communications Letters*, vol. 5, no. 6, pp. 640–643, 2016.

[22] 3GPP, “LTE; Evolved Universal Terrestrial Radio Access (E-UTRA); Physical layer procedures”, TS 36.213, 2019, Version 14.9.

[23] I. Csiszár and P. Shields, “Information Theory and Statistics: A Tutorial”, *Foundations and Trends in Communications and Information Theory*, vol. 1, no. 4, pp. 417–528, 2004.

[24] M. Karmoose, K. G. Seddik, and H. El Kamchouchi, “Censoring for Type-Based Multiple Access Scheme in Wireless Sensor Networks”, in *2012 IEEE Vehicular Technology Conference (VTC Fall)*, 2012, pp. 1–6.

[25] T. Rappaport, *Wireless Communications: Principles and Practice*, 2nd. USA: Prentice Hall PTR, 2001.

[26] 3GPP, “Evolved Universal Terrestrial Radio Access (E-UTRA); User Equipment (UE) radio transmission and reception”, TS 36.101, 2019, Version 14.8.

[27] R. Ratasuk, N. Mangalvedhe, Z. Xiong, M. Robert, and D. Bhatoolaul, “Enhancements of narrowband IoT in 3GPP Rel-14 and Rel-15”, in *2017 IEEE Conference on Standards for Communications and Networking (CSCN)*, 2017, pp. 60–65.

[28] P. Gao and C. Tepedelenlioglu, “On the Parameter Estimation over Fading Channels with TBMA Wireless Sensor Networks”, in *2006 IEEE International Conference on Communications*, vol. 8, 2006, pp. 3408–3413.

[29] F. Tufvesson, M. Faulkner, and T. Maseng, “Pre-compensation for Rayleigh Fading Channels in Time Division Duplex OFDM Systems”, in *Proceedings of 6th IEEE International Workshop on Intelligent Signal Processing and Communication Systems, Melbourne, Australia, Nov 5-6, 1998*, 1998, pp. 53–57.

A

Probability distribution of devices' observations

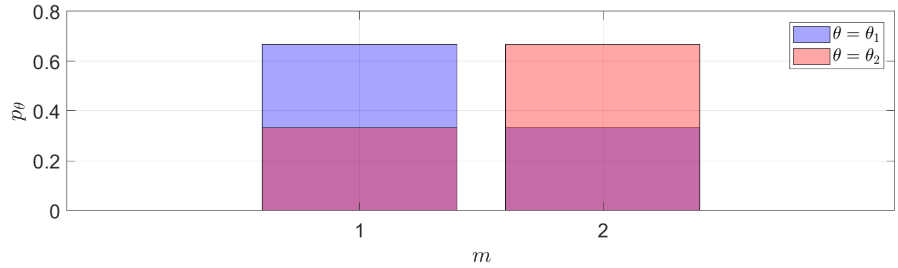
Hereby we illustrate the probability distributions of the observations from devices on Scenario A and B which are described on Sections 4.1.1 and 4.1.2.

A.1 Scenario A

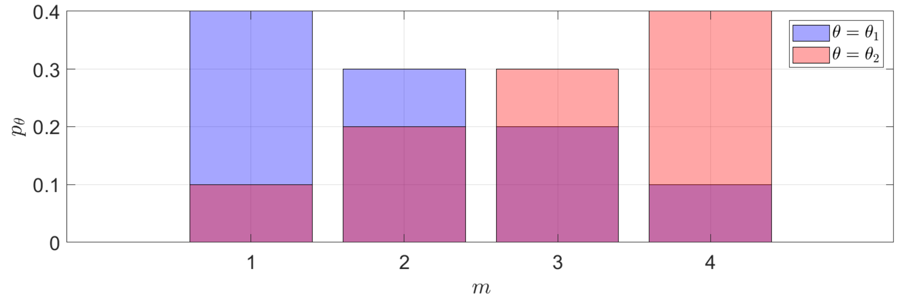
For the evaluated number of preambles $M = 2, 4, 12, 24$ and 48, the observations' distributions are illustrated on Figure A.1.

A.2 Scenario B

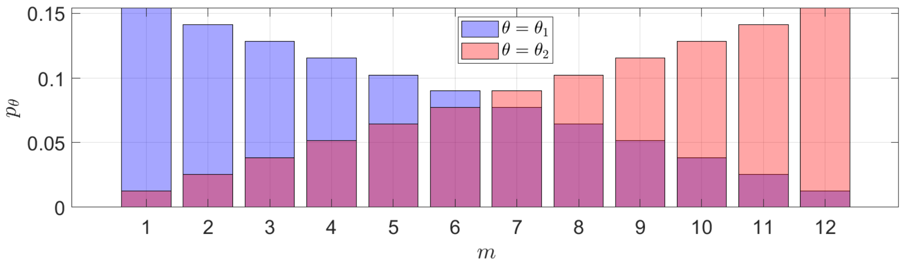
Figure A.2 illustrates the probability distribution of each θ value evaluated on Scenario B. Additionally, Figure A.3 illustrates the distribution of the states considered on each Q case on Scenario B.



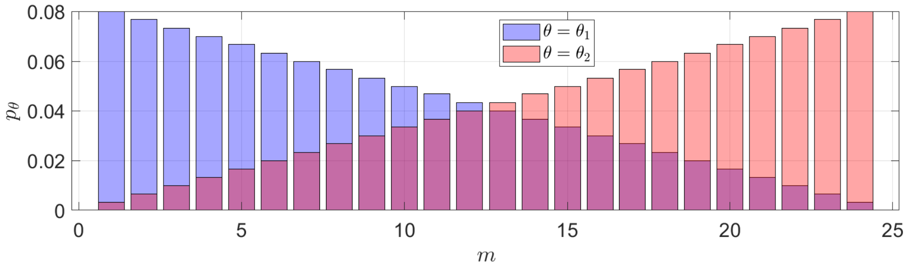
(a) $M = 2$



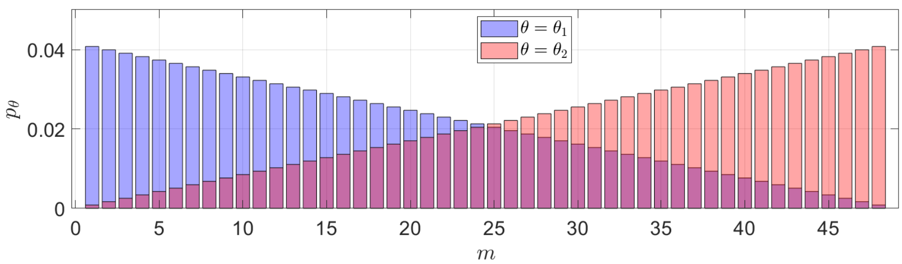
(b) $M = 4$



(c) $M = 12$

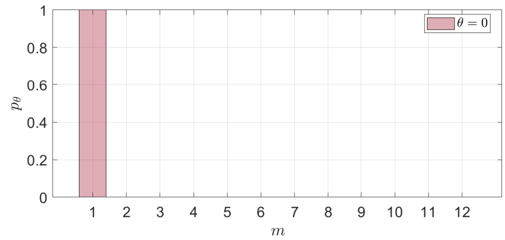


(d) $M = 24$

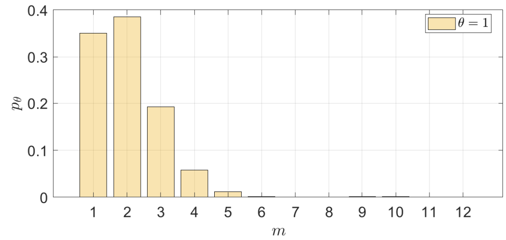


(e) $M = 48$

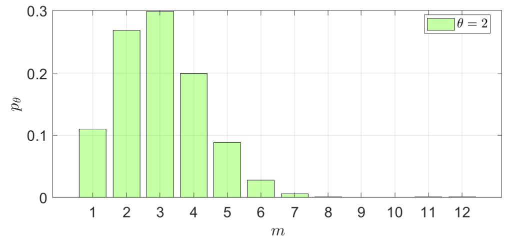
Figure A.1: Probability distribution of the observations on Scenario A.



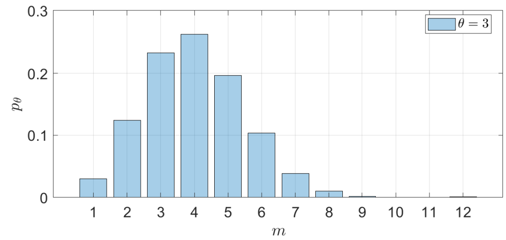
(a) $\theta = 0$



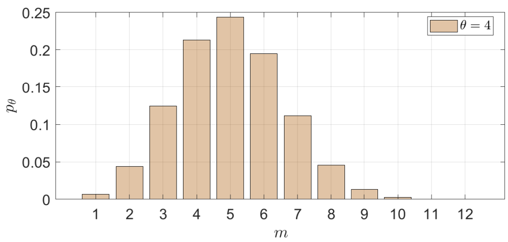
(b) $\theta = 1$



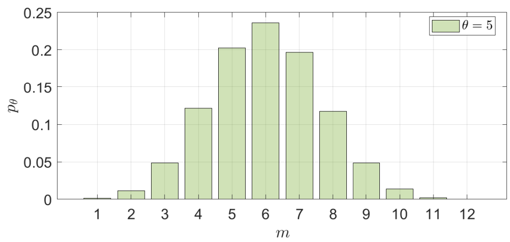
(c) $\theta = 2$



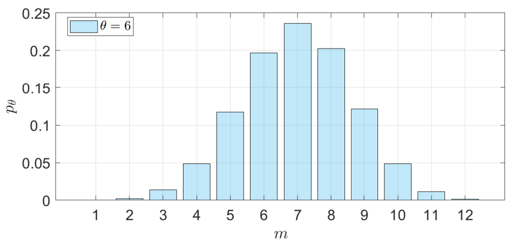
(d) $\theta = 3$



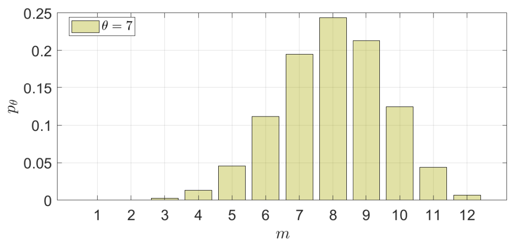
(e) $\theta = 4$



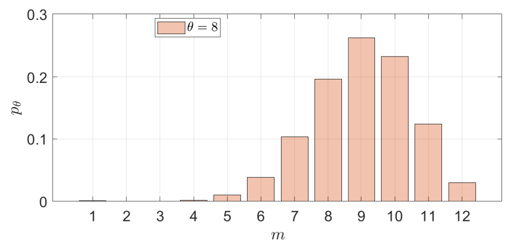
(f) $\theta = 5$



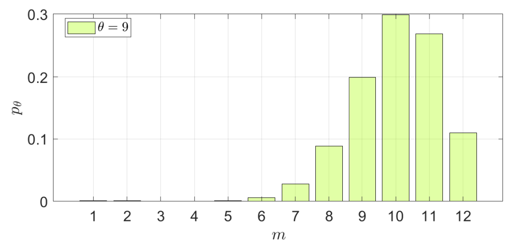
(g) $\theta = 6$



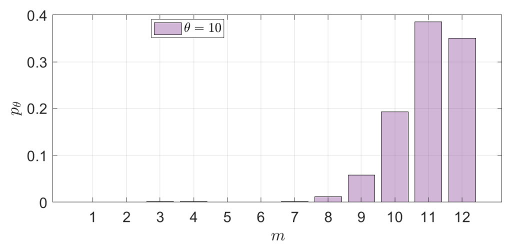
(h) $\theta = 7$



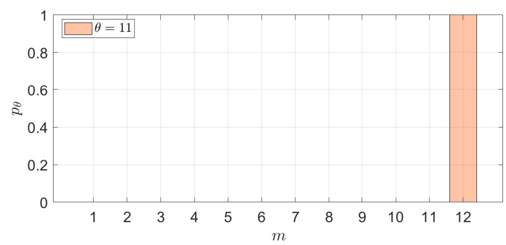
(i) $\theta = 8$



(j) $\theta = 9$

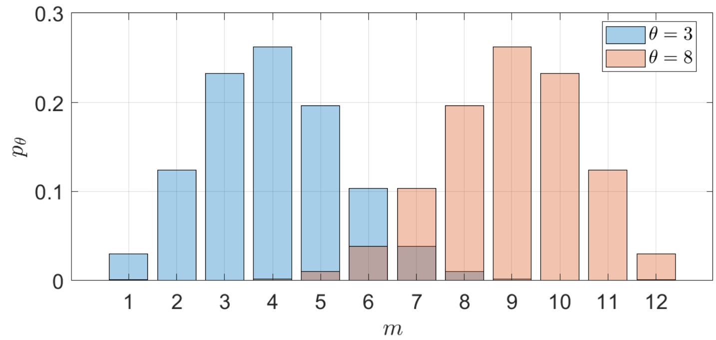


(k) $\theta = 10$

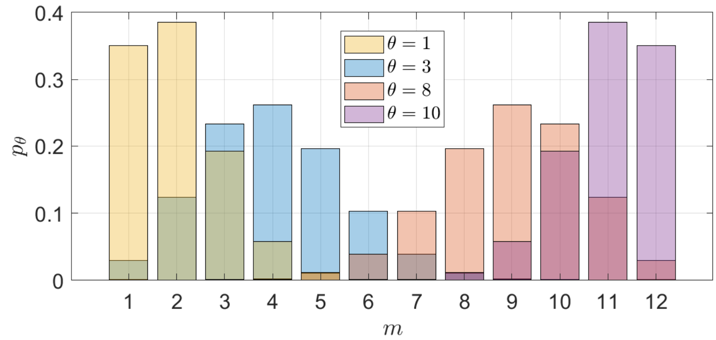


(l) $\theta = 11$

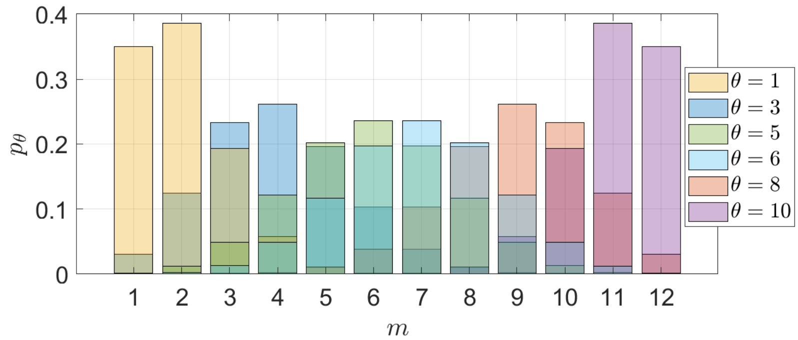
Figure A.2: Probability distribution of the observations on Scenario B for each state θ.



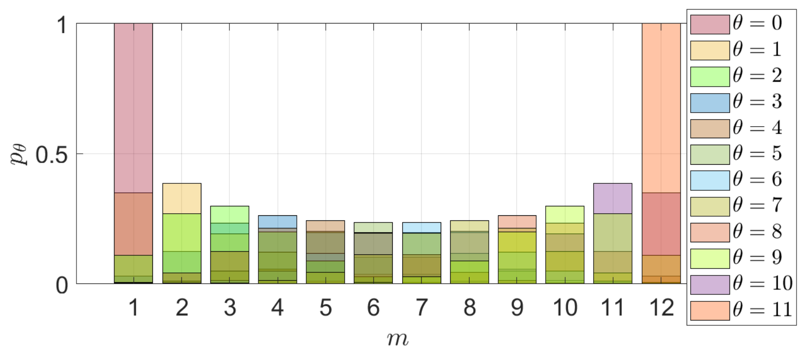
(a) $Q = 2$



(b) $Q = 4$



(c) $Q = 6$



(d) $Q = 12$

Figure A.3: Probability distribution of the observations on Scenario B.