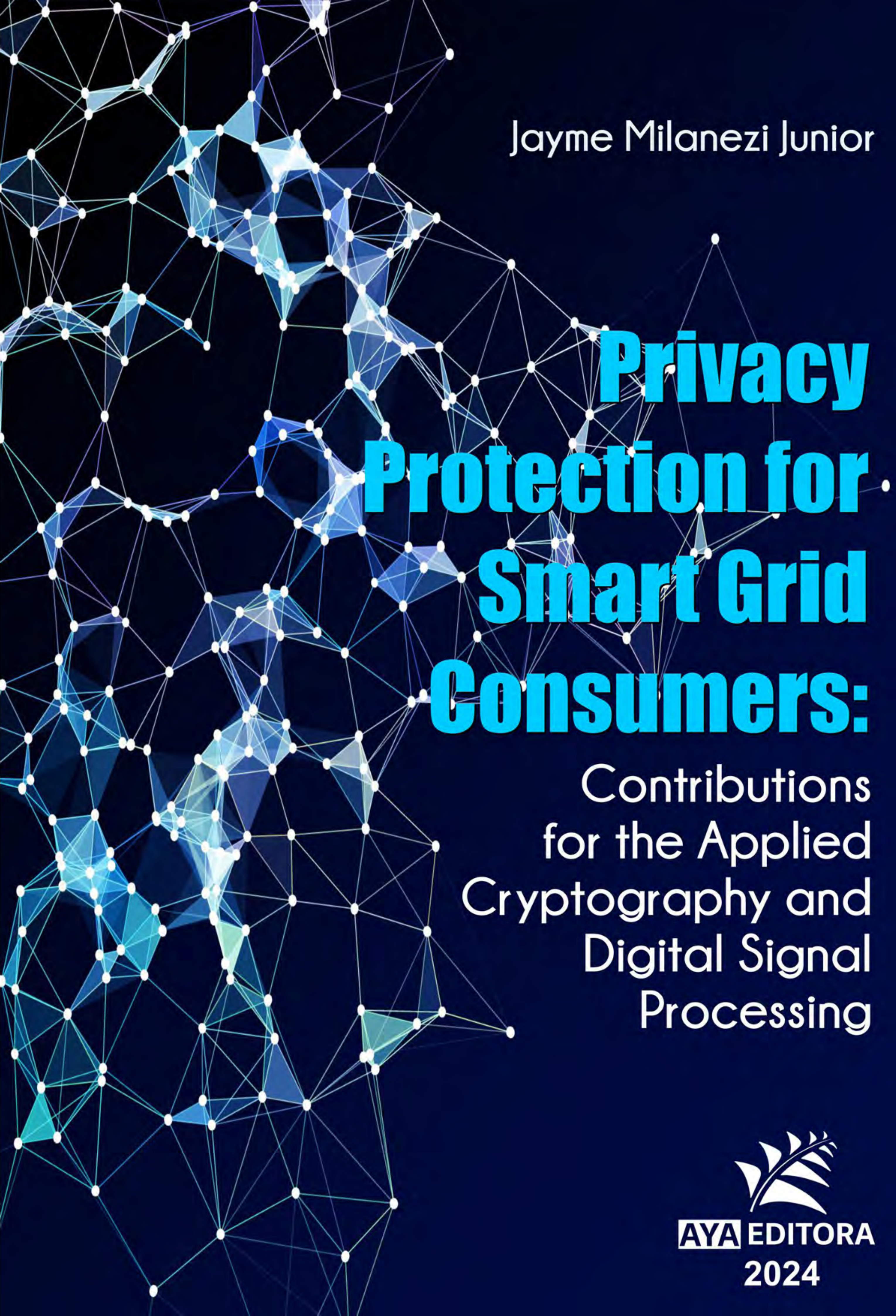


Jayme Milanezi Junior

Privacy Protection for Smart Grid Consumers:

Contributions for the Applied Cryptography and Digital Signal Processing

AYA EDITORA
2024

Jayme Milanezi Junior

Privacy Protection for Smart Grid Consumers: Contributions for the Applied Cryptography and Digital Signal Processing

Ponta Grossa
2024

Editorial Direction
Ph.D. Adriano Mesquita Soares

Author
Ph.D. Jayme Milanezi Junior

Cover
AYA Editora©

Revision
The author

Business Executive
Ana Lucia Ribeiro Soares

Editoral Production
AYA Editora©

Cover Images
br.freepik.com

Knowledge Area
Engineering

Editorial Board

Ph.D. Adilson Tadeu Basquerote Silva
Universidade para o Desenvolvimento do Alto Vale do Itajaí

Ph.D. Aknaton Toczek Souza
Centro Universitário Santa Amélia

Ph.D. Andréa Haddad Barbosa
Universidade Estadual de Londrina

Ph.D. Andreia Antunes da Luz
Faculdade Sagrada Família

Ph.D. Argemiro Midonês Bastos
Instituto Federal do Amapá

Ph.D. Carlos López Noriega
Universidade São Judas Tadeu e Lab. Biomecatrônica - Poli - USP

Ph.D. Clécio Danilo Dias da Silva
Centro Universitário FACEX

Ph.D. Daiane Maria de Genaro Chiroli
Universidade Tecnológica Federal do Paraná

Ph.D. Danyelle Andrade Mota
Universidade Federal de Sergipe

Ph.D. Déborah Aparecida Souza dos Reis
Universidade do Estado de Minas Gerais

M.Sc. Denise Pereira
Faculdade Sudoeste – FASU

Ph.D. Eliana Leal Ferreira Hellvig
Universidade Federal do Paraná

Ph.D. Emerson Monteiro dos Santos
Universidade Federal do Amapá

Ph.D. Fabio José Antonio da Silva
Universidade Estadual de Londrina

Ph.D. Gilberto Zammar
Universidade Tecnológica Federal do Paraná

Ph.D. Helenadja Santos Mota
Instituto Federal de Educação, Ciência e Tecnologia Baiano, IF Baiano - Campus Valença

Ph.D. Heloísa Thaís Rodrigues de Souza
Universidade Federal de Sergipe

Ph.D. Ingridi Vargas Bortolaso
Universidade de Santa Cruz do Sul

M.Sc. Jaqueline Fonseca Rodrigues
Faculdade Sagrada Família

Ph.D. Jéssyka Maria Nunes Galvão
Faculdade Santa Helena

Ph.D. João Luiz Kovaleski
Universidade Tecnológica Federal do Paraná

Ph.D. João Paulo Roberti Junior *Universidade Federal de Roraima*

M.Sc. Jorge Soistak
Faculdade Sagrada Família

Ph.D. José Enildo Elias Bezerra
Instituto Federal de Educação Ciência e Tecnologia do Ceará, Campus Ubajara

Ph.D. Karen Fernanda Bortoloti
Universidade Federal do Paraná

Ph.D. Leozenir Mendes Betim
Faculdade Sagrada Família e Centro de Ensino Superior dos Campos Gerais

M.Sc. Lucimara Glap
Faculdade Santana

Ph.D. Luiz Flávio Arreguy Maia-Filho
Universidade Federal Rural de Pernambuco

M.Sc. Luiz Henrique Domingues

Universidade Norte do Paraná

Ph.D. Milson dos Santos Barbosa

Instituto de Tecnologia e Pesquisa, ITP

Ph.D. Myller Augusto Santos Gomes

Universidade Estadual do Centro-Oeste

Ph.D. Pauline Balabuch

Faculdade Sagrada Família

Ph.D. Pedro Fauth Manhães Miranda

Universidade Estadual de Ponta Grossa

Ph.D. Rafael da Silva Fernandes

Universidade Federal Rural da Amazônia, Campus Parauapebas

Ph.D. Regina Negri Pagani

Universidade Tecnológica Federal do Paraná

Ph.D. Ricardo dos Santos Pereira

Instituto Federal do Acre

M.Sc. Rosângela de França Bail

Centro de Ensino Superior dos Campos Gerais

Ph.D. Rudy de Barros Ahrens

Faculdade Sagrada Família

Ph.D. Saulo Cerqueira de Aguiar Soares

Universidade Federal do Piauí

Ph.D. Silvia Aparecida Medeiros Rodrigues

Faculdade Sagrada Família

Ph.D. Silvia Gaia

Universidade Tecnológica Federal do Paraná

Ph.D. Sueli de Fátima de Oliveira Miranda Santos

Universidade Tecnológica Federal do Paraná

Ph.D. Thaisa Rodrigues

Instituto Federal de Santa Catarina

© 2024 - AYA Editora - The content of this book was submitted by the author for open access publication under the terms and conditions of the Creative Commons Attribution 4.0 International License **(CC BY 4.0)**. This book, including all illustrations, information, and opinions contained therein, is the result of the exclusive intellectual creation of the author. The author holds full responsibility for the content presented, which solely and entirely reflects their personal perspective and interpretation. It is important to emphasize that the content of this book does not necessarily represent the views or opinions of the publisher. The role of the publisher was strictly technical, limited to the services of layout and registration of the work, without any influence over the content presented or opinions expressed. Therefore, any questions, interpretations, or inferences arising from the content of this book should be directed exclusively to the author.

M637 Milanezi Junior, Jayme

Privacy Protection for Smart Grid Consumers: Contributions for the Applied Cryptography and Digital Signal Processing [recurso eletrônico]. / Jayme Milanezi Junior. -- Ponta Grossa: Aya, 2024. 165 p.

Texto em inglês
Inclui biografia
Formato: PDF
Requisitos de sistema: Adobe Acrobat Reader
Modo de acesso: World Wide Web
ISBN: 978-65-5379-506-8
DOI: 10.47573/aya.5379.1.262

1. Engenharia – Estudo e ensino. 2. Engenharia elétrica. 3. Redes elétricas inteligentes. 4. Energia - Conservação I. Título

CDD: 621.3

Ficha catalográfica elaborada pela bibliotecária Bruna Cristina Bonini - CRB 9/1347

International Scientific Journals Publicações de Periódicos e Editora LTDA
AYA Editora©
CNPJ: 36.140.631/0001-53
Fone: +55 42 3086-3131
WhatsApp: +55 42 99906-0630
E-mail: contato@ayaeditora.com.br
Web site: https://ayaeditora.com.br
Address: Rua João Rabello Coutinho, 557
Ponta Grossa - Paraná - Brasil
84.071-150

Summary

This book aims at revisiting a research line carried out between 2014 and 2019, when we tackled challenges related to the privacy of consumers in automated electrical power grids, the so-called Smart Grids. The result of that effort is the PhD Thesis PPGEE.TD-157/2019, by Universidade de Brasília (UnB), defended by this Author in August 2019. The Thesis' title is re-edited in this book in an attempt of improving its coherence with the content, since it addresses problems still observed in these grids. However, while part of the premises adopted remain valid, some technical issues do no longer hold, mainly from the standpoint of hardware.

Two main research lines related to the privacy of Smart Grid clients were exploited: (i) the technology concerning data transmission and reception through the physical layer in wireless contexts, where we discuss undersampling techniques for electromagnetic (EM) waveforms; and (ii) contributions in cryptography, with the proposal of a framework for a local Neighborhood Area Network (NAN), due to limitations of the Smart Meters (SM), namely the power meters that can send and receive data from/to the client's household and the utility company.

More specifically, we developed a framework for preserving the identity and privacy of the prosumer, a participant of the grid who is simultaneously consumer and producer of electrical energy – think of a neighbor that owns solar photovoltaic panels on his roof and batteries deployed elsewhere. This neighbor can storage energy and trade it in different instants of the day. Prosumers are a subset of consumers, i.e., even if someone does not generate energy, they can still purchase it. Our framework manages to provide privacy to consumers in general, hence improving data protection for every client, including prosumers.

From the hardware standpoint, it is well known that the electrical power meters, while migrating from the purely analogic status to the digital paradigm, have gained processing power, not only in terms of memory, but also speed. When we concluded the Thesis in 2019, the predominant literature reported a certain difficulty for the off-the-shelf SMs to execute highly complex calculations due to hardware limitations. One of the most known problems was related to processing cryptographic asymmetric keys, whose calculations demand a noticeable processing power, even nowadays. As a result, by 2018-2019, we have adopted symmetric keys-based solutions to SMs in our Smart Grid model.

However, after 5 years, it has been observed a substantial evolution in the SM hardware, overcoming the former hindrances related to calculations that were assumed unfeasible at that time. These SM machines can currently tackle such operations satisfactorily as well as connect to household local networks, possibly delegating a set of numeric calculations to hosts in their vicinity.

A collateral consequence of the evolution in hardware and processing capabilities of the SMs is the possibility of threats to the privacy protection of the Smart Grid user, the personal data owner, also named by in legal texts as the data subject. Despite the several forms of preserving user privacy in Internet applications, the SM itself remains the most vulnerable point of the grid. As stated on page 75, a sufficiently high data granularity (i.e., the arbitrarily little intervals between snapshots of instantaneous power consumption, when the sampling rate increases) can describe the whole lifestyle of the energy client. Given that the current SMs manage to produce progressively higher resolution, anyone who spots measurements from the household SM can infer practically everything that happens in there – even which film the residents are watching on TV. The Smart Grid user privacy can be in jeopardy due to the SM enhanced capabilities.

We propose the employment of SMs combined with Digital Processing undersampling techniques and a cryptographic framework for privacy improvement. Undersampling is achieved when different samplers work together to enhance their detectable frequency range, by adopting co-prime sampling rates. To achieve so, we employed the Chinese Remainder Theorem (CRT) in two different perspectives. For physical applications, the remainder have errors, and undersampling is performed as developed in Chapter II, within the realm of Digital Signal Processing.

Differently, the CRT shown in Chapter IV reassumes its classical form, with no-error remainders, with the goal of improving a cryptographic technique known as Secret Sharing Schemes (SSS). In SSS, a secret is mathematically divided into shares. A minimum amount of t among n share owners must cooperate as a condition to reconstituting the secret, constituting a (n,t) SSS. Considering the Asmuth-Bloom based SSS, which is based on CRT, we have broadened up to 10103 times the size of the secret, when considering the number of bits used in the both models – Asmuth-Bloom's and ours. Our contribution creates obstacles of exponential magnitude for Brute Force Attackers (BFA). Note that the size of the secret can be expanded arbitrarily with our approach, so that the mentioned 10103 times are just an example.

All in all, after more than 5 years, we emphasize the need for bringing back the described solutions to the present context, while highlighting as still promising the proposed cryptographic contributions. The discussion about the sufficiency of the adopted models for personal data protection in the Smart Grid environment is still worth it, given a scenario in which the SM can be the point of vulnerability.

Contents

List of Tables

List of Figures

SYMBOLS

$a, b, c, A, B, C, \sigma, \tau$	scalars
$\mathbf{a}, \mathbf{b}, \mathbf{c}$	vectors
$\mathbf{A}, \mathbf{B}, \mathbf{C}$	matrices
$\boldsymbol{\mathcal{A}}, \boldsymbol{\mathcal{B}}, \boldsymbol{\mathcal{C}}$	tensors
$i \in \{1, 2, \ldots, L\}$	scalar i can assume all values in the set $\{1, 2, \ldots, L\}$
$A = \{1, 2, \ldots, L\}$	set A contains the values $\{1, 2, \ldots, L\}$
$a \bmod b$	the remainder of the division a/b
M_i, r_i	i-th modulo and remainder
$\tilde{r}_i$	i-th erroneous remainder
$\hat{N}$	estimated N
$\overline{a \bmod b}$	modular multiplicative inverse of a modulo b
$d_C(x, y)$	circular distance between x and y over the ring C
$\lfloor a \rfloor$	flooring operation over a
$[\ a\ ]$	rounding operation over a
$\otimes$	Kronecker product
$\times_n$	n-th mode tensor product
$\boldsymbol{\mathcal{A}}_{(n)}$	n-th mode tensor unfolding
$\text{diag}\{\cdot\}$	extract diagonal to a vector
$\text{circshift}(\mathbf{a}, j)$	circular shift of vector $\mathbf{a}$ from the $(j+1)$-th element on
$\text{find}\{\mathbf{a}\}$	extract the cardinalities of the non-zero values in $\mathbf{a}$
$\overline{U}$	average value of variable U
$x_0[t]$	value of variable x_0 at instant time t
ϕ_i	phase of arrival of the impinging waveform at the i-th receiver terminal
P_{rec}	power harvested by rectennas
P_{ant}	power harvested by antennas
$\mathcal{O}(\cdot)$	big O operator
$\odot$	Schur product (entrywise product)
$\sqcup_n$	tensor concatenation along the n-th dimension
$\oplus$	XOR boolean operator

ACRONYMS

ANN	Artificial Neural Networks
ARMAX	Auto-Regressive Moving Average with eXogeneous Inputs Filter
BFA	Brute-Force Attack
BP	Back Propagation Double Layer ANN
CDD	Cooling Degree-Days
CFR-CRT	Closed-Form Robust Chinese Remainder Theorem
CRT	The Chinese Remainder Theorem
DER	Distributed Energy Resource
DFT	Dicrete Fourier Transform
DLP	Discrete Logarithm Problem
DOA	Direction of Arrival
DR	Demand-Response
DSO	Distribution System Operator
ELD	Enthalpy Latent Days
GCD	Greatest Common Divisor
GDP	Gross Domestic Product
HDD	Heating Degree-Days
KF	Kalman Filter
KME-CRT	Kronecker based M-Estimator
LFSR	Linear-feedback shift register
LOS	Line-of-sight
MAPE	Mean Average Percentual Error
ME	M-Estimator
METAR	METeorological Aerodrome Reports
MLE-CRT	Maximum Likelihood Based Robust Chinese Remainder Theorem
MPE	Maximum Percentual Error
MSE	Mean Squared Error
NAN	Neighborhood Area Networks

NEM	Net Energy Metering
NLOS	Non-line-of-sight
PCA	Principal Component Analysis
RF	Radiofrequency (Energy Harvesting)
RX	Receiver
SG	Smart Grids
SM	Smart Meter
SSS	Secret Sharing Scheme
SVD	Singular Value Decomposition
TTP	Trusted Third Party
TX	Transmitter
WSN	Wireless Sensor Network
XOR	Exclusive OR

I

INTRODUCTION

The power grid is a crucial large scale infrastructure. In order to allow high level of automation, information security, distributed energy control and robust load fluctuation management of the power grids, Smart Grids (SG) are essential. Historically in the electricity grid, the flow of electrical power as well as the corresponding consumption measurements and prices have been imposed in a vertical frame, from the companies/producers to the consumers. Nowadays, power, information and prices flow bi-directionally. Ubiquitously produced data in end-points flow upwards in the grid, reporting to the electricity company about all sort of actions [11] [12]. Power is generated inside the boundaries of final user real states — integrating the micro scale power generation — and exported to the company or other consumers [13] [14]. As an illustration, the total worldwide PV panels installations have reached 300 GW by 2016, of which about 28 % are decentralized grid connected worldwide [15]. From those, a great fraction consist of captive clients of the power utility company, which until recently were not allowed to export power. The capability of these clients to trade energy is a matter of increasing interest in the state-of-the-art SG. As a reflex of these phenomena, prices are undergoing a process of decentralization [16] [17], with the possibility of cooperative or non-cooperative frameworks, including auctions [18].

According to the International Energy Agency (IEA), investments in SG by technology area in the past five years have evolved as illustrated in Fig. 1.1, which depicts the investments in SG by technology areas [19]. In absolute values, 2018 was an year of decrease in power equipments expenditures, falling from 135 billion USD in 2017 to 131 billion USD in 2018 worldwide. However, investments in Smart Meters (SM) have been in constant increase in the entire period, from 11 billion USD in 2014 to 19 billion USD in 2019. China is approaching full SM deployment, while Japan, Spain and France are poised to achieve full rollouts in the next few years. In the United States and the European Union, SM have been deployed in over half of the market. Note that the share of digital infrastructure has been rising constantly from 2015 to 2018, which illustrates the tendency of investments in digital equipments and technologies for SG in the short term.

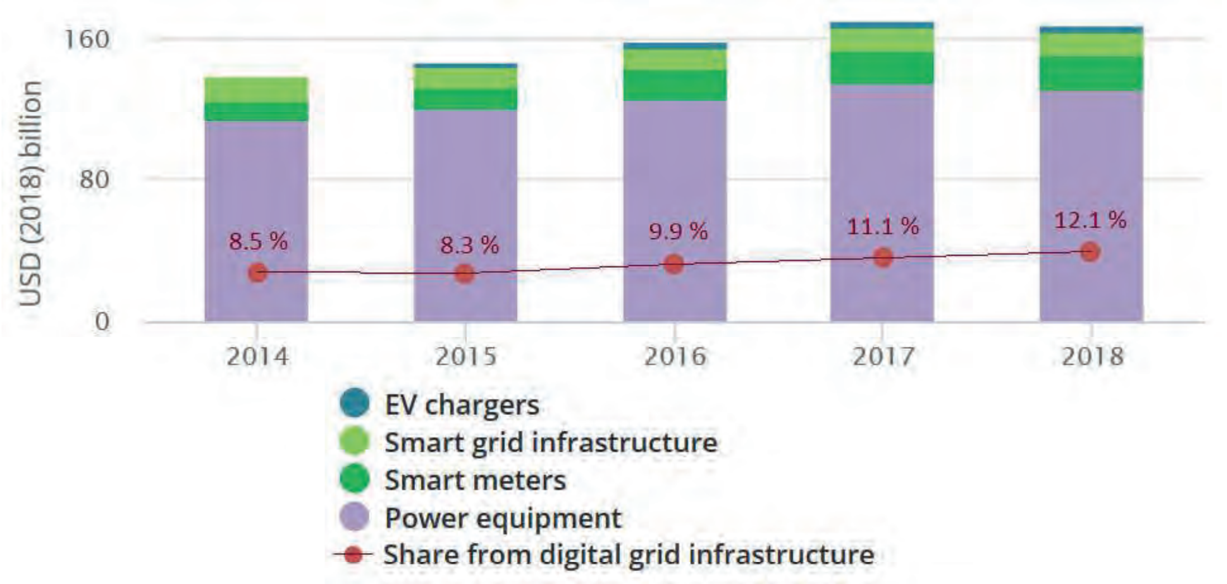


Figure 1.1: Investments in SG by technology areas [19]

Still in accordance with [19], digital energy networks reduce the need to build new power lines and to invest in physical network assets. In this sense, all the ongoing changes in electricity market demand improvements in wireless communication technologies as well as for data protection. In terms of communication technologies, undersampling systems [20] [21] [22], i.e., systems whose sampling rates are below the Nyquist limit [23] [24], provide a reduction of the necessary sampling rates for achieving single-tone signals frequency values. Theoretically, only three samples are sufficient for featuring frequency, phase and amplitude of a wavelength, provided that such samples are close enough. This closeness implies, however, that the sampling rate must be sufficiently high. On the other hand, undersampling systems are suitable for the acquisition of high-frequency signal components that are sparse in the frequency domain when it is not possible to sample such high-frequency components at the Nyquist rate because of limitations of the sampling rate in the hardware in use [25].

As a second SG application, electrical load forecasting is an important instrument in order to make the energy supplied by utilities compatible with the load plus the energy lost in the system. Load forecasting is basically defined as the technique of predicting the future load on a given system for a specified period of time ahead. An accurate load model is required to mathematically represent the relationship between the load and influential variables such as time, weather, economic factors, and so further. The precise relationship between the load and these variables is usually determined by their role in the load model. Based on these premises, we propose a Kalman filter based short term load forecasting system that benefits from known dependencies and data mining techniques to extract optimized sets of input variables from a selection of candidate time series collected from distinct sources. In parallel, we consider that RF energy recycling systems are installed next to the sensors used to measure weather data in order to enhance their autonomy without the need of accessing the nodes. The RF recycling systems use rectennas instead of antennas with adaptive circuits. We show the superiority of employing rectennas to this aim. As a way of proving the feasibility of this solution, we perform a RF incidence measurement campaign in Brasilia, Brazil.

We also address the problem of data protection in local trading systems, as SG are still vulnerable to cyber attacks. Current power grids should be further improved to fit the demands regarding to data security [26] and energy trade between prosumers [27] [28]. There are open issues in the literature with respect to data protection associated with the energy exchange between final clients. For instance, the main efforts towards privacy consists of obfuscating the instantaneous consumption pattern of each consumer [29] [30]. However, the profile of traded energy also delivers relevant information about the prosumers to his neighbors. Models dealing with energy trade directly among prosumers [27] [31] limit themselves to exploit the trade environment without discussing in details data security aspects related to the identities of the traders in relation to their neighbors. In this particular, Secret Sharing Schemes (SSS) are one of the most important cryptographic techniques that allow for sensitive data protection and, as a cryptographic technique, is a topic of continuous improvement.

A common feature between undersampling systems and SSS is the Chinese Remainder Theorem (CRT). With regards to the undersampling systems, we present a novel CRT algorithm that results in improvement towards the correct estimation of the unknown impinging frequency value when compared to the state-of-the-art methods. With regards to SSS, we propose a new form of CRT that enhances drastically the size of the secret in terms of the size of the shares used to hide the sensitive information. It is worth noticing that the CRT applications can be divided into two main types, i.e., stochastic and deterministic applications. In the former, the remainders contain errors originated from physical measurement processes, such as in cognitive radio networks (CRN) [32], polynomial reconstruction [33], electric encoders for motion control [34] and radio interferometric positioning system [35]. Undersampling systems are circumscribed to this type of application and are addressed in Chapter II. On the other hand, in deterministic scenarios, the remainders are always integer numbers, as in cryptography [36] [37], in which the CRT is the basis for the state-of-the-art Asmuth-Bloom's SSS [38] [39]. We present in Chapter IV a novel CRT based SSS that allows for a potentially unlimited ratio between the size of the secret and that of the shares. Note that, in cryptography, error-free remainders are a condition for CRT use. Table I.1 shows the main CRT features and the characteristics of stochastic and deterministic applications in the context of this work.

Table I.1: The Chinese Remainder Theorem Techniques State-of-the-Art Features

	Stochastic	Deterministic
CRT remainders	remainders contain errors	remainders are error-free
Unknown value N	$N \in \mathbb{R}$ is estimated	$N \in \mathbb{Z}$ is calculated
Realm	physical measurement systems	cryptography, modular algebra
Applications	Undersampling systems [20] [40], cognitive radio networks (CRN) [32], radio interferometric positioning system [35]	cryptography [36] [37], secret sharing scheme (SSS) [38] [39]
State-of-the-Art	equal variances: CFR-CRT [41] different variances: MLE-CRT [42]	Traditional CRT [41] [43] Mixed Radix Conversion (MRC) [44]
Content in this work	Chapter II [1]	Chapter IV [2]

Hence, our contributions towards CRT in SG not only relate to applied signal processing solutions, but also to protecting the data content itself. We do not concretely explicit how these tools are used in the SG scenarios, however, we present the overall conditions and features of SG, which let clear the applicability of our CRT based contributions.

This chapter is divided into three sections. In Section 1.1 describes the objectives of this dissertation. In Section 1.2 we describe the main contributions of this work and Section 1.3 shows the used mathematical notation.

1.1 Objectives

As a cyber physical system, the SG demands solutions in terms of data transmission and protection. The broad objective of this work is to offer alternative ways for data secrecy and transmission in the context of SG.

The first aspect is physical, consisting of data processing and transmission issues. Undersampling of single-tone signals is designed as part of the solutions that can be introduced in the SG, as they are suitable for the acquisition of high-frequency signal components that are sparse in the frequency domain. This solution is desirable when it is not possible to sample such high-frequency components at the Nyquist rate due to limitations of the sampling rate in the hardware in use [25]. The implementation of high frequency sampling rates for SM and other commercial equipments in real scenarios for SG may represent some challenges. One of the expected issues is the cost and capacity of data signal samplers belonging to consumers in a Neighborhood Area Networks (NAN), which can be understood as a set of households located in a small urban area. As a first specific objective of this work, we present in [1] a novel CRT technique that outperforms the state-of-the-art CRT methods by estimating the frequency value from impinging waveforms by means of undersampling sensors.

The second goal of this dissertation is related to load forecast is urban areas. For this part of the study, we use data obtained in the city of Brasília, Brazil. The effect of weather on electricity consumption is researched since the first half of the 20th century. Degree-days are used as tool for energy consumption forecast, as showcased in [45, 46]. Currently, heating (HDD) and cooling degree-days (CDD) have been featured in several load forecast methods. Based on these variables, we propose a Kalman filter based short term load forecasting system. Kalman filters benefit from known dependencies and data mining techniques to extract optimized weights of input variables from a selection of candidate time series collected from distinct sources. Note that all used data about humidity, dew point, temperature, weather phenomena and altimeter barometric pressure are obtained via sensors, which are often installed over a relatively great area of difficult access for people. Nevertheless, sensors consume energy continuously, which imposes the achievement of an optimal energetic management. In this context, the concept of energy recycling plays an important role, constituting a parallel objective of this work at this point. Hence, we handle the issue of forecasting load behavior while improving the physical measurement systems, dedicating attention to the energy consumed by sensors. Among several forms of energy recycling, radiofrequency (RF) harvesting has been suggested due to its wide availability mainly in urban areas. We show that RF recycling systems based on rectennas outperform those RF harvesters based on antennas.

Beyond the problem of data transmission and load forecast, there is the issue concerning privacy protection, trading systems and data secrecy against malicious attackers. Therefore, the third specific objective of this work is two-fold: we seek producing a trading system framework for a NAN in which power is traded directly between final nodes, where all the information about the energy bids, such as identities of the bidders, types and number of bids, etc., are kept secret. In parallel, as a cryptographic solution, we introduce the CRT technique that enhances the range of secret values and is used in the framework. In doing so, our SSS [2] is to be applied in the SG environment as a way of protecting sensitive data.

1.2 Overview and Contributions

This thesis is divided into five chapters including this introduction. In Chapter II, we propose a Kronecker based M-Estimation (ME) approach (KME-CRT) [1]. By exploiting the Kronecker product that yields a mapping vector, we drastically reduce the computational complexity of the ME approach, thus allowing its practical application. Due to its routines, the here presented method is specially suitable for CRT systems with few remainders, which is equivalent to networks with few sensors. In our proposal, the errors in the remainders of CRT system may have the same or different variances, allowing our work to be compared with state-of-the-art methods CFR-CRT [41] and MLE-CRT [42]. KME-CRT outperforms both methods in terms of estimation of the real-valued number, according to results tested over 10^5 realizations.

Chapter III proposes a Kalman filter for short term load forecasting system, using sets of input variables from a selection of candidate time series data [3]. We consider that RF energy recycling systems [4] [5] are installed next to the sensors used to measure weather data, in order to enhance their stand by autonomy without the need of accessing the nodes. As a way of proving the feasibility of this solution, we perform a RF incidence measurement campaign in Brasília, Brazil. Inspired by antenna array communication systems, our second contribution in this chapter is to propose an improved RF energy recycling system based on a rectenna array [6]. We show that, by increasing the amount of rectennas, a significant gain due to the array is achieved. We also show that rectenna arrays can outperform standard antenna arrays as energy harvesting systems, quantifying the improvement as a function of the number of employed receiver terminals. This chapter also aims at establishing a connection of this work with our Master Thesis [47], in which we have applied load forecasting over the consumption data of Leipzig, Germany. We have employed Auto Regressive Moving Average with eXogeneous inputs (ARMAX) filters combined with micro-generators such as indoor light energy recycling and RF harvesting circuits to determine the relevance of each chosen power recycling resource.

Chapter IV presents the problem of providing data privacy for self-interested players that trade energy in the context of a NAN. The energy is sold by local micro-generators and locally purchased by their neighbors, also known as the final users. As in this chapter we deal with the problem of energy trade, we indirectly aggregate the scope of our recent works in [4–7], in which we exploit alternative energy resources with commercial aspects, as well as with load forecasting models [3] and communication architectures [8] for integrated power generators in urban environments. Furthermore, as individual energy producers must cope with island-mode operation due to security reasons, in [9] we propose a patent-pending island-mode detector system that disconnects the prosumer automatically when there is an outage in the grid side. Our framework in [10] integrates all these previous works as it deals simultaneously with SG data security requirements and energy trade systems. As a first contribution, the proposed framework has a privacy-preserving model which has a low computational complexity and avoids completely an unauthorized party to identify the bidders, the number or types of them, and even if the bids achieve or not a deal. As a second contribution, all the bids are made clear to the NAN participants, with all SM owners having access to how many bids are proposed, their types, prices and quantities. Nevertheless, the proposed framework avoids totally any access to the bidders identities. In order to improve data secrecy robustness, the chapter describes a perfect CRT based SSS that has a potentially unlimited ratio between the sizes of the shares and the secret than in the Asmuth-Bloom's SSS. Hence, in [2] the secret is an integer $S \in \{0, 1, \ldots, (k-1)!\}$ that is associated with each permutation of k elements of a vector $\mathbf{s} \in \mathbb{Z}^k$. A bijective relationship between S and the permutation in $\mathbf{s}$ is guaranteed by means of the Lehmer Code.

As a summary of the above, Fig. 1.2 illustrates the correspondence between SG features and the respective chapters. Chapter II presents an innovation for the state of the art of CRT for single-tone signals. This innovation can be used for by any communication link that makes use of undersampling technique (a). Chapter III presents a study based on load forecast (b) usually carried out by the electricity company. Sensors used in the data collection are energized via RF energy harvesting systems (c). Chapter IV bears the problem of energy trade in a NAN, where (d) the dealer is a Trusted Third party (TTP) that processes the bids forwarded by (e) the prosumers with a SM. The cryptographic technique of SSS is also a matter of innovation. Finally, Chapter V draws the conclusions about this work.

1.3 Notation

Along this work, the following notation is used. Scalars are denoted by lower-case letters $(a, b, \cdots)$, upper-case letters $(M, N, \cdots)$ and Greek letters $(\sigma, \mu, \cdots)$. Vectors are written as boldface lower-case letters $(\mathbf{a}, \mathbf{b}, \cdots)$, matrices as boldface capitals $(\mathbf{A}, \mathbf{B}, \cdots)$, and tensors as boldface calligraphic letters $(\boldsymbol{\mathcal{A}}, \boldsymbol{\mathcal{B}}, \cdots)$. The notation $\mathbf{A}(:, i) \in \mathbb{C}^{R \times 1}$ represents a column vector denoting the i-th column of $\mathbf{A} \in \mathbb{C}^{R \times I}$. The operator $\mathrm{vec}(\mathbf{A})$ results in a vector by concatenating the columns of the matrix $\mathbf{A}$ one on top of the other. The notation $[\boldsymbol{\mathcal{T}}]_{(r)}$ is the r-mode matrix unfolding of $\boldsymbol{\mathcal{T}}$ and $\boldsymbol{\mathcal{T}} \times_r \mathbf{A}$ is the $r-$th mode product between the tensor $\boldsymbol{\mathcal{T}}$ and the matrix $\mathbf{A}$. Moreover, the Kronecker product and outer product operators are denoted by $\otimes$ and $\circ$, respectively. The operator $\mathrm{E}\{\cdot\}$ stands for the expected value operation.

Sets of elements are signalized with calligraphic font as $\mathcal{N}$. Estimated numbers are written

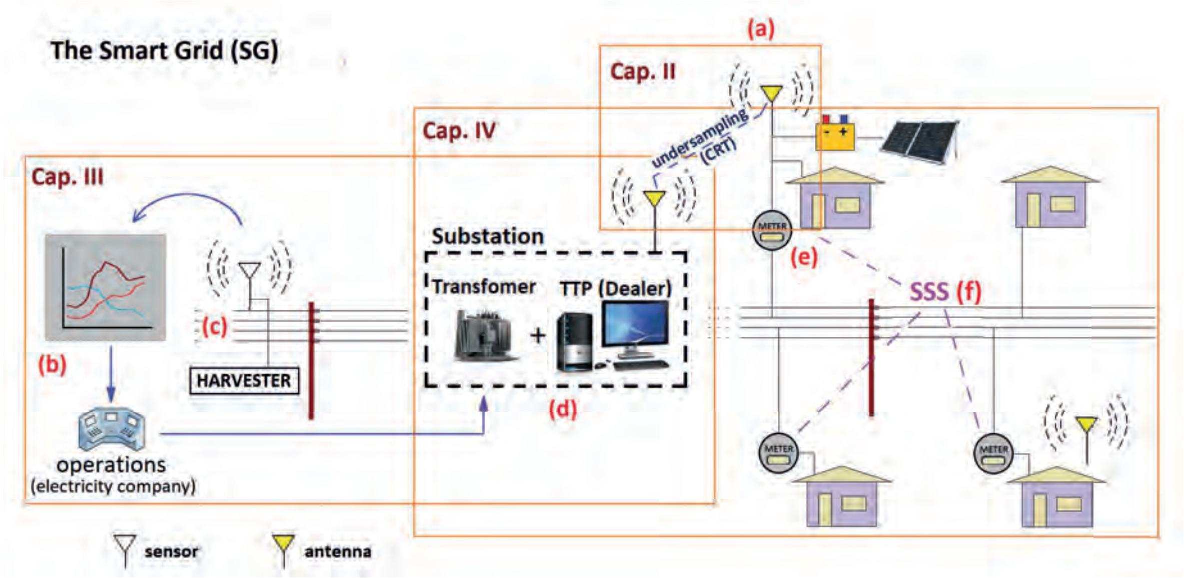


Figure 1.2: Summary of this work in terms of the SG features and the respective chapters

as in $\hat{N}$, whereas variables that contain errors are notated as in $\tilde{r}$. Vectors of estimations and of values with errors are notated respectively as in $\hat{\mathbf{n}}$ and $\tilde{\mathbf{r}}$. We also assume that, differently from the rigorous mathematical standpoint, non-integer values have the modulus operation applicable only on its integer part. Thus, when a real-valued $N = N_{\text{int}} + N_{\text{dec}}$, where the parcels refer respectively to the integer and decimal parts of N, we assume that $N \bmod M_i = N_{\text{int}} \bmod M_i + N_{\text{dec}}$.

We introduce the Kronecker product based on [48]. Let $\mathbb{S}^{p\times q}$ denote the space of real or complex matrices. The (i, j)-th entry of a matrix $\mathbf{A} \in \mathbb{S}^{p\times q}$ is a_{ij}. The Kronecker product is defined for two matrices of arbitrary sizes over any ring. For instance, consider matrices $\mathbf{A} \in \mathbb{S}^{p\times q}$ and $\mathbf{B} \in \mathbb{S}^{u\times v}$. Their Kronecker product is defined as

$$\mathbf{A} \otimes \mathbf{B} = \begin{bmatrix} a_{11}\mathbf{B} & \dots & a_{1q}\mathbf{B} \\ \vdots & & \vdots \\ a_{p1}\mathbf{B} & \dots & a_{pq}\mathbf{B} \end{bmatrix} \in \mathbb{S}^{pu\times qv}, \tag{1.1}$$

where the symbol "$\otimes$" stands for the defined Kronecker product.

The Schur product is an element-wise product between two tensors $\mathcal{A}$ and $\mathcal{B}$ of same dimensions. The product $\mathcal{C} = \mathcal{A} \odot \mathcal{B}$ yields

$$\mathcal{C}(a_1, a_2, \dots, a_n) = \mathcal{A}(a_1, a_2, \dots, a_n) \cdot \mathcal{B}(a_1, a_2, \dots, a_n), \tag{1.2}$$

for all a_j, $j \in \{1, 2, \dots, n\}$. Therefore, each entry in $\mathcal{C}$ is the product of the corresponding entries of tensors $\mathcal{A}$ and $\mathcal{B}$.

II

IMPROVED MATRIX AND TENSOR BASED M-ESTIMATOR FOR THE CHINESE REMAINDER THEOREM

The Chinese Remainder Theorem (CRT) explains how to solve an algebra problem in which an integer-valued N is determined from its remainders, as in

$$\begin{cases} N \text{ mod } M_1 = r_1, \\ N \text{ mod } M_2 = r_2, \\ \quad \vdots \qquad \vdots \qquad \vdots \\ N \text{ mod } M_L = r_L, \end{cases} \tag{2.1}$$

where $M_i = \{M_1, M_2, \ldots, M_L\}, 0 < M_1 < M_2 < \cdots < M_L$, are the L moduli, and mod stands for the modulus operator. The remainders are r_i, for $i = \{1, 2, \ldots, L\}$. All remainders respect $0 \leq r_i < M_i$ for $i = \{1, 2, \ldots, L\}$. Given any i-th row of (2.1), an equivalent expression is

$$n_i M_i + r_i = N, \tag{2.2}$$

where the n_i, for $i = \{1, 2, \ldots, L\}$ are the folding integers, also unknown. From the knowledge of M_i and r_i, the CRT offers the straightforward calculation integer-valued N when the remainders are free of errors [41] [43] [49] [50].

Particularly in signal processing, the CRT given by the system in (2.1) can be used to solve the problem of estimating the frequency of a desired signal in undersampling systems [20] [21] [22], i.e., systems whose sampling rates are co-prime and all below the Nyquist limit [40] [51]. The frequency is estimated in terms of their remainders given the sampling rates.

For illustration, consider that a single-tone signal whose frequency value is $f = 2177$ MHz is sampled by four distinct synchronized sensors, whose sampling rates are respectively $F_{s,1} = 11$ MHz, $F_{s,2} = 13$ MHz, $F_{s,3} = 15$ MHz and $F_{s,4} = 17$ MHz. The four sampling rates form a co-prime system since the greatest common divisor (GCD) of any pair $\text{GCD}\{F_{s,i}, F_{s,j}\} = 1$, for $i \neq j$. Hence, the peaks in the DFT taken by every sensor are respectively $\{10, 6, 2, 1\}$, which are the remainders in

$$\begin{cases} f \bmod 11 = 10, \\ f \bmod 13 = 6, \\ f \bmod 15 = 2, \\ f \bmod 17 = 1, \end{cases} \tag{2.3}$$

where the moduli M_i are the sampling rates. For instance, in the undersampling system whose CRT scheme is given by (2.3), the sensor of 15 MHz reads exactly the same sequence of snapshots of a hypothetical impinging waveform of 2 MHz, since the same vector of snapshots is obtained with the frequency value of $f = 2177$ MHz or $f = 2$ MHz when sampled at $F_{s,3} = 15$ MHz. The main goal of undersampling schemes is reducing the necessary sampling rates for achieving the values of frequency. Theoretically, only three samples are sufficient for featuring frequency, phase and amplitude of a wavelength, provided that such samples are close enough. This closeness implies that the sampling rate must be sufficiently high. Undersampling systems, on the other hand, are suitable for the acquisition of high-frequency signal components that are sparse in the frequency domain when it is not possible to sample such high-frequency components at the Nyquist rate because of limitations of the sampling rate in the hardware in use [25]. For instance, in [52], with a monostatic synthetic aperture radar (SAR) for terahertz (THz), the wavelength is in the order of millimeter or submillimeter, and hence the requirement on subwavelength interval of spatial sampling by the Nyquist theory aggravates the measurement difficulty. Sparsity, in terms of CRT based undersampling systems, means the existence of a single DFT peak that is easily identifiable in the surveyed spectrum. This implies that each sensor must be able to identify a single peak in their undersampling scales, which for practical purposes means that there is only one single-tone impinging waveform. Note that if each sensor reads two peaks, the result is a relevant reduction of the dynamic range, i.e., the interval in which the unknown frequency value is uniquely determinable. If the dynamic range for a single unknown is D, two unknown numbers reduce it to $2\sqrt{D}$ [53]. When the number of DFT peaks surpasses two, the dynamic range is lowered progressively at each new peak [54]. Therefore, the CRT solution presented in this chapter presupposes the identification of a single-tone signal read by each sensor.

Clearly, measurements in the realm of digital signal processing are liable to jitter and phase noise as part of real applications. These errors affect the remainders in a CRT originated by an undersampling system. It is important to notice that any deviations in the remainders cannot be reduced or improved by the CRT system itself. A better CRT technique can, however, estimate the unknown number - in the case of undersampling systems, the frequency - with better approximation given the mentioned pre-existing errors.

In engineering applications, either the unknown is an integer or a real-number, and errors may exist or not. In the latter, we have deterministic applications, related for instance to cryptography [36] [37], Digital Signature Standard (DSS) [55], image processing and security [56] [57], secret sharing schemes [38] [39] and E-Voting systems [58]. However, not only the remainders have errors, but also the unknown value is a real-valued number. Hence, beyond undersampling systems, CRT is also employed to estimate unknown numbers in the presence of errors such as in cognitive radio networks (CRN) [32], polynomial reconstruction [33]-[59], electric encoders (EE) for motion control [34], and radio interferometric positioning system [35]. Another signal processing application of CRT is related to phase unwrapping based systems for distance estimation [49] [50], where the remainders stand for the phase of arrival in terms of wavelength, and the moduli represent the wavelengths of each component.

The state-of-the-art approaches for CRT estimation include the traditional CRT [41] [43], the robust CRT [41] [60] [61] [62], the closed-form robust CRT [41] [49] and the maximum likelihood based robust CRT [63] [42]. In the latter, an optimization of the search routine for a real-valued number is proposed assuming Gaussian distributed errors with different variances, whereas in the closed-form robust CRT the variances are presumed constant. There is still the Multi-Stage Robust CRT, which is proposed in [50] and consists in splitting the moduli of a CRT system over different moduli groups in accordance with the GCD of each set. The number of resulting groups is the number of stages. In [64], a generalization of the two-stage robust CRT algorithm to a multi-stage system is also presented. Splitting the moduli in groups with different GCD by group can improve the remainder error bound for a given set of moduli in terms of the remainder error bound of the entire CRT system. However, such a split is based on adopting for each moduli set the same concepts of the CFR-CRT. When all moduli share the same pairwise GCD as presumed in (2.1), applying the Multi-Stage Robust CRT does not improve the estimation of N in terms of the maximum tolerable error.

In this chapter, we propose a Kronecker based M-Estimation (ME) approach (KME-CRT). By exploiting the Kronecker product that yields a mapping vector, we drastically reduce the computational complexity of the ME approach allowing its practical application. Due to its routines, the here presented method is specially suitable for CRT systems with few remainders, which is equivalent to networks with few sensors. Furthermore, our proposed technique enhances the probability of estimating an unknown number accurately even when the errors in the remainders surpass $1/4$ of the greatest common divisor of all moduli. We also provide a version of the mapping vectors based on tensorial n-mode products, delivering in the end the same information as the original method.

The remainder of this chapter is organized as follows. The CRT systems and the state-of-the-art Closed-Form Robust CRT (CFR-CRT) and Maximum Likelihood Estimator Based Robust CRT (MLE-CRT) are reviewed in Section 2.1. Section 2.2 presents the proposed KME-CRT for assembling the mapping vector, along with its tensorial versions. Simulations and results are presented in Section 2.3, and Section 2.4 concludes the chapter.

2.1 The State-of-the-Art CRT Based Techniques

In practical applications, differently from solving (2.1), CRT problems consist in estimating a real-valued N with erroneous remainders as in

$$\begin{cases} N \bmod M_1 + \Delta_1 = \tilde{r}_1, \\ N \bmod M_2 + \Delta_2 = \tilde{r}_2, \\ \quad \vdots \quad\; \vdots \quad\; \vdots \\ N \bmod M_L + \Delta_L = \tilde{r}_L, \end{cases} \tag{2.4}$$

where $\tilde{r}_i = r_i + \Delta_i$, for $i \in \{1, 2, \ldots, L\}$, with Δ_i denoting the deviation or error in the i-th remainder originated from noise or inaccuracy in measurement. All remainders respect $0 \leq r_i < M_i$ for $i \in \{1, 2, \ldots, L\}$. Solving (2.4) is generally far more complex than (2.1). CRT is not a robust system due to the fact that small errors in any remainder may cause a large reconstruction error [41]. Note that, given (2.2), we can also write

$$n_i \; M_i + \tilde{r}_i - \Delta_i = N. \tag{2.5}$$

A variable τ is defined as the remainder error bound, or the maximum absolute value for every existing error Δ_i, hence $\tau = \max\limits_{i \in \{1,\ldots,L\}} |\Delta_i|$. Assuming M as the GCD of all moduli M_i, provided that

$$\tau < \frac{M}{4}, \tag{2.6}$$

the calculation of the folding integers n_i is guaranteed [41] [65]. The estimated values of n_i in CFR-CRT are designed as $\hat{n}_i$ and $\hat{N}$ is the estimation for N.

The dynamic range D of a CRT system delimits the value until which N can be uniquely determined [53] [54]. As a consequence, the search for the value of N is performed only in the range of D. If all moduli M_i are co-prime such that $M = 1$, $D = \prod_{i=1}^{L} M_i$. Otherwise, if $M > 1$, let

$$\Gamma_i = \frac{M_i}{M}, \quad i \in \{1, 2, \dots, L\}, \tag{2.7}$$

so that all Γ_i, for $i \in \{1, 2, \dots, L\}$, are co-prime, and the dynamic range D is given by

$$D = M\Gamma, \tag{2.8}$$

where $\Gamma = \prod_{i=1}^{L} \Gamma_i$.

In this chapter, the goal is to estimate a desired real-valued N given the information about M_i and $\tilde{r}_i$, for $i \in \{1, ..., L\}$. One of the state-of-the-art CRT techniques is the Robust CRT [60] [61], whose main disadvantage is that the order of $2(L-1)\Gamma_i$ searches is necessary even in the 1-D searching scheme. As a consequence, when L or Γ_i gets large, the searching complexity is still high [41] [50]. Therefore, we consider the CFR-CRT and the MLE-CRT in terms of benchmark to compare with our proposed approach. In the CFR-CRT, the variances of the errors are presumed equal, whereas MLE-CRT addresses scenarios in which the variances of errors are different and known at prior.

Note that CRT is a deterministic problem and that, given a remainder error bound τ, all errors are confined to a fixed interval $[N - \tau, N + \tau]$. In the following state-of-the-art methods, the errors are presumed to have Gaussian distribution. In a Gaussian distribution, any error that surpasses τ can be arbitrarily high with non-zero probability, thus violating the assumption of the existence of τ. In order to keep the Gaussian distribution yet with reasonable error profile, we assume that $\tau \approx 3\sigma$, hence assuring that more than 99% of the values of Δ_i lie in the interval $[-\tau, +\tau]$. In doing so, we also preserve the possibility of changing the variances of errors by modifying the value of τ.

2.1.1 Closed-Form Robust CRT

According to [41] [49] [64], the CFR-CRT is summarized in (2.9)-(2.17) and Algorithm 6. First an auxiliary variable γ_i is defined using Γ_i from (2.7) and Γ from (2.8) as follows,

$$\gamma_i = \frac{\Gamma}{\Gamma_i}, \quad 1 \leq i \leq L. \tag{2.9}$$

Note that all possible pairs $\{\gamma_i, \Gamma_i\}$ are co-prime. The modular multiplicative inverse (MMI) of a number γ_i modulo Γ_i is the smallest number $\overline{\gamma}_i$ that satisfies $\gamma_i\overline{\gamma}_i = k\Gamma_i + 1$, for some $k \in \mathbb{Z}$. In [65], the Qin's Algorithm for the calculation of the MMI by means of a fast matrix based technique is presented. We notate the MMI function as

$$\overline{\gamma}_i = \overline{\gamma_i \bmod \Gamma_i}. \tag{2.10}$$

According to [41], the co-primality between $\{\gamma_i, \Gamma_i\}$ assures the existence of a MMI $\overline{\gamma}_i$. Next, we define q_i:

$$q_i = \left\lfloor \frac{\tilde{r}_i}{M} \right\rfloor, \tag{2.11}$$

where $\lfloor . \rfloor$ stands for the flooring operation. N_0 is defined as

$$N_0 = \sum_{i=1}^{L} (\overline{\gamma}_i \gamma_i q_i) \bmod \Gamma. \tag{2.12}$$

In fact, the sequence (2.9)-(2.12) is used to calculate integer-valued N when the remainders are free of errors based on $N = MN_0 + r^c$, where $r^c = r_i$, for $i \in \{1, 2, \ldots, L\}$, stands for the common remainder in the error-free case. This method receives different names in the literature, such as Conventional CRT [41] [60] [66] [53], Gauss's Algorithm [67] [68], CRT Standard version [58] and Classical CRT formula [63] [42]. In [43], it is called Traditional CRT, as well as in [41] [63] [42]. We adopt the latter terminology. The Traditional CRT is based on the extended Euclidean algorithm [65] and is part of the CFR-CRT in [41].

It is still worth noting that, according to the CRT theory,

$$N_0 = n_i \Gamma_i + q_i, \tag{2.13}$$

for $i \in \{1, 2, \ldots, L\}$, i.e., N_0 is an integer-valued number whose value is obtained by means of (2.13) independently of the chosen i. The CFR-CRT is developed with sequential subtractions of $(n_i \Gamma_i + q_i)$ from a reference remainder whose row is z, yielding $L - 1$ results that are generated by applying

$$n_z \Gamma_z - n_i \Gamma_i = q_{i,z} \tag{2.14}$$

where $q_{i,z} = q_i - q_z$, for $i \in \{1, 2, \ldots, z-1, z+1, \ldots L\}$. In order to choose the reference remainder, first we define the circular distance of two real numbers x and y for a non-zero positive number C as

$$d_C(x, y) \triangleq x - y - \left[\frac{x - y}{C} \right] C. \tag{2.15}$$

In [41], the z-th row, which is the row of the reference remainder, is obtained via

$$\hat{r}^c \triangleq \underset{0 \leq m \leq M-1}{\arg\min} \sum_{i=1}^{L} d_M^2(\tilde{r}_i^c, m) \tag{2.16}$$

where $\tilde{r}_i^c = \tilde{r}_i \bmod M$, for $i \in \{1, 2, \ldots, L\}$, and

$$z = \underset{j \in \{1,2,\ldots,L\}}{\arg\min}\, d_M^2(\tilde{r}_j^c, \hat{r}_c). \tag{2.17}$$

The estimation of $\hat{N}$ obtained from CFR-CRT according to [41] is summarized in Algorithm 6, where the results of (2.7)-(2.17) are presumed available. In line 3 of Algorithm 6, [.] stands for the rounding operator.

Algorithm 1 State-of-the-art technique: CFR-CRT

```
 1: procedure CFR-CRT ($M_i, \tilde{r}_i$)
 2:     for $i = 1 : L, i \neq z$ do
 3:         $\hat{q}_{i,z} \leftarrow \left[\frac{\tilde{r}_i - \tilde{r}_z}{M}\right]$   % The reference remainder $\tilde{r}_z$ is chosen by means of
 4:                           % (2.15)-(2.17), and $\hat{q}_{i,z}$ follows the definition of (2.14)
 5:         $\overline{\Gamma}_{i,z} \leftarrow \overline{\Gamma_z \mathrm{mod} \Gamma_i}$   % MMI operation as specified in (2.10).
 6:                              % Note that all $\Gamma_i$ are defined in (2.7)
 7:         $\hat{\xi}_{i,z} \leftarrow (\hat{q}_{i,z}\overline{\Gamma}_{i,z}) \bmod \Gamma_i$
 8:         $b_{i,z} \leftarrow \overline{\gamma_z/\Gamma_i \bmod \Gamma_i}$   % Recall that all $\gamma_i$ are computed as in (2.9)
 9:     $\hat{n}_z \leftarrow \sum_{i=1, i\neq z}^{L} (\hat{\xi}_{i,z} b_{i,z} \frac{\gamma_z}{\Gamma_i}) \bmod \gamma_z$
10:     for $i = 1 : L, i \neq z$ do
11:         $\hat{n}_i \leftarrow \frac{\hat{n}_z \Gamma_z - \hat{q}_{z,1}}{\Gamma_i}$
12:     $\hat{N} \leftarrow \frac{1}{L} \sum_{i=1}^{L} (\hat{n}_i M_i + r_i)$
```

The selection of the optimal reference remainder $\tilde{r}_z$ is based on the reference common remainder $\tilde{r}_i^c$, which can only be appropriately determined when $\sigma_1^2 = \sigma_2^2 = \cdots = \sigma_L^2$. When σ_i^2 are different for each Δ_i, for $i \in \{1, 2, \ldots, L\}$, the calculations based on $\tilde{r}_z$ as in (2.16) and (2.17) may be ineffective [63].

Note that, although the Traditional CRT is part of the CFR-CRT, the latter is the method that makes the estimation of N achievable when the remainders have errors. CFR-CRT can be used for integer-valued N free of errors; however, in this case, it consists in a simple application of the above commented Traditional CRT.

2.1.2 Maximum Likelihood Based Robust CRT

Aiming to solve the case of different variances in the errors, a Maximum Likelihood based CRT is proposed in [63] [42]. MLE-CRT is basically a method for determining the best $\hat{r}^c$, the estimation of the common remainder r^c.

In [63] [42], the standard deviations are $\sigma_i = \mu M_i$, for $i \in \{1, 2, \dots, L\}$, where μ is a small arbitrary positive factor. A set Ω is then assembled as

$$\Omega = \left\{ \left(\sum_{i=1}^{L} w_i \tilde{r}_i^c + M \sum_{i=1}^{t} w_{\rho(i)} \right) \mod M \right\}, \text{ for } t \in \{1, 2, \dots, L\}, \tag{2.18}$$

where ρ is a permutation of the set $\{1, 2, \dots, L\}$ such that $\tilde{r}^c_{\rho(1)} \leq \dots \leq \tilde{r}^c_{\rho(L)}$, and

$$w_i = \frac{1/\sigma_i^2}{\sum_{i=1}^{L} 1/\sigma_i^2}. \tag{2.19}$$

The estimated common remainder $\hat{r}^c$ is then achieved by

$$\hat{r}^c \triangleq \arg\min_{x \in \Omega} \sum_{i=1}^{L} w_i d_M^2(\tilde{r}_i^c, x) \tag{2.20}$$

Algorithm 2 shows the MLE-CRT, whose input arguments are M_i, $\tilde{r}_i$ and μ.

2.2 Kronecker Product Based Mapping Vector

In Section 2.2.1, we exploit the CRT system with error-free remainders, with $N \in \mathbb{Z}$ and $\Delta_i = 0$, for $i \in \{1, 2, \dots, L\}$, in which N is not estimated but rather calculated in a deterministic way. Next, Section 2.2.2 handles the case of remainders with errors, when $\hat{N}$, the estimated N, is obtained. In Section 2.2.3, we provide a tensorial model based on n-mode products for delivering the same information of the mapping vector with regards to the error-free case of Section 2.2.1 and the remainders with errors of Section 2.2.2. Section 2.2.4 presents the study of how the proposed mapping vector enables correct estimations $\hat{N}$ even when $M/4 \leq \tau < M/2$. Recall that, when any error surpasses $M/4$, the reconstruction of the folding integers $\hat{n}_i$ is not guaranteed in accordance with the literature.

If an ME routine is performed over the entire dynamic range D in order to find the most appropriate value of $\hat{N}$ that minimizes all deviations with regards to the remainders, the result is a computationally expensive task. In order to mitigate this hindrance, we propose a mapping vector $\mathbf{v}$ that indicates on which parts of D the search for N should be made. From the knowledge of the values M_i and $\tilde{r}_i$, L auxiliary vectors $\mathbf{c}_i$, for $i \in \{1, 2, \dots, L\}$ are assembled, which jointly yield $\mathbf{v}$.

Algorithm 2 State-of-the-art technique: MLE-CRT

```
procedure MLE-CRT($M_i, \tilde{r}_i, \mu$)
    for $i = 1 : L$ do
        $\tilde{r}_i^c \leftarrow \tilde{r}_i \bmod M$
        $\sigma_i \leftarrow \mu M_i$
    for $i = 1 : L$ do
        $w_i \leftarrow \frac{1/\sigma_i^2}{\sum_{i=1}^{L} 1/\sigma_i^2}$
    $\tilde{\mathbf{r}}' \leftarrow \begin{bmatrix} \tilde{r}_1 & \tilde{r}_2 & \dots & \tilde{r}_L \end{bmatrix}$
    $\mathbf{w}' \leftarrow \begin{bmatrix} w_1 & w_2 & \dots & w_L \end{bmatrix}$
    $\mathbf{R_W} \leftarrow \begin{bmatrix} \tilde{\mathbf{r}} & \mathbf{w} \end{bmatrix}$
    $\mathbf{R_W} \leftarrow \text{sortrows}(\mathbf{R_W})$
    for $i = 1 : L$ do
        for $t = 1 : i$ do
            $\Omega(i) \leftarrow \left\{ \left( \sum_{j=1}^{L} w_j \tilde{r}_j^c + M \sum_{j=1}^{t} \mathbf{R_W}(j, 2) \right) \bmod M \right\}$
    $\hat{\mathbf{r}} \leftarrow \text{zeros}[L \times 1]$
    for $i = 1 : L$ do
        $x \leftarrow \Omega(i)$
        $\hat{\mathbf{r}}(i) \leftarrow \sum_{j=1}^{L} w_j d_M^2(\tilde{r}_i^c, x)$
    $[\sim, \text{ index}] \leftarrow \min(\hat{\mathbf{r}})$
    $\hat{r}^c = \Omega(\text{index})$
    for $i = 1 : L$ do
        $\hat{q}_i \leftarrow \left[ \frac{\hat{r}_i - \hat{r}^c}{M} \right]$
    $\hat{N}_0 \leftarrow \sum_{i=1}^{L} \overline{\gamma_i} \gamma_i \hat{q}_i \bmod \Gamma$
    $\hat{N} \leftarrow M\hat{N}_0 + \hat{r}^c$
```

As it will be shown in Algorithm 3 of Section 2.2.2, our M-Estimator for a given $\hat{N}$ is based on the minimization of θ in (2.21). Using the circular distances of (2.15), for the case of all variances with the same value, the estimator is given by

$$\theta = \arg\min_{\hat{N}} \sum_{i=1}^{L} (d_{M_i}(\hat{N} \bmod M_i, \tilde{r}_i))^2, \tag{2.21}$$

whereas if the variances are a function of the respective modulus M_i,

$$\theta = \arg\min_{\hat{N}} \sum_{i=1}^{L} \left(\frac{d_{M_i}(\hat{N} \bmod M_i, \tilde{r}_i)}{M_i} \right)^2, \tag{2.22}$$

where the values of $\hat{N}$ are to be selected according to the content of the mapping vector $\mathbf{v}$.

Note that either in (2.21) and (2.22) the intermediate estimation of the folding integers $\hat{n}_i$ is avoided, differently of the state-of-the-art CFR-CRT and MLE-CRT.

2.2.1 Data Structure for Remainders Without Errors

In this Section, we address the problem of calculating $N \in \mathbb{Z}$ and $\Delta_i = 0$, for $i \in \{1, 2, \dots, L\}$. Prior to the CRT system itself, we explain how N can be determined by the remainders and moduli under the perspective of group theory as a way of featuring the proposed method. In any CRT system, each of the L rows informs the possible values for N over the dynamic range D. We then assemble the sets S_i, for $i \in \{1, 2, \dots, L\}$, where the k-th component is of the form

$$r_i + (k-1)M_i, \tag{2.23}$$

for $k \in \{1, 2, \dots, \gamma_i\}$. Therefore, the values in each set S_i represent the sufficient and necessary set of possibilities for N in the dynamic range respecting the condition $N \bmod M_i = r_i$. Hence,

$$S_i = \{r_i,\ (r_i + M_i),\ (r_i + 2M_i),\ \dots,\ (r_i + (\gamma_i - 1)M_i)\}, \tag{2.24}$$

and we can write

$$\cap_{i=1}^{L} S_i = \{N\}, \tag{2.25}$$

which is equivalent to stating that N is the unique number that figures simultaneously in all sets S_i, for $i \in \{1, 2, \dots, L\}$.

Proposition 1. *The set of values in each S_i specified as in (2.24) not only cover all the possibilities of values for the unknown N, given M_i and r_i over D, but also N cannot be excluded from any set S_i.*

Proof of Sufficiency. In case of Proposition 1, sufficiency is proven if and only if any additional insertions of $(r_i{+}kM_i)$ terms beyond the limit shown in (2.24), i.e., $(r_i{+}(\gamma_i{-}1)M_i)$, lead to repetition of terms in the set S_i. Let $(r_i{+}kM_i)$, $k \in \{0, 1, \dots, (\gamma_i{-}1)\}$ be the $(k{+}1)$-th possible value for N in the set S_i. We show that extending the content of S_i by inserting a k'-th entry, where

$$k' = q_1\gamma_i + q_2, \tag{2.26}$$

where $q_1 \in \mathbb{Z}^+$ and $q_2 \in \mathbb{Z}$, entails repetition of values in S_i, since the inserted values are of the form

$$r_i + (q_1\gamma_i + q_2)M_i = r_i + q_1 D + q_2 M_i, \tag{2.27}$$

as $M\gamma_i\Gamma_i = D$. Given the definition of the dynamic range, $(r_i+D+q_2M_i) \bmod D = r_i+q_2M_i$, if $q_2 < D/M_i$. If, otherwise, $q_2 \geq D/M_i$ in (2.27), then

$$\text{mod } (k'M_i, D) = \text{mod } (q_2M_i, D) = \text{mod } (D + q_2'M_i, D) = \text{mod } (q_2'M_i, D), \tag{2.28}$$

so that it is possible to substitute $q_2'M_i = k'$ and recalculate (2.26) until $q_2 < D/M_i$ in (2.27). Note that, since $q_2 \in \mathbb{Z}$, D/M_i is also an integer, it is equivalent to state that $q_2 \leq \gamma_i - 1$, as $D/M_i = \gamma_i$. Therefore, the dynamic range is surpassed given a set of values for S_i if any arbitrary value $(r_i + k'M_i)$, with $k' \in \{\gamma_i, \gamma_i + 1, \dots\}$, is included in S_i. Due to the minimum values that q_1 and q_2 can assume in (2.26), namely $q_1 = 1$ and $q_2 = 0$, no further elements are to be included in (2.24), proving the limit $k' = \gamma_i$. As a consequence, each set S_i has γ_i terms. The smallest value for S_i cannot be lower than r_i due to the non-admission of negative numbers. Hence, the analysis of the boundary to the left is dismissed.

Proof of necessity is the requirement that no possible value of N can be excluded from at least one set S_i, $i \in \{1, 2, \dots, L\}$. This proof is straightforward, since $N \bmod M_i = r_i$, $k \in \{0, 1, \dots, \gamma_i - 1\}$. □

Defining $\mathbf{e}_i \in \mathbb{Z}^{M_i}$ as

$$\begin{cases} \mathbf{e}_i(p) = 1, & \text{if } p = r_i \text{ and } r_i \neq 0, \\ \mathbf{e}_i(p) = 1, & \text{if } p = M_i \text{ and } r_i = 0, \\ \mathbf{e}_i(p) = 0, & \text{otherwise,} \end{cases} \tag{2.29}$$

and the vector $\mathbf{u}_k$ as

$$\mathbf{u}_k \in \mathbb{Z}^k, \mathbf{u}_k(j) = 1 \text{ for } j \in \{1, 2, \dots, k\}, \tag{2.30}$$

each column-vector $\mathbf{c}_i \in \mathbb{Z}^D$ is obtained via Kronecker product as

$$\mathbf{c}_i = \mathbf{w}_i \otimes \mathbf{e}_i, \tag{2.31}$$

where, defining the auxiliary set $\mathcal{Y}_i \in \{1, \dots, i-1, i+1, \dots, L\}$,

$$\mathbf{w}_i = \mathbf{u}_{\Gamma_{\mathcal{Y}_i(1)}} \otimes \mathbf{u}_{\Gamma_{\mathcal{Y}_i(2)}} \otimes \cdots \otimes \mathbf{u}_{\Gamma_{\mathcal{Y}_i(L-1)}}, \tag{2.32}$$

which is the same as writing

$$\mathbf{w}_i = \mathbf{u}_{\gamma_i}. \tag{2.33}$$

Carrying out the product in (2.31) is the same as assembling a vector which is formed by γ_i stacked vectors $\mathbf{e}_i$. Fig. 2.1 shows the concept for visualization with an example where $L = 3$. Note that the length of the resultant $\mathbf{c}_i$ is $M_i\gamma_i = M\Gamma = D$, for $i \in \{1, 2, \ldots, L\}$, in accordance with (2.8).

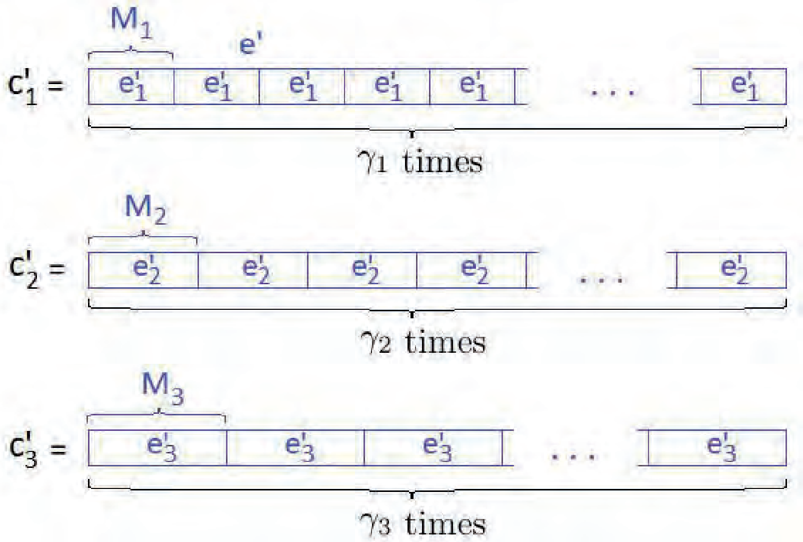


Figure 2.1: Visual equivalent description of (2.31) with $L = 3$ as example

Every vector $\mathbf{e}_i \in \mathbb{Z}^{M_i}$ in Fig. 2.1 has zeros in all entries, except in entry $\mathbf{e}_i(r_i) = 1$ according to (2.29). A generic representation of $\mathbf{e}_i$ in a system with $L = 3$ is shown in Fig. 2.2.

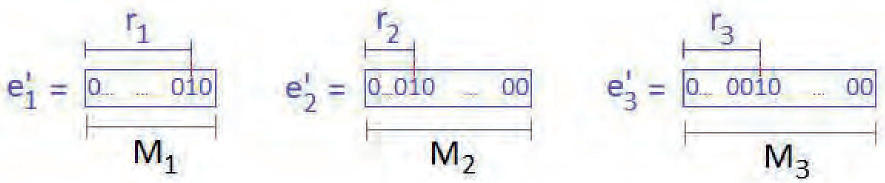


Figure 2.2: Visual equivalent description of $\mathbf{e}_i$ as specified in (2.29)

In the context of undersampling systems, Fig. 2.1 combined with Fig. 2.2 has the representation of Fig. 2.3 for $L = 3$. The sampling rates $F_{s,i}$ are the length of the i-th DFT window and numerically correspond to M_i, for $i \in \{1, 2, 3\}$. The r_i-th entry of each $\mathbf{e}_i$ is 1 due to (2.29) as in Fig. 2.2. Here, r_i indicates the frequency f undersampled at the rate

$M_i = F_{s,i}$, as f mod $M_i = r_i$, in accordance with (2.3). Note that the value of N is achieved only when $\mathbf{c}_1(N) = \mathbf{c}_2(N) = \mathbf{c}_3(N) = 1$, and that over the dynamic range D only one value of N fits in this definition. This simultaneousness derives from (2.25). Since all peaks in Fig. 2.3 are of the form $r_i + kM_i$, for $k \in \{0, 1, 2, \ldots, \gamma_i - 1\}$, they obey the rule of (2.23).

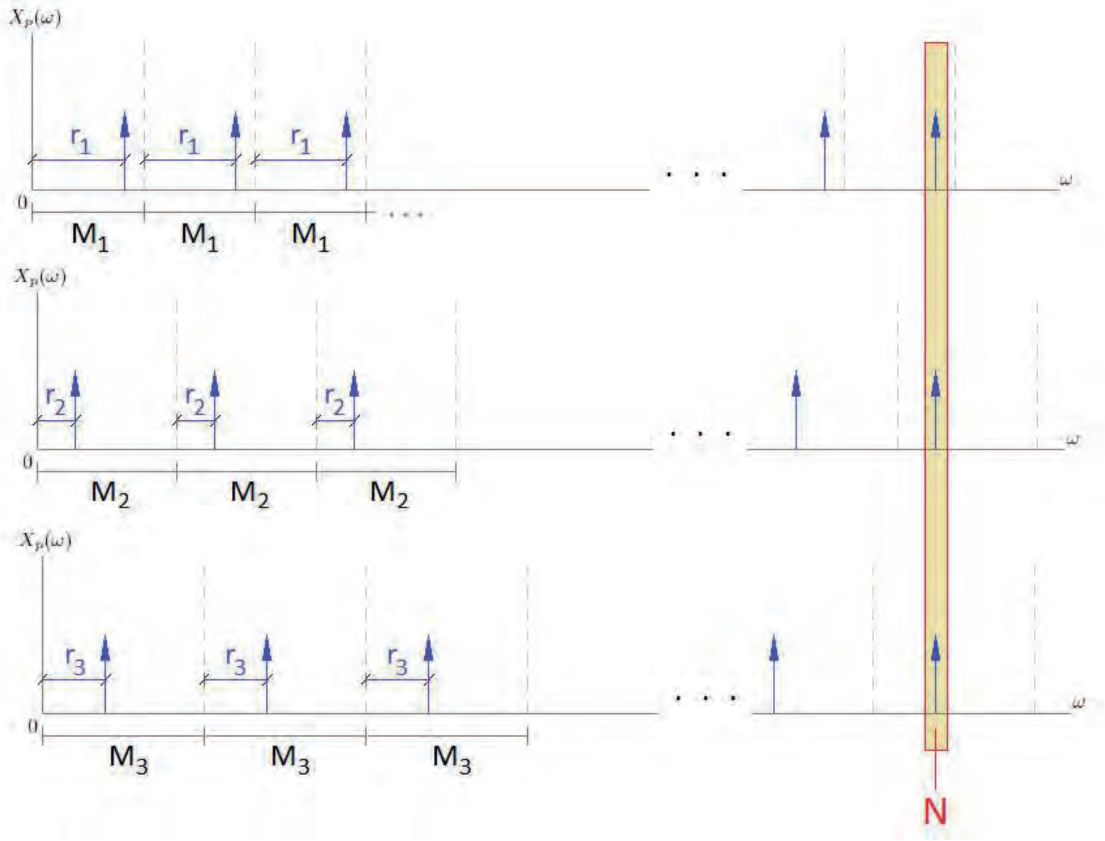


Figure 2.3: Visual equivalent description of CRT for an undersampling system with $L = 3$. Each graphic reproduces the values in $\mathbf{c}_i$ through repetitions of the DFT window, and the value of N is achieved when all the peaks occur simultaneously, due to (2.25).

Each vector $\mathbf{c}_i \in \mathbb{Z}^D$ is obtained by means of (2.31). Gathering all $\mathbf{c}_i$ together, the mapping vector $\mathbf{v} \in \mathbb{Z}^D$ is given by

$$\begin{cases} \mathbf{v}(p) = 1, & \text{if } \prod_{i=1}^{L} \mathbf{c}_i(p) = 1, \\ \mathbf{v}(p) = 0, & \text{otherwise.} \end{cases} \tag{2.34}$$

Note that the each vector $\mathbf{c}_i$ contains a sequence of entries whose values are either 0 or 1. The relevant information is the cardinality of the entries 1 throughout the vector $\mathbf{c}_i$, which indicate the possible values of N according to the content of each set S_i, for $i \in \{1, 2, \ldots, L\}$.

as in (2.24). Hence, (2.34) aims at obtaining the set intersection specified in (2.25), which returns N.

2.2.2 Data Structure for Remainders With Errors

If $N \in \mathbb{Z}$ and $\Delta_i = 0$, for $i \in \{1, 2, \ldots, L\}$, the data structure explained in (2.29)-(2.34) suffices for determining the value of N. However, if $\Delta_i \neq 0$ for any $i \in \{1, 2, \ldots, L\}$, the entries of the unitary vectors $\mathbf{e}_i$ must inform not the integer remainders, but the interval in which the remainders lie given that the errors are continuous variables. Fig. 2.4 illustrates the assemblage of vectors $\mathbf{e_i}$, for $i \in \{1, 2, \ldots, L\}$, in terms of Sections 2.2.1 and 2.2.2. Hereafter, each entry in $\mathbf{e_i}$ is determined in accordance with the model of Fig. 2.4(b), i.e., the entry that satisfies the criterion in terms of r_i assumes the value 1, while the others remain with value zero. In this chapter, the interval is 1/4 for each entry. Hence, while in Section 2.2.1 all vectors $\mathbf{c}_i \in Z^D$, now $\mathbf{c}_i \in Z^{4D}$ due to the fact that the vector $\mathbf{e}_i \in \mathbb{Z}^{4M_i}$. The length 1/4 is chosen as an example, so that other implementations of the here proposed method can have different values for this length. Note that in (a) the r_i stand for a discrete variable, whereas in (b) the r_i are continuous variables due to the fact that N is also a continuous number.

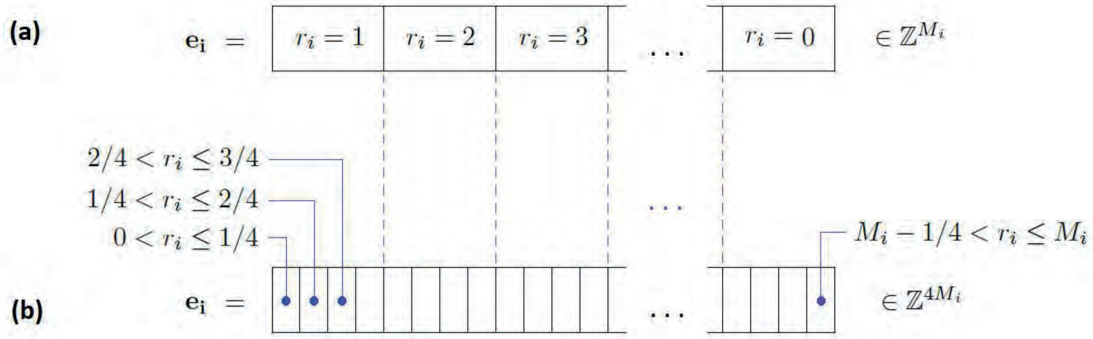


Figure 2.4: Assemblage of vectors $\mathbf{e_i}$, for $i \in \{1, 2, \ldots, L\}$, (a) in terms of Section 2.2.1 and (b) in terms of Section 2.2.2, where the value 1 is inserted in the entry that can contain the true r_i. In (a), the r_i stand for discrete values, whereas in (b) the entries refer to the interval of r_i, which are now continuous variables.

In the proposed KME-CRT with errors in the remainders, all $\tilde{r}_i$ and M_i are normalized to the case $M = 1$, so that $M_i = \Gamma_i$, for $i \in \{1, 2, \ldots, L\}$. This is achieved by dividing all rows in (2.4) by M as in (2.35),

$$\begin{cases} N_m \bmod \Gamma_1 = \tilde{r}_1/M, \\ N_m \bmod \Gamma_2 = \tilde{r}_2/M, \\ \quad \vdots \quad \vdots \quad \vdots \\ N_m \bmod \Gamma_L = \tilde{r}_L/M, \end{cases} \tag{2.35}$$

where $N_m = N/M$. The assemblage of $\mathbf{e}_i \in \mathbb{Z}^{4M_i}$, for $i \in \{1, 2, \ldots, L\}$, is modified to

$$\begin{cases} \mathbf{e}_i(p) = 1, \qquad \text{if } p = \lceil 4r_i \rceil \\ \text{For } k \in \{-2, -1, 1, 2\} \text{ do} \\ \qquad \mathbf{e}_i(p) = 1, \qquad \text{if } p = (\lceil 4r_i \rceil + k) \bmod 4\Gamma_i \text{ and } \lceil 4r_i \rceil + k \neq 4\Gamma_i, \\ \qquad \mathbf{e}_i(p) = 1 \qquad \text{if } p = 4\Gamma_i \text{ and } \lceil 4r_i \rceil + k = 4\Gamma_i, \\ \text{End For} \\ \mathbf{e}_i(p) = 0, \qquad \text{otherwise}, \end{cases} \tag{2.36}$$

where $\lceil . \rceil$ stands for the ceil operator, which turns the input number to next integer towards plus infinity. Eq. (2.36) adapts (2.29) for the case of errors in the remainders.

Algorithm 3 specifies the routine for the estimation of $\hat{N}$ in the proposed method for remainders with errors. The set $\Gamma \in \{\Gamma_1, \Gamma_2, \ldots, \Gamma_L\}$ has its values selected with aid of the same auxiliary set $\mathcal{Y}$ used in (2.32). The input arguments are $M_i, \tilde{r}_i, s_i, H$, where s_i is the incremental step of the ME realizations and H is the hypothesis of error variances, i.e., $H = 1$ for $\sigma_i = \tau/3$, and $H = 2$ for $\sigma_i = (M_i\tau)/(3M_1)$, for $i \in \{1, 2, \ldots, L\}$. The corresponding adjustment is between lines 33 and 37 of the Algorithm, following the stated in (2.21) and (2.22) respectively.

Note that, according to the commands in lines 32, 33 and 34 of the Algorithm 3, the proposed method evaluates the minimal squared errors in terms of the M-Estimators in (2.21) and (2.22). A set of numbers is tested in intervals in accordance with the incremental step s_i, which is the resolution of test. It is indeed a test made over discrete values, however, any real number can be sufficiently approximated with a sufficiently small s_i.

2.2.3 Tensorial Models for the Proposed Mapping Vector

In this Section, we present equivalent approaches to the set of Eqs. (2.29)-(2.34) involving tensorial operations as a way of enriching the proposed technique. In Section 2.2.3.1, we explain the sequence of operations for the error-free case of Section 2.2.1, and in Section

Algorithm 3 Proposed Kronecker Based Mapping Vector for ME-CRT

```
procedure Proposed KME-CRT $(M_i, \tilde{r}_i, s_i, H)$
    $\mathbf{v} \leftarrow$ zeros$[4D \times 1]$        % $\mathbf{v}$ is the mapping vector
                                          to be obtained
    for $i = 1 : L$ do
        $\mathbf{e}_i \leftarrow$ zeros$[4\Gamma_i \times 1]$
        $\mathbf{e}_i(j_i) \leftarrow 1$, where $j_i = \lceil 4r_i \rceil$
        for $k \in \{-2, -1, 1, 2\}$ do
            if $j_i + k \neq 4\Gamma_i$ then
                $\mathbf{e}_i((j_i + k) \bmod 4\Gamma_i) \leftarrow 1$
            else $j_i + k = 4\Gamma_i$
                $\mathbf{e}_i(4\Gamma_i) \leftarrow 1$
        $\mathcal{Y}_i \leftarrow \{1, 2, \ldots, i-1, i+1, \ldots, L\}$
        $\mathbf{w}_i = \mathbf{u}_{\Gamma_{\mathcal{Y}_i(1)}} \otimes \mathbf{u}_{\Gamma_{\mathcal{Y}_i(2)}} \otimes \cdots \otimes \mathbf{u}_{\Gamma_{\mathcal{Y}_i(L-1)}}$        % Eq. (2.32)
        $\mathbf{c}_i \leftarrow \mathbf{w}_i \otimes \mathbf{e}_i$        % Eq. (2.31)
    for $p = 1 : 4D$ do
        if $\prod_{i=1}^{L} \mathbf{c}_i(p) = 1$ then        % Conditions stated in (2.34)
            $\mathbf{v}(p) \leftarrow 1$
    $\mathbf{p} \leftarrow$ find$(\mathbf{v}(p))$        % $\mathbf{p}$ informs which rows of $\mathbf{v}$ have all elements 1
    $L_p \leftarrow$ length$(\mathbf{p})$
    $\mathbf{p} \leftarrow \mathbf{p}/4 - 1/4$        % As each row of $\mathbf{v}$ in fact spans $1/4$.
    $L_q \leftarrow \lceil 1/4s_i \rceil$
    $\mathbf{Q} \leftarrow$ zeros$[L_q \times L_p]$
    $\mathbf{W} \leftarrow$ zeros$[L_q \times L_p]$
    for $j_1 = 1 : L_q$ do
        for $j_2 = 1 : L_p$ do
            $\mathbf{Q}(j_1, j_2) \leftarrow \mathbf{p}(j_2) + (j_1 - 1)s_i$
            for $i = 1 : L$ do
                if $H = 1$ then        % Hypothesis of constant variances
                    $\mathbf{W}(j_1, j_2) \leftarrow \mathbf{W}(j_1, j_2) + (d_{\Gamma_i}(\mathbf{Q}(j_1, j_2) \bmod \Gamma_i), \tilde{r}/M)^2$
                else $H = 2$        % Hypothesis of $\sigma_i^2$ as a function of $M_i$
                    $\mathbf{W}(j_1, j_2) \leftarrow \mathbf{W}(j_1, j_2) + (d_{\Gamma_i}((\mathbf{Q}(j_1, j_2) \bmod \Gamma_i), \tilde{r}/M)/\Gamma_i)^2$
    $[x_1, x_2] \leftarrow \min(\mathbf{W})$
    $[\sim, x_3] \leftarrow \min(x_1)$        % $\mathbf{W}(x_2(x_3))$ is the minimal deviation in the ME algorithm
    $\hat{N} = M\mathbf{Q}(\mathrm{x}_2(\mathrm{x}_3), \mathrm{x}_3)$        % Hence, $\mathbf{Q}(x_2(x_3))$ is the optimal value of $N/M$.
```

2.2.3.2, for the case of a system with erroneous remainders of Section 2.2.2.

The motivation for testing solutions with a tensorial approach involves three aspects. First, it is worth noting that $\mathbf{e}_i$ described in (2.36) contains $4M_i$ entries in a single dimension. It is possible that overflow problems arise depending on the software architecture and the value M_i itself. A second reason is that gains in processing speed should be tested, as tensorial operations have a relatively low computational cost in the main available programs such as Matlab. Furthermore, CRT systems are very prone to be represented in a tensorial fashion. Note that any CRT system with L moduli M_i can be represented by a tensor with L dimensions $M_1, M_2, \dots, M_L$, where each entry of the tensor is uniquely described by the remainders $r_1, r_2, \dots, r_L$ that stand for the coordinates of the entry. Hence, our motivation to verify CRT solutions by means of a tensorial approach relies on their significant similarity.

According to [69], the n-mode product between a tensor $\boldsymbol{\mathcal{A}} \in \mathbb{R}^{x_1 \times x_2 \times \cdots \times x_N}$ with a matrix $\mathbf{E} \in \mathbb{R}^{J \times x_n}$ over the n-th dimension of $\boldsymbol{\mathcal{A}}$ is denoted by

$$\boldsymbol{\mathcal{B}} = \boldsymbol{\mathcal{A}} \times_n \mathbf{E}, \tag{2.37}$$

where $\boldsymbol{\mathcal{B}} \in \mathbb{R}^{x_1 \times x_2 \times \cdots \times x_{n-1} \times J \times x_{n+1} \times \cdots \times x_N}$. Note that, in terms of matrix based expressions, we have

$$\boldsymbol{\mathcal{B}}_{(n)} = \mathbf{E} \boldsymbol{\mathcal{A}}_{(n)}, \tag{2.38}$$

where the subscript in tensor $\boldsymbol{\mathcal{B}}_{(n)}$ denotes that it is unfolded over its n-th dimensional fibers [69]. Hence, $\boldsymbol{\mathcal{A}}_{(n)}$ and $\boldsymbol{\mathcal{B}}_{(n)}$ are matrices, so that the first size of $\boldsymbol{\mathcal{A}}_{(n)}$ is x_n, while in $\boldsymbol{\mathcal{B}}_{(n)}$ the first size is J. The second size of $\boldsymbol{\mathcal{A}}_{(n)}$ and $\boldsymbol{\mathcal{B}}_{(n)}$ is the product of all remaining dimensions of the original tensor, i.e., $(x_1 x_2 \dots x_{n-1} x_{n+1} \dots x_N)$.

One of the properties of the n-mode product shown in (2.37) is that one can select the entries of a tensor over one of its dimensions by means of a diagonal matrix whose elements are properly chosen. If the matrix used in (2.37) is $\mathbf{E} \in \mathbb{R}^{x_n \times x_n}$ where $\mathbf{E}(l, l) = 1$, for $1 \leq l \leq x_n$, and all other entries are zero, then the resulting tensor $\boldsymbol{\mathcal{B}}_{(n)}$ is

$$\begin{cases} \boldsymbol{\mathcal{B}}(x_1, x_2, \dots, x_n, \dots, x_N) = \boldsymbol{\mathcal{A}}(x_1, x_2, \dots, x_n, \dots, x_N), \text{ if } x_n = l, \\ \boldsymbol{\mathcal{B}}(x_1, x_2, \dots, x_n, \dots, x_N) = 0, \text{ otherwise.} \end{cases} \tag{2.39}$$

As an example, let the tensor $\boldsymbol{\mathcal{A}} \in \mathbb{R}^{5 \times 6 \times 4}$ and the matrix $\mathbf{E} \in \mathbb{R}^{5 \times 5}$ have an n-mode product. If $\mathbf{E}(i, j) = 0$ for $i \neq j$ and the main diagonal of $\mathbf{E}$ is the vector

$$\mathbf{e} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \end{bmatrix}, \tag{2.40}$$

then

$$\mathbf{E} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}. \tag{2.41}$$

and the n-mode product of (2.37) is written as $\boldsymbol{\mathcal{B}} = \boldsymbol{\mathcal{A}} \times_1 \mathbf{E}$. Fig. 2.5 provides a visual interpretation of this tensorial operation, with the original entries in gray and zero values in white. The diagonal of $\mathbf{E}$ is applied over the first dimensions, i.e., column-wise. Tensor $\boldsymbol{\mathcal{A}}$ has its entries in the fourth slice of the first dimension preserved, while all other entries are set to zero, yielding thereby tensor $\boldsymbol{\mathcal{B}}$.

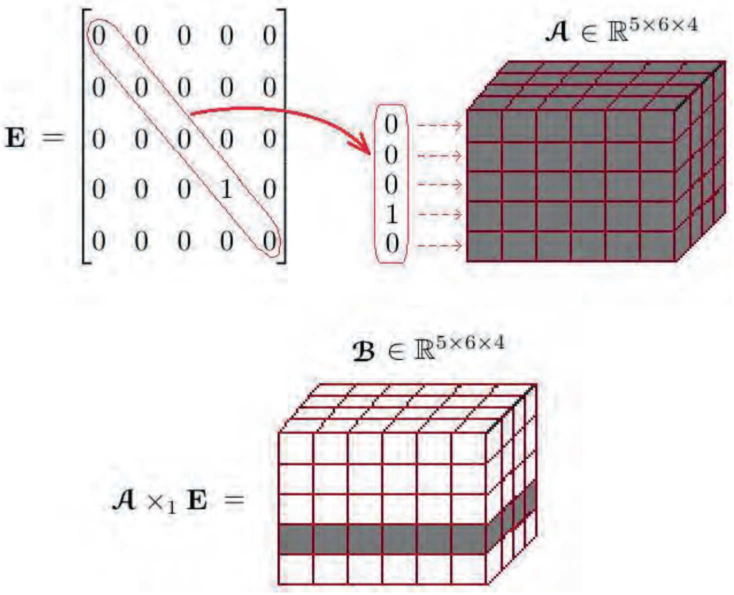

$\mathbf{E} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$
$\begin{matrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{matrix}$
$\boldsymbol{\mathcal{A}} \in \mathbb{R}^{5\times6\times4}$
$\boldsymbol{\mathcal{B}} \in \mathbb{R}^{5\times6\times4}$
$\boldsymbol{\mathcal{A}} \times_1 \mathbf{E} =$

Figure 2.5: Visual interpretation of the tensorial n-mode product $\boldsymbol{\mathcal{B}} = \boldsymbol{\mathcal{A}} \times_1 \mathbf{E}$. By convention, original entries are in gray and zero values in white.

2.2.3.1 Tensorial Model for the Error-Free Case

The tensorial algorithm for the proposed KME-CRT starts with the moduli M_i and the remainders r_i. Initially, we set up a tensor $\boldsymbol{\mathcal{T}} \in \mathbb{Z}^{M_1 \times M_2 \times \cdots \times M_L}$, with $\boldsymbol{\mathcal{T}}(j_1, j_2, \ldots, j_L) = 1$,for all $i \in \{1, 2, \ldots, L\}$ and $j_i \in \{1, 2, \ldots, M_i\}$, i.e., the value 1 in all entries. Now define matrices $\mathbf{E}_i \in \mathbb{Z}^{M_i \times M_i}$ following

$$\mathbf{E}_i = \operatorname{diag}\{\mathbf{e}_i\}, \tag{2.42}$$

using the vectors $\mathbf{e}_i$ defined in (2.29). Algorithm 4 shows the sequence of steps in order to accomplish the calculation of N. Note that, in line 8, the circshift$(\mathbf{a}, j)$ command makes a circular shift over a vector $\mathbf{a} \in \mathbb{R}^L$ as in

$$\operatorname{circshift}(\mathbf{a}, j) = \Big[\mathbf{a}(L - j + 1 : L) \quad \mathbf{a}(1 : L - j)\Big], \tag{2.43}$$

while, in line 9, $\boldsymbol{\mathcal{T}}$ is unfolded in its $(L+1)$-th dimension, yielding vector $\mathbf{v}$.

Algorithm 4 Proposed Tensorial KME-CRT for remainders free of errors

```
 1: procedure TENSORIAL KME-CRT 1 $(M_i, r_i)$
 2:     $\mathfrak{T} \leftarrow \text{ones}([M_i])$
 3:     for $i = 1 : L$ do
 4:         $\mathbf{E}_i \leftarrow \text{diag}\{\mathbf{e}_i\}$                    % Vectors $\mathbf{e}_i$ defined in (2.29)
 5:     $\mathfrak{T}^{(1)} \leftarrow \mathfrak{T} \times_1 \mathbf{E}_1$
 6:     $M'_i \leftarrow M_i$
 7:     for $i = 2 : L$ do
 8:         $M'_i \leftarrow \text{circshift}\{M'_i, (L-1)\}$
 9:         $\mathbf{v} \leftarrow \mathfrak{T}^{(i-1)}_{(L+1)}$
10:         Reshape the vector $\mathbf{v}$ into a tensor $\mathfrak{T}^{(i)} \in \mathbb{N}^{[M'_i]}$
11:         $\mathfrak{T}^{(i)} \leftarrow \mathfrak{T}^{(i)} \times_1 \mathbf{E}_i$
12:     $\mathbf{v} \leftarrow \mathfrak{T}^{(L)}_{(L+1)}$
13:     $N \leftarrow \text{find}\{\mathbf{v}\}$
```

As visual example of application for Algorithm 4, let a CRT system with $L = 3$ and M_1, M_2, M_3 be generic moduli. The tensor $\mathfrak{T} \in \mathbb{Z}^{M_1 \times M_2 \times M_3}$ is set up with all entries 1. Fig. 2.6 shows the sequence of steps in this specific case. Note that, along the three steps (a), (b) and (c), the tensor $\mathfrak{T}^{(i)}$, $i \in \{1, 2, 3\}$, is the reshape of vector $\mathbf{v}$ as stated in line 10 of Algorithm 4. The n-mode products along the first dimension successively filter the elements of the reshaped tensor $\mathfrak{T}^{(i)}(x_1, x_2, x_3)$ which are located at the slice $x_1 = r_i$ at each step i.

2.2.3.2 Tensorial Model for the Case of Remainders with Errors

In order to adapt the Algorithm 3 for the case when there are errors in the remainders, we rewrite matrices $\mathbf{E}_i \in \mathbb{Z}^{4M_i \times 4M_i}$ as in (2.42), however now using vectors $\mathbf{e}_i \in \mathbb{Z}^{4M_i}$ defined as in (2.36). Algorithm 5 gives the pathway for the best estimation of $\hat{N}$. Note that, by the end of Algorithm 5, the last steps of Algorithm 3 are needed to achieve the definitive estimation.

The routine in Algorithm 5 is shown in Fig. 2.7, where the scheme in Fig. 2.6 is adapted in order to provide the estimation of $\hat{N}$ when at least one of the errors is not equal to zero. Note that the final vector is $\mathbf{p}$ which, differently of vector $\mathbf{v}$ in Fig. 2.6, has several intervals of possible locations of $\hat{N}$, and that the first dimension of the reshaping tensor $\mathfrak{T}^{(i)}$ at each stage i is $4M_i$.

2.2.4 Possibility of Correct Estimation when $M/4 \leq \tau < M/2$ by Means of the Proposed Mapping Vector

In accordance with [64] [70], CRT is a robust method when the remainders have an error bound τ and the reconstruction error is also bounded to $|N - \hat{N}| \leq \tau$. Note that, as

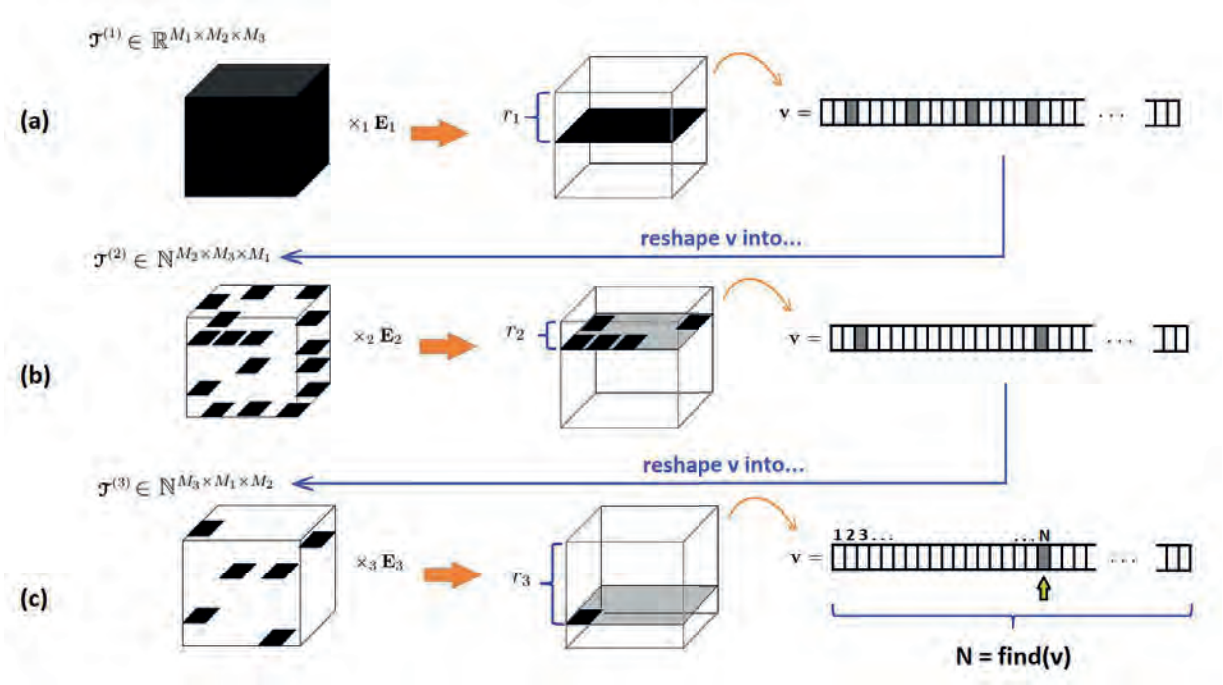

$\mathcal{T}^{(1)} \in \mathbb{R}^{M_1 \times M_2 \times M_3}$
(a)
$\times_1 \mathbf{E}_1$
r_1
v =
...
reshape v into...
$\mathcal{T}^{(2)} \in \mathbb{N}^{M_2 \times M_3 \times M_1}$
(b)
$\times_2 \mathbf{E}_2$
r_2
v =
...
reshape v into...
$\mathcal{T}^{(3)} \in \mathbb{N}^{M_3 \times M_1 \times M_2}$
(c)
$\times_3 \mathbf{E}_3$
r_3
1 2 3 ...
... N
v =
...
N = find(v)

Figure 2.6: Example of application of Algorithm 4 with $L = 3$. The tensor $\mathcal{T}^{(i)}$, $i \in \{1, 2, 3\}$, is at each step i the reshape of vector $\mathbf{v}$ as stated in line 10. The n-mode products along the first dimension filter the elements of the reshaped tensor $\mathcal{T}^{(i)}(x_1, x_2, x_3)$ where $x_1 = r_i$.

extensively studied in the literature, the maximum value for τ is $\tau < M/4$ as given in (2.6). In [65], the sharpness of the boundary in (2.6) is illustrated by means of an example. The boundary in (2.6) is indeed the sufficient one in order to ensure that all folding integers n_i are correctly reconstructed. However, the actual boundary from a theoretical standpoint is given by

$$-\frac{M}{2} \leq \Delta_1 - \Delta_i < \frac{M}{2}, \tag{2.44}$$

so that (2.44) is automatically guaranteed when $\tau < M/4$, as extensively proven in [41] [50] [64] [70]. Thus, while (2.6) is a sufficient condition for the uniqueness of the solution of the folding numbers n_i, (2.44) is the necessary and sufficient one, constituting the general boundary for a robust CRT.

The main drawback in working with (2.44), in accordance with [41] [50], is that it involves two remainder errors, which are in practice hard to check. That is the reason why the limit of (2.6) is largely adopted in the literature instead of (2.44). On the other hand, a consequence of using (2.6) is that selecting the remainder of least variance is crucial for the success of the CRT methods based on a reference remainder [41] [64]. In order to illustrate

Algorithm 5 Proposed Tensorial KME-CRT for remainders with errors

```
procedure PROPOSED TENSORIAL KME-CRT 2 $(M_i, \tilde{r}_i)$
    $M_i(1) = 4M_i(1)$
    $\mathfrak{T} \leftarrow \text{ones}([M_i])$
    for $i = 1 : L$ do
        $\mathbf{E}_i \leftarrow \text{diag}\{\mathbf{e}_i\}$                % Vectors $\mathbf{e}_i$ defined in (2.36)
    $\mathfrak{T} \leftarrow \mathfrak{T} \times_1 \mathbf{E}_1$
    $M'_i \leftarrow M_i$
    for $i = 2 : L$ do
        $M'_i(1) \leftarrow M'_i(1)/4$
        $M'_i \leftarrow \text{circshift}\{M'_i, (L-1)\}$
        $M'_i(1) \leftarrow 4M'_i(1)$
        $\mathbf{v} \leftarrow \mathfrak{T}_{(L+1)}$
        Reshape the vector $\mathbf{v}$ into a tensor $\mathfrak{T}^{(i)} \in \mathbb{N}^{M'_i}$
        $\mathfrak{T}^{(i)} \leftarrow \mathfrak{T}^{(i)} \times_1 \mathbf{E}_i$
    $\mathbf{p} \leftarrow \text{find}\{\mathfrak{T}^{(L)}_{(L+1)}\}$
    Resume Algorithm 3 at line 24
```

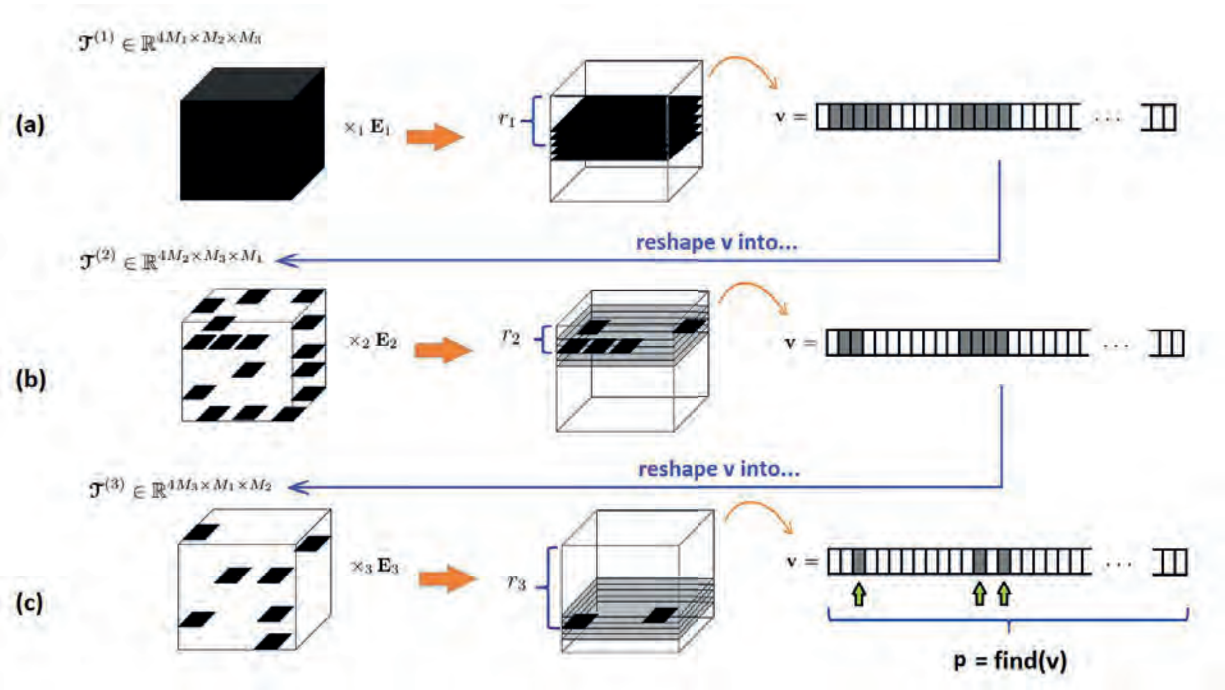

$\mathfrak{T}^{(1)} \in \mathbb{R}^{4M_1 \times M_2 \times M_3}$
(a)
$\times_1 \mathbf{E}_1$
r_1
v =
reshape v into...
$\mathfrak{T}^{(2)} \in \mathbb{R}^{4M_2 \times M_3 \times M_1}$
(b)
$\times_2 \mathbf{E}_2$
r_2
v =
reshape v into...
$\mathfrak{T}^{(3)} \in \mathbb{R}^{4M_3 \times M_1 \times M_2}$
(c)
$\times_3 \mathbf{E}_3$
r_3
v =
p = find(v)

Figure 2.7: Example of application of Algorithm 5 with $L = 3$, providing a version of Fig. 2.6 for the case of remainders with errors as described in Algorithm 5. Note that the final vector is $\mathbf{p}$ which, differently of vector $\mathbf{v}$ in Fig. 2.6, has several intervals of possible locations of $\hat{N}$.

this aspect, Fig. 2.8 shows an example with the remainders errors of a CRT system of $L = 4$,

$\Delta_i \in \mathbb{R}$, for $i \in \{1, 2, 3, 4\}$, normalized at each length M. In this example, we assume $M/4 < |\Delta_3| < M/2$ and $M/4 < |\Delta_4| < M/2$. In Fig. 2.8(a), the chosen reference remainder is r_2, so that $|\Delta_2 - \Delta_i| < M/2$, with $i \in \{1, 3, 4\}$. In Fig. 2.8(b), an alternative case is shown where r_1 is erroneously selected as the reference remainder. Suppose, as suggested by the figure, that $|\Delta_1 - \Delta_4| > M/2$ through the continuous red line, i.e., inside the length M. In order to obtain $|\Delta_1 - \Delta_4| < M/2$, the resulting minimal distance $|\Delta_1 - \Delta_4|$ must be calculated through the green dashed line to the left. However, computing the distance $|\Delta_1 - \Delta_4|$ in this way corresponds to changing the remainder error Δ_4, yielding $\Delta'_4 = \Delta_4 - M$, whereas Δ_4 is the true deviation. Since originally $\tilde{r}_4 = r_4 + \Delta_4$, the change in the pathway entails $r_4 + \Delta'_4 + M = r'_4 + \Delta'_4$, where $r'_4 = r_4 + M$, thus erroneously changing the value of the remainder r_4. As a consequence, the choice of r_1 as the reference remainder in this example leads to the violation of the rule in (2.44).

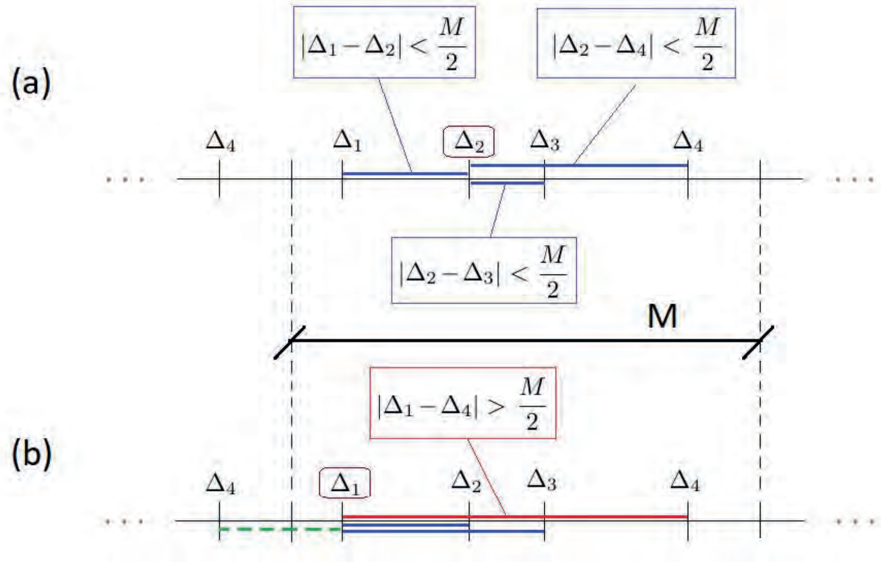

(a)
$|\Delta_1 - \Delta_2| < \frac{M}{2}$
$|\Delta_2 - \Delta_4| < \frac{M}{2}$
Δ_4
Δ_1
Δ_2
Δ_3
Δ_4
$|\Delta_2 - \Delta_3| < \frac{M}{2}$
M
$|\Delta_1 - \Delta_4| > \frac{M}{2}$
(b)
Δ_4
Δ_1
Δ_2
Δ_3
Δ_4

Figure 2.8: Magnitudes of the errors Δ_i along the distance M, where (a) the chosen reference remainder is r_2, when all distances respect the limit $|\Delta_2 - \Delta_i| < M/2$, whereas in (b) r_1 is erroneously selected as the reference remainder, yielding $|\Delta_1 - \Delta_4| > M/2$, which violates the criterion in (2.44), that is a necessary and sufficient condition for solving the CRT system.

We now explain how the proposed mapping vector addresses this problem. Lines 5 to 11 of Algorithm 3 are dedicated to set up vectors $\mathbf{e}_i \in \mathbb{Z}^{4M_i}$, for $i \in \{1, 2, \ldots, L\}$, while lines 13 and 14 provide their successive concatenations until the vectors $\mathbf{c}_i$ are assembled. As explained with Fig. 2.4, each vector $\mathbf{e}_i$ contains entries (or slots) that inform the occurrence of the $\tilde{r}_i$ value in terms of slots whose length is 1/4. In accordance with (2.36), two slots on each side of the original entry in $\mathbf{e}_i$ are also filled with 1. As a consequence, any possible slot corresponding to the correct r_i has 1 as its entry in $\mathbf{e}_i$, since two slots cover 1/2 on each side of the entry related to $\tilde{r}_i$. Fig. 2.9 shows how the criterion in (2.44) is maintained. Independently of how near to the border of the original slot the value $\tilde{r}_i$ is, the values within the distance $\pm M/2$ have also entries 1. Points a and b in Fig. 2.9 illustrate two hypotheses for the location of the value r_i near border points. Points a' and b', which show the spatial boundary for the value of $\tilde{r}_i$ in each case, with a maximum deviation of $M/2$ or two slots, are also within a slot of value 1, and their interval are then included in the M-Estimator test of either (2.21) or (2.22) due to (2.36).

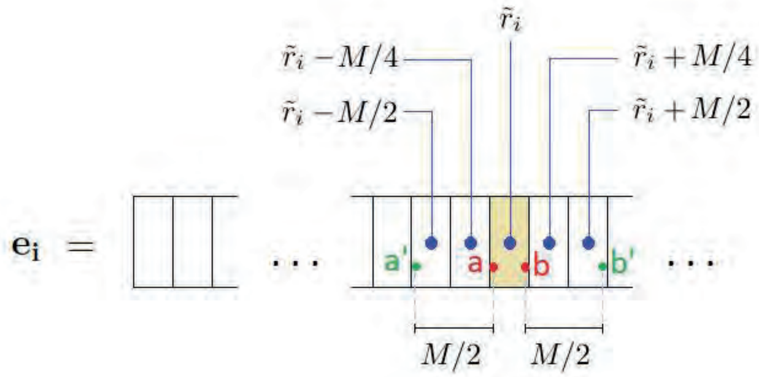


Figure 2.9: Assemblage of vector $\mathbf{e}_i$ in terms of $\tilde{r}_i$ value, with two entries on each side of the original entry also filled with 1. Points a and b are examples of near border points, and the points a' and b' are in any case within a slot of value 1.

After assembling all vectors $\mathbf{e}_i \in \mathbb{Z}^{4M_i}$ and $\mathbf{c}_i \in \mathbb{Z}^{4D}$, commands in lines 16 and 17 of Algorithm 3 yield the scheme in Fig. 2.10, where each vector $\mathbf{c}_i$ is set up according lines 13 and 14 of the same algorithm. As a consequence, only a restricted set of entries in the resulting vector $\mathbf{v} \in \mathbb{Z}^{4D}$ is obtained. Since $\mathbf{v}$ contains solely the intersection of all 1 entries from the vectors $\mathbf{c}_i$, for $i \in \{1, 2, \ldots, L\}$, the cardinality of these entries indicate in which points of the dynamic range the application of the M-Estimator is to be made, thus constituting the mapping vector $\mathbf{v}$ over the dynamic range D. Note that there are many intersections of this type throughout the vector $\mathbf{v}$, such that in Fig. 2.10 only one intersection, p_1, is shown due to practical reasons.

As a result of schemes in Figs. 2.9 and 2.10, every value r_i that follows $|r_i - \tilde{r}_i| \leq M/2$ is considered for the analysis of the M-Estimator. Therefore, the criterion in (2.44) is fully preserved, and the reconstruction of N is possible without necessarily choosing the best reference remainder. Note that there is no guarantee the value of $\hat{N}$ that minimizes (2.21) or (2.22) reproduces the real value of N. However, the value of N is necessarily in the vicinity of at least one entry of the vector $\mathbf{p}$, which informs the cardinalities of the mapping vector $\mathbf{v}$ in accordance with the line 18 in Algorithm 3.

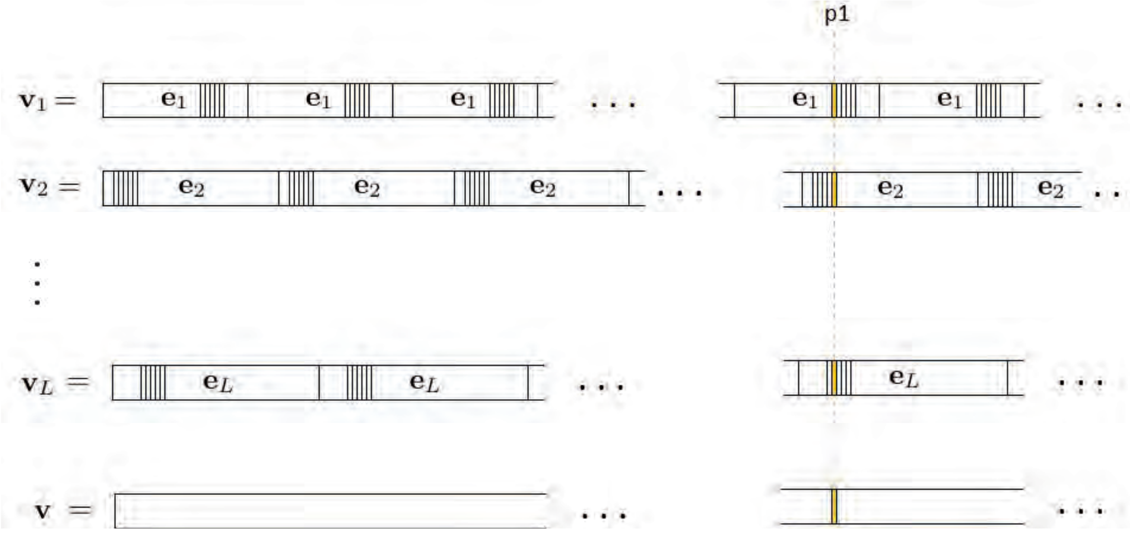


Figure 2.10: Assemblage of vectors $\mathbf{c}_i$. Only a restricted set of entries in $\mathbf{v}$ are to be tested with the M-Estimator. In this example, the intersection of all entries of the same cardinality with value 1 is indicated in yellow. Note that there are many intersections of this type throughout the vector $\mathbf{v}$, while in this Figure only the first one, p_1, is shown.

2.3 Experiments and Results

In this section, we first develop an example with an undersampling system in Section 2.3.1. In Section 2.3.2, results of general simulations are presented for the case of $\sigma_1 = \sigma_2 = \cdots = \sigma_L$ and, in sequel, for the case of $\sigma_i = \mu\Gamma_i$, for $i \in \{1, 2, \ldots, L\}$. In Section 2.3.3, we present the a comparison of the computational cost of CFR-CRT, MLE-CRT and the proposed KME-CRT.

2.3.1 System Validation - Example

In a system with $L = 3$ sensors, the sampling frequencies are $M_i \in \{55 \text{ kHz}, 65 \text{ kHz}, 85 \text{ kHz}\}$. The impinging frequency value N is any real number within the dynamic range, $0 < N < 12.1550$ MHz, and must be estimated. Each sensor reads the peaks in the DFT of the frequency N, whose values are $r_i \in \{17.81 \text{ kHz}, 15.45 \text{ kHz}, 13.24 \text{ kHz}\}$, yielding the system

$$\begin{cases} N \bmod 55 = 17.81, \\ N \bmod 65 = 15.45, \\ N \bmod 85 = 13.24. \end{cases} \tag{2.45}$$

Prior to further steps, proceed to the division of the original system as in (2.45) by M, the

GCD of all moduli M_i. In (2.45), $M = 5$, yielding

$$\begin{cases} N_m \bmod 11 = 3.562, \\ N_m \bmod 13 = 3.09, \\ N_m \bmod 17 = 2.648, \end{cases} \tag{2.46}$$

where $N_m = N/M$. Note that in (2.46) $M = 1$, and thus $M_i = \Gamma_i$, for $i \in \{1, 2, 3\}$.

We assemble vectors $e_1 \in \mathbb{Z}^{44}$, $e_2 \in \mathbb{Z}^{52}$ and $e_3 \in \mathbb{Z}^{68}$ with zeros in all entries. In sequel, we make $e_1(15) = 1$, $e_2(13) = 1$ and $e_3(11) = 1$, filling with value 1 also the following entries: $e_1(13) = e_1(14) = e_1(16) = e_1(17) = 1$, $e_2(11) = e_2(12) = e_2(14) = e_2(15) = 1$ and $e_3(9) = e_3(10) = e_3(12) = e_3(13) = 1$. Such steps correspond to lines 4-13 of Algorithm 3. By this moment, we have the definitive e_i, for $i \in \{1, 2, 3\}$.

We set up the mapping vector $\mathbf{v} \in \mathbb{Z}^{9724}$ with zero in all entries. For $i \in \{1, 2, 3\}$, we organize the vectors $\mathbf{c}_1$ in (2.47), $\mathbf{c}_2$ in (2.48) and $\mathbf{c}_3$ in (2.49),

$$\mathbf{c}_1 = \mathbf{u}_{13} \otimes \mathbf{u}_{17} \otimes \mathbf{e}_1, \tag{2.47}$$

$$\mathbf{c}_2 = \mathbf{u}_{11} \otimes \mathbf{u}_{17} \otimes \mathbf{e}_2, \tag{2.48}$$

$$\mathbf{c}_3 = \mathbf{u}_{11} \otimes \mathbf{u}_{13} \otimes \mathbf{e}_3, \tag{2.49}$$

where the $\mathbf{u}_k$ are defined as in (2.30). It is worth noting that the vectors $\mathbf{c}_i$ in (2.47)-(2.49) correspond to the form of (2.31), where the Kronecker product of all vectors $\mathbf{u}_k$ in (2.32) impose a fixed size for the resulting vector $\mathbf{c}_i$, for $i \in \{1, 2, \ldots, L\}$. Furthermore, since the dimension of the mapping vector $\mathbf{v}$ is the same of all vectors $\mathbf{c}_i$, with $\mathbf{v}, \mathbf{c}_i \in \mathbb{Z}^{4D}$, there is a relationship between the size of $\mathbf{v}$ and the CRT system rows as D derives straightforwardly from the co-prime moduli Γ_i and the GCD M given by (2.8).

We now apply (2.34) in order to obtain mapping vector $\mathbf{v}$ in accordance with the lines 18-22 of Algorithm 3. With line 23, we extract the cardinality of such entries in $\mathbf{v}$, which returns the following vector $\mathbf{p} \in \mathbb{Z}^8$ in (2.50),

$$\mathbf{p}' = \begin{bmatrix} 13 & 895 & 896 & 897 & 6877 & 6878 & 6879 & 7761 \end{bmatrix}. \tag{2.50}$$

Line 25 of Algorithm 3 converts $\mathbf{p}$ of (2.50) into

$$\mathbf{p}' = \begin{bmatrix} 3 & 223.5 & 223.75 & 224 & 1719 & 1719.25 & 1719.5 & 1940 \end{bmatrix}. \tag{2.51}$$

We choose $s_i = 0.02$, thus avoiding that any final estimated $\hat{N}$ lie further than 0.01 from the optimal point. The determination of the optimized value of s_i is beyond the scope of this work, but it suffices to note that infinite other values of s_i are possible. Clearly, the smaller the s_i, the higher the accuracy, but also the computational cost, since smaller s_i entails more values to test in each selected slot. On the other hand, shortening the interval 1/4 yields more resultant cells of test, yet of smaller length. All in all, we have in the choice of the length (in our case, 1/4) and incremental step s_i two tuning parameters that aid us to control the complexity of our system. We then set the matrix $\mathbf{Q} \in \mathbb{R}^{13\times 8}$, whose values are shown in (2.52). The first row of $\mathbf{Q}$ is $\mathbf{p}'$ in (2.51).

$$\mathbf{Q} = \begin{bmatrix} 3.00 & 223.50 & 223.75 & 224.00 & 1719.00 & 1719.25 & 1719.50 & 1940.00 \\ 3.02 & 223.52 & 223.77 & 224.02 & 1719.02 & 1719.27 & 1719.52 & 1940.02 \\ 3.04 & 223.54 & 223.79 & 224.04 & 1719.04 & 1719.29 & 1719.54 & 1940.04 \\ 3.06 & 223.56 & 223.81 & 224.06 & 1719.06 & 1719.31 & 1719.56 & 1940.06 \\ 3.08 & 223.58 & 223.83 & 224.08 & 1719.08 & 1719.33 & 1719.58 & 1940.08 \\ 3.10 & 223.60 & 223.85 & 224.10 & 1719.10 & 1719.35 & 1719.60 & 1940.10 \\ 3.12 & 223.62 & 223.87 & 224.12 & 1719.12 & 1719.37 & 1719.62 & 1940.12 \\ 3.14 & 223.64 & 223.89 & 224.14 & 1719.14 & 1719.39 & 1719.64 & 1940.14 \\ 3.16 & 223.66 & 223.91 & 224.16 & 1719.16 & 1719.41 & 1719.66 & 1940.16 \\ 3.18 & 223.68 & 223.93 & 224.18 & 1719.18 & 1719.43 & 1719.68 & 1940.18 \\ 3.20 & 223.70 & 223.95 & 224.20 & 1719.20 & 1719.45 & 1719.70 & 1940.20 \\ 3.22 & 223.72 & 223.97 & 224.22 & 1719.22 & 1719.47 & 1719.72 & 1940.22 \\ 3.24 & 223.74 & 223.99 & 224.24 & 1719.24 & 1719.49 & 1719.74 & 1940.24 \end{bmatrix} \quad (2.52)$$

We then calculate $\mathbf{W}$ according to the lines 29-40 of the Algorithm 3, entrywise in terms of $\mathbf{Q}$. Hence, given every entry $\mathbf{Q}(i,j)$, the entry $\mathbf{W}(i,j)$ informs the respective error, either according to (2.21) or (2.22).

As a last step, search the minimum absolute value in $\mathbf{W}$. Since this is $\mathbf{W}(2,3) = 0.1605$ and $M = 5$, $\hat{N} = 5\mathbf{Q}(2,3) = 5 \times 223.77 = 1118.85$, i.e., the estimated frequency value is $\hat{N} = 1118.85$ kHz. This value lies in the tolerance interval around the true value $N = 1120$ kHz, as $|N - \hat{N}| < M/2$. It is worth it to remark that the result of the same system given by CFR-CRT is $\hat{N} = 8597.2$ kHz and by MLE-CRT is $\hat{N} = 8597.1$ kHz, values significantly distant from the true N.

2.3.2 Results of Experiments in Terms of Errors Variance

Next, we compare the results of the proposed KME-CRT with the state-of-the-art CFR-CRT and MLE-CRT for the case of same and different variances of errors. This comparison is based on the fact that CFR-CRT is the state-of-the-art method for constant variances of the errors Δ_i, whereas MLE-CRT is suitable for different variances of Δ_i, $i \in \{1, 2, \ldots, L\}$.

In both scenarios, we also compare the proposed KME-CRT with the complete ME, which is the application of the ME routine over the entire dynamic range D without the help of the Kronecker mapping vector. Hence, with the complete ME, the total number of realizations is D/s_i. This comparison is performed in order to make clear how effective is the proposed KME-CRT due to the Kronecker product of the specified vectors.

The evaluation of the computational time required to perform the proposed KME-CRT, the CFR-CRT and the MLE-CRT is carried out in Matlab by means of commands tic and toc, which are used to measure the time the computer takes to perform a sequence of commands. The processor used to undertake the measurements is a dual core i7-5500U at 2.4 GHz, with double precision in Matlab settings. We do not employ parallel structures as the goal is to illustrate the total required computational effort. nevertheless, we highlight that CFR-CRT and MLE-CRT can be deployed in parallel structures. The proposed KME-CRT can be executed in a parallel fashion with regards to the vectors assemblage of (2.47)-(2.49) and in testing each value of matrix $\mathbf{Q}$. However, since several possible arrangements are possible which depend on the number of processors available, for the sake of simplicity we execute all routines sequentially.

In our simulations, the moduli are $\Gamma_i \in \{11, 13, 17\}$. Recall that, if $M \neq 1$, the previous normalization with regards to M is assumed. Errors Δ_i are generated under the Gaussian distribution in the range $[N - \tau, N + \tau]$ for each value of τ. At each realization, a real valued N with two decimals is randomly chosen from 0 up to $D = 2431$ over 10^5 realizations. The estimation is accepted as correct when $N - M/2 \leq \hat{N} \leq N + M/2$.

In Fig. 2.11, the percentages of correct estimation of N are shown for constant variances for the proposed KME-CRT, CFR-CRT [41], MLE-CRT [42] and the complete ME. Throughout the values of τ, the percentages of correct calculations of N are higher with the proposed KME-CRT than with CFR-CRT or MLE-CRT. The results are shown for $\tau \in \{0.25, 0.30, 0.35, 0.40, 0.45, 0.50\}$ since for $\tau \in \{0.05, 0.10, 0.15, 0.20\}$ the correct estimation is 100% for all the three compared methods. The proposed KME-CRT outperforms the CFR-CRT due to its capacity of always finding the best reference remainder with the criterion of (2.44) as a consequence of the exhaustive search performed by the assemblage of matrix $\mathbf{Q}$ in (2.52).

Defining the root mean squared error of N as

$$N_{\text{RMSE}} = \sqrt{\mathrm{E}\{|\hat{N} - N|^2\}}, \tag{2.53}$$

Fig. 2.12 shows the results of N_{RMSE} for the proposed KME-CRT, CFR-CRT, MLE-CRT

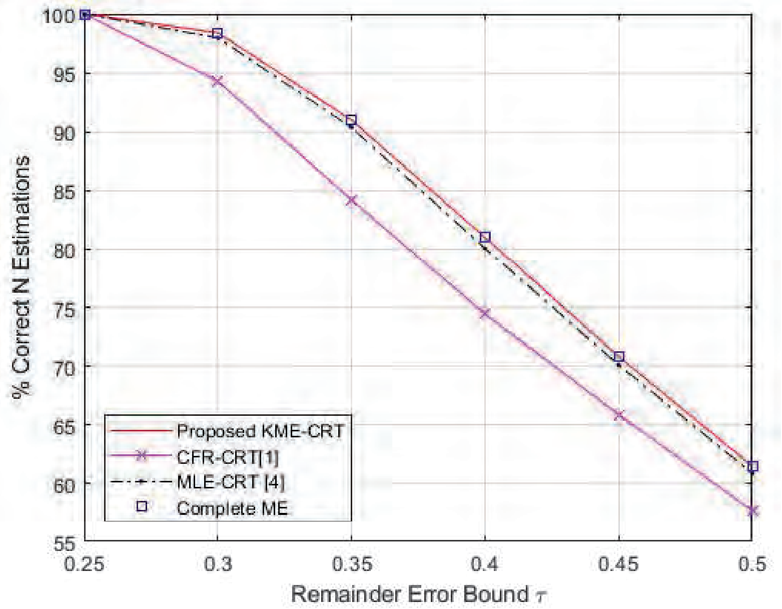


Figure 2.11: Percentages of successful estimations of N with the proposed KME-CRT, the CFR-CRT [41], MLE-CRT [42] and complete ME for constant variances in the errors.

Table II.1: % of Correct $\hat{N}$ Estimation with 2 or 3 Remainders with Errors $M/4 \leq |\Delta_i| < M/2$ over 10^5 Realizations

CRT Method	2 Remainders	3 Remainders
Proposed KME-CRT	63.49%	8.01 %
CFR-CRT	60.58 %	7.41 %
MLE-CRT	61.81 %	7.41 %

and the complete ME. The smallest values of N_{RMSE} are in the proposed KME-CRT and complete ME results, even considering the interval where the limit $\tau < M/4$ is observed, i.e., $\tau \in \{0.05, 0.10, 0.15, 0.20, 0.25\}$.

When all variances are presumably equal, Table II.1 informs the percentage of correct estimation of N in the case of two or three moduli with error Δ_i that are in the range $M/4 \leq |\Delta_i| < M/2$. The results are derived from 10^5 realizations.

The same analysis is repeated, now focusing on the case of $\sigma_i \neq \sigma_j$, for $i, j \in \{1, 2, \ldots, L\}, i \neq j$, and establishing $\sigma_i = \tau M_i$. In Fig. 2.13(a), percentages of successful estimations of N with the proposed KME-CRT, the CFR-CRT [41], MLE-CRT [42] and complete ME for different σ_i^2, with $i \in \{1, 2, \ldots, L\}$. In 2.13(b), the values around $\tau = 0.3$ in different scale. Throughout the values of τ, the percentages of correct calculations of N are higher with

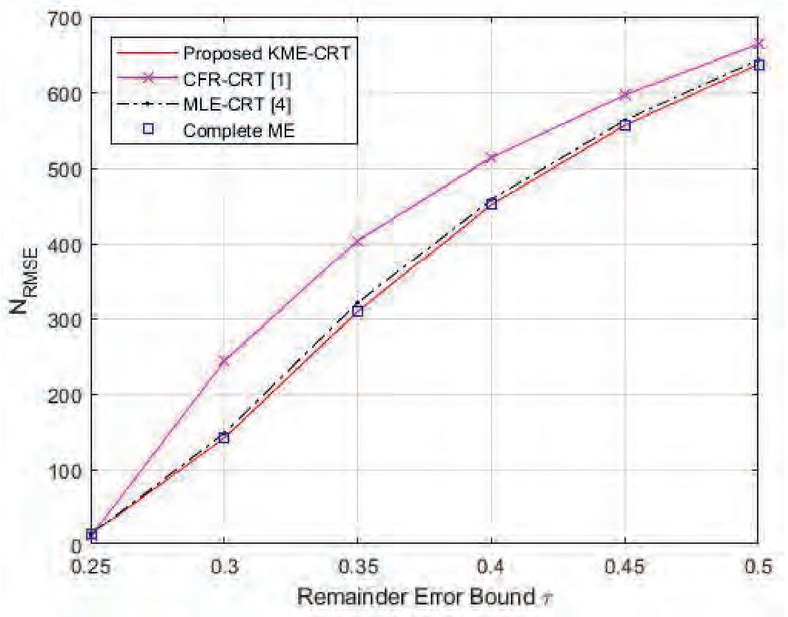


Figure 2.12: N_{RMSE} values with the proposed KME-CRT, the CFR-CRT [41], MLE-CRT [42] and complete ME for constant variances in the errors.

the proposed KME-CRT than with CFR-CRT or MLE-CRT. However, the results of the proposed KME-CRT and the MLE-CRT present significant proximity.

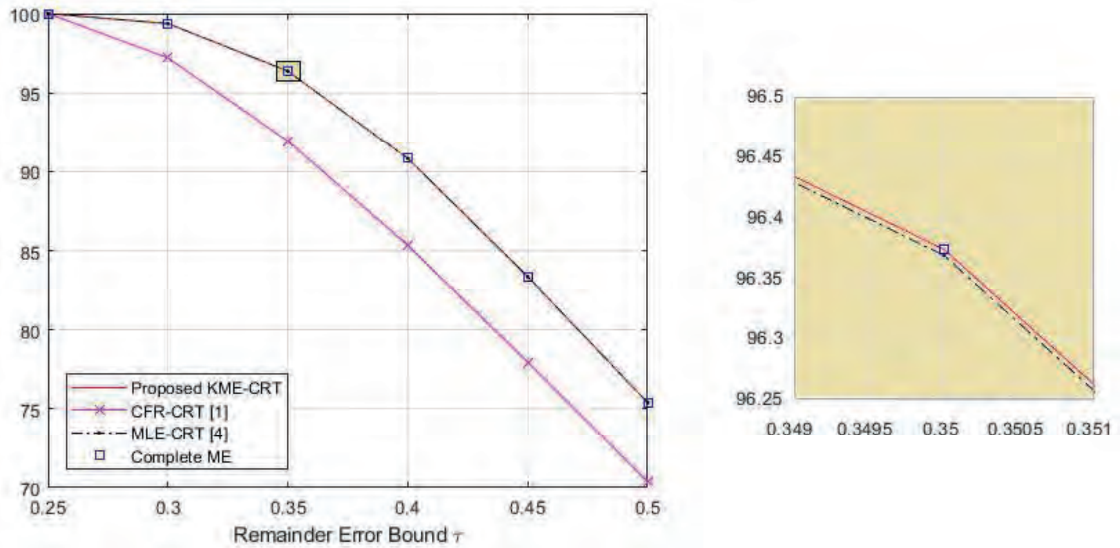


Figure 2.13: In (a), percentages of successful estimations of N with the proposed KME-CRT, the CFR-CRT [41], MLE-CRT [42] and complete ME for different σ_i^2, with $i \in \{1, 2, \ldots, L\}$. In (b), the values around $\tau = 0.3$ in different scale.

Fig. 2.14 shows the evolution of N_{RMSE} values obtained via (2.53) for the proposed KME-CRT, CFR-CRT, MLE-CRT and the complete ME. The smallest values of N_{RMSE} are in the proposed KME-CRT and complete ME results. Note that once more the results of the proposed KME-CRT and MLE-CRT are very similar.

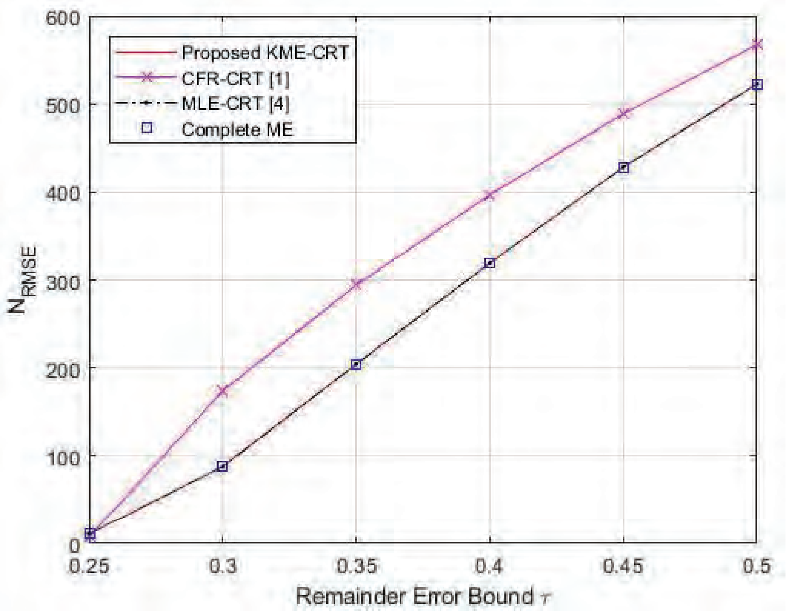


Figure 2.14: N_{RMSE} values with the proposed KME-CRT, the CFR-CRT [41], MLE-CRT [42] and complete ME for different σ_i^2, with $i \in \{1, 2, \ldots, L\}$.

We have also tested the performance of the methods for different sets of moduli through 10^4 realizations with different variances in errors considering only the case $\tau = 0.45$. Table II.2 shows the sets of moduli, which are analogous to sensors sampling rates if the CRT system is applied to undersampling systems.

Table II.2: Sets of moduli for the results shown in Fig. 2.15 for $\tau = 0.45$

Set of Moduli	Moduli of the CRT System
1	$M_i \in \{8, 11\}$
2	$M_i \in \{8, 11, 13\}$
3	$M_i \in \{8, 11, 13, 15\}$
4	$M_i \in \{8, 11, 13, 15, 17\}$

Finally, we monitor the disparity between the performances of the proposed KME-CRT and complete ME. At each of the 10^5 realizations, we count the times in which the proposed KME-CRT returns a successful estimation but the complete ME does not, and vice versa, summing it to each entry corresponding to each value of τ. Over 10^5 realizations, and in each

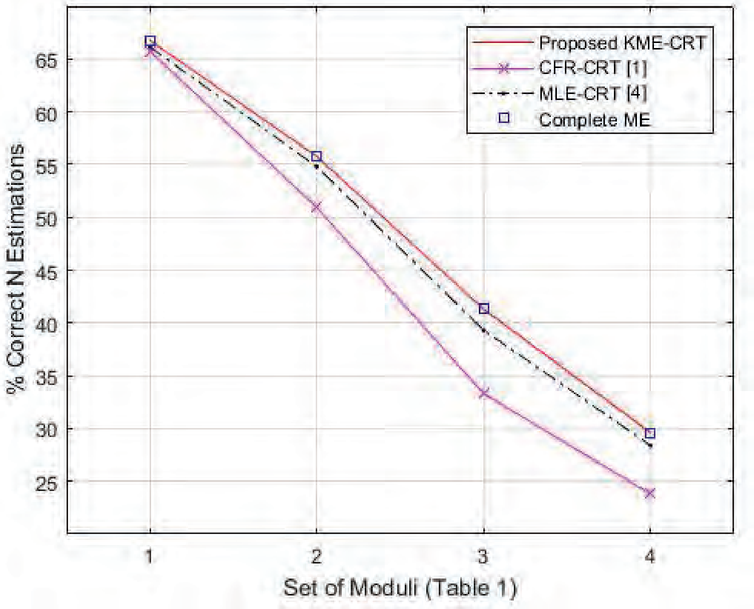


Figure 2.15: % of Correct Estimations with the CFR-CRT [41], MLE-CRT [42], the proposed KME-CRT and complete ME for the sets of moduli shown in Table II.2 with $\tau = 0.45$.

$\tau \in \{0.25, 0.30, 0.35, 0.40, 0.45, 0.50\}$, it is possible to conclude that the proposed KME-CRT and complete ME have the identical performance for practical applications. As a consequence, the proposed KME-CRT provides an optimized version of complete ME.

2.3.3 Computational Cost

The comparison of computational cost of the proposed KME, the CFR-CRT and the MLE-CRT comprises the following cases: Case 1 with $M_i \in \{7, 11\}$, Case 2 with $M_i \in \{7, 11, 13\}$, Case 3 with $M_i \in \{7, 11, 13, 15\}$, Case 4 with $M_i \in \{7, 11, 13, 15, 17\}$, and Case 5 with $M_i \in \{7, 11, 13, 15, 17, 19\}$. In each case, the time of computational processing is taken as the mean of 10^4 realizations. In sequel, we highlight the reduction of computational effort of the proposed KME-CRT in terms of complete ME. All premises adopted in Section 2.3.2 with regards to time measurements are maintained, such as avoiding parallel settings and the Matlab commands used.

Table II.3 shows the time of computational processing in milliseconds (ms) for each case. The incremental step of the proposed KME-CRT is kept as $s_i = 0.02$. The time of processing for the proposed KME-CRT increases significantly with the addition of further values in M_i as all $\mathbf{c}_i \in \mathbb{Z}^{4D}$, i.e., vectors $\mathbf{c}_i$ have the length of the dynamic range. Due to this fact, KME-CRT is supposed to be suitable for systems with low number of sensors. Table II.3

also includes the average time processing of tensorial version of the proposed KME-CRT for comparison.

Table II.3: Time of computational processing in milliseconds (ms) averaged over 1000 realizations

Moduli Set	Proposed KME	Proposed Tensor-KME	CFR	MLE	Complete ME
1: $\{7, 11\}$	0.250	0.458	0.188	0.101	0.496
2: $\{7, 11, 13\}$	0.788	1.304	0.212	0.103	8.7
3: $\{7, 11, 13, 15\}$	8.5	2.5	0.225	0.108	179.4
4: $\{7, 11, 13, 15, 17\}$	136.8	94.5	0.243	0.145	4259.1
5: $\{7, 11, 13, 15, 17, 19\}$	3068	2111	0.265	0.170	109508.4

According to the results of Table II.3 and taking into account the moduli sets 1 and 2, the proposed KME-CRT by means of the Kronecker product in Algorithm 3 outperforms the proposed KME-CRT processed in the tensorial version of Algorithm 5 in terms of processing time. However, for 4 or more moduli, the performance of the tensorial framework has better results. This can be assigned to the tensorial characteristic of lowering the maximum length of the greatest dimension involved in the routines. Combined with the parallel processing features of multidimensional data, this yields a gain in terms of computational time.

Sets of few moduli are frequently encountered in the literature, as for instance in simulations where $L = 2$ as in [60] [61] [70] [71], $L = 3$, as in [22] [43] [53] [72] [73], and $L = 4$ as in [66]. Even in [41], L assumes different values, from 3 up to 12, hence starting with low values. The case of few remainders is thus a matter of attention in the state-of-the-art, which is the most indicated case of application for the proposed KME-CRT. One can also notice that the proposed tensorial version of the KME-CRT has a lower computational time than the purely vectorial form in the cases $L = 4$ and $L = 5$, i.e., for greater values of L.

In comparison with the complete ME, however, the proposed KME-CRT shows great capability of computational economy as shown in Table II.4, where the number of rows filtered by the proposed KME-CRT and the respective reduction of computational effort are shown. The economy in comparison with complete ME is based on the fraction of the filtered rows in relation to the complete number of rows $4D$ in $\mathbf{v}$.

For instance, in case 4, $\mathbf{v} \in \mathbb{Z}^{1021020}$, but at maximum 35 rows of $\mathbf{v}$ can have 1 as value. This filtering reduces the computational effort to $35/1021020 = 0.00343$ % of the undertaken

Table II.4: Reduction of computational effort in terms of rows - Proposed KME-CRT and complete ME

Case	Proposed KME-CRT - rows of test	Economy - complete ME
1: $M_i \in \{7, 11\}$	7	97.7272 %
2: $M_i \in \{7, 11, 13\}$	11	99.7352 %
3: $M_i \in \{7, 11, 13, 15\}$	19	99.9684 %
4: $M_i \in \{7, 11, 13, 15, 17\}$	35	99.9965 %
5: $M_i \in \{7, 11, 13, 15, 17, 19\}$	67	99.9996 %

by complete ME. Nevertheless, the result that the proposed KME-CRT delivers is the same of the complete ME, as shown in Section 2.3.2.

2.4 Conclusion

In this chapter, a novel method for estimating a real number using the CRT was presented. The method is based on an ME scheme that is optimized by means of a mapping vector that indicates in which parts of the dynamic range the search for the real number should occur. This mapping vector is assembled via tensorial operations, i.e., Kronecker product of previously defined vectors. We also provide a version of the mapping vectors based on tensorial n-mode products, delivering in the end the same information of the original method. For its characteristics, it is suitable overall for CRT systems with few moduli, which in the case of sensors networks corresponds to low quantity of sensors.

In our proposal, the errors in the remainders of CRT system may have the same or different variances, allowing our work to be compared with state-of-the-art methods CFR-CRT [41] and MLE-CRT [42]. According to results tested over 10^5 realizations, in the case of equal variances, the proposed KME-CRT is consistently superior to the state-of-the-art methods in terms of percentage of correct estimations. On the other hand, with regards to the case of different variances, the superiority of our proposal is comparatively small, not outperforming the state-of-the-art MLE-CRT significantly. Hence, in this particular, both methods can be considered as of equivalent performances. However, for all that was shown, our proposed technique enhances the probability of estimating an unknown number accurately even when the errors in the remainders surpass $1/4$ of the greatest common divisor of all moduli. A drawback is that, as shown in Table II.3, the computational cost of the proposed KME-CRT increases more than linearly and surpasses the costs of CFR-CRT and MLE-CRT in certain scenarios.

The proposed KME-CRT was also compared with the complete ME, which by definition cannot be outperformed. KME-CRT has the same results of complete ME, while reduces their necessary computational effort in at least 97 %, thus offering a decisive advantage in terms of computational effort.

For future works, we envisage the need for optimization of the mapping vector as a searching method. Furthermore, errors with distributions different from the Gaussian one should be investigated. Tensor based mapping vector routines of Algorithms 3 and 5 are still under development and are also a matter of concern for future studies. In terms of CRT techniques, the possibility of applying the mapping vector to the Multi-Stage Robust CRT and studies involving CRT in a probabilistic way, as in the case of unrestricted errors of [59], are also supposed to have a good applicability towards the technique proposed here.

III

PCA KALMAN LOAD PREDICTION SYSTEM WITH RF-BASED DISTRIBUTED GENERATION FOR SENSORS IN SMART GRID

3.1 Introduction

Electrical load forecasting is an essential activity for formulating strategies, planning and operation of electric systems. Several factors affect the load behavior in both spatial and time domains. Relevant variables include weather, demographics, economic production, cultural habits and calendar events. Evidently, there are random load components related to the natural variations observed in industrial processes and to the unpredictable character of the human behavior. These components are also usually nonlinear in nature, which creates further difficulties in the selection and calibration of suitable models.

The effect of weather on electricity consumption is researched since the first half of the 20th century. In [74], the influence of weather variables and conditions over the South East England power system is discussed, stressing the effects of decreased temperature over mean and peak load. The concept of degree-day relates to a variable that, given the external temperature value, measures the amount of energy needed to heat or cool a building to a comfortable temperature. Since the last quarter of the 20th century, degree-days are used as tool for energy consumption forecast [45, 46]. Currently, heating (HDD) and cooling degree-days (CDD) have been featured in several load forecast methods, such as [75–77].

Other important aspects of electricity demand are associated with economic activities, energy prices, industrial production and running stock of electric appliances. The causal relationship between economic growth, characterized in diverse indicators, and the electricity consumption is investigated in numerous papers. In [78], Granger tests indicate short-run causality between energy consumption and income for India and Indonesia, while the test points to bi-diretional relationship for Thailand and the Philippines. Several variables are used to assert the dependencies between energy consumption and economic activities: Gross Domestic Product (GDP), population and price indexes [79].

After the selection of a suitable set of input variables related to the electricity demand, a prediction algorithm can be chosen. The survey presented in [80] discusses the most relevant studies on electric demand prediction over the last 40 years. Still in [80], classification of different models based on input variables, forecasting horizon and linear or non-linear prediction are shown. In terms of linear methods, they refer to multiple regression [75], Box and Jenkins (B&J) models [47,81] and State Space approaches [82,83], while, in terms of non-linear methods, [80] cites Artificial Neural Networks (ANN) [84], fuzzy logic and Grey theory models. The conclusions in the survey are favorable to the nonlinear approaches, praising the ability to generalize and detect nonlinearities. According to [80], nonlinear approaches benefit from their ability to generalize and detect nonlinearities, while linear approaches are suitable for large amounts of data and benefit from precise formulas to discover nonlinearities.

As a forecast technique, this chapter proposes a Kalman filter based short term load forecasting system that benefits from known dependencies to extract optimized weights of input variables, which are selected from candidate time series of distinct sources. The Kalman filter is an algorithm permitting exact inference in a linear dynamical system, where the State Space of the latent variables is continuous and all latent and observed variables have a Gaussian distribution [85]. The great success of the Kalman filter consists of its small computational requirement, elegant recursive properties, and its status as the optimal estimator for one-dimensional linear systems with Gaussian error statistics [86]. In comparison with the ARMAX filters used in [47], the superiority of Kalman stems from its adaptive features, the ability in treating time series data as a Markovian process – as the last time series snapshot condenses all previous information – and dealing with non-stationary data series. We combine Kalman with Principal Component Analysis (PCA), which exploits whether certain variables of a process are correlated. If some degree of redundancy exists, PCA algorithms provide a reduction in dimensionality of the data source by removing the redundant content. Hence, the PCA algorithm reduces data size while preserving its core information for application of the Kalman filter.

In real applications, data about relative humidity, dew point, temperature, weather phenomena and altimeter barometric pressure feature measurements of phenomena such as fog, rain, thunderstorms and snow. All the data about such phenomena are obtained via sensors, which are often installed over relatively wide areas whose access for people is impracticable. However, sensors consume energy continuously, which imposes the achievement of an optimal energetic management. In this context, the concept of energy recycling plays an important role. Among several forms of recycling energy, radiofrequency (RF) harvesting [4, 5] has been suggested due to its wide availability mainly in urban areas. Its applications range from sensor nodes to charging low power consumption portable devices and depend on the amount of antennas.

We consider that RF energy recycling systems are installed next to the sensors used to measure weather data in order to enhance their stand-by autonomy without the need of accessing the nodes frequently. As a way of proving the feasibility of this solution, we perform a RF incidence measurement campaign in Brasilia, Brazil. Inspired by antenna array communication systems, our second contribution in this chapter is to propose an improved RF energy recycling system based on a rectenna array. We show that, by increasing the amount of rectennas, a significant gain due to the array is achieved. We also show that rectenna arrays can outperform standard antenna arrays as energy harvesting systems, quantifying the improvement as a function of the number of employed receiver terminals.

The remainder of this chapter is organized as follows: Section 3.2 shows the variables adopted as candidate inputs. Section 3.3 explains the proposed Kalman based load forecasting system. Section 3.4 presents the RF energy recycling state-of-the-art concepts and the proposed power harvesting system based on rectennas. Section 3.5 presents the prediction results compared to benchmarking predictors. Section 3.6 concludes the paper and also indicates directions for future work.

3.2 Candidate Input Variables

Several factors are known to affect energy demand: temperature, climate events, energy tariffs, demographic indicators, economic indexes, social conventions and cultural traditions. Worldwide, with the ample access to information technology, a diversity of data can be collected and substantial volumes of time series related to electricity demand can be processed. In Sections 3.2.1, 3.2.2, 3.2.3 and 3.2.4, we present respectively the variables related to weather, socioeconomic factors, energy tariffs and calendar events that are used in our forecast model. All these variables stand for exogeneous inputs.

Fig. 3.1 depicts the candidate variables per types, numbers and sections. The sets are referred to through the letters A to H. Set A comprise 3 variables – minimum, average and maximum temperature of the day – while $B \supset A$ and contains further 19 variables, thus with

22 variables in total. Set C includes 10 variables related to weather features. Set D bears 32 variables since $D = B \cup C$. From the socioeconomic part, 13 socioeconomic variables integrate set E. Tariffs and calendar variables are in the sets F and G, respectively. In our framework, sets D, E and F count on data for the day of load prediction and the day before. Hence, for variable sets D, E and F there are two input values at each realization. Variable set G is boolean, containing only values of the day of load prediction, i.e., $H = D \cup E \cup F \cup G =$ 250 variables are part of the set H, which considers all the previous sets together. As a consequence, without PCA pre-processing, the number of exogeneous variables is 250.

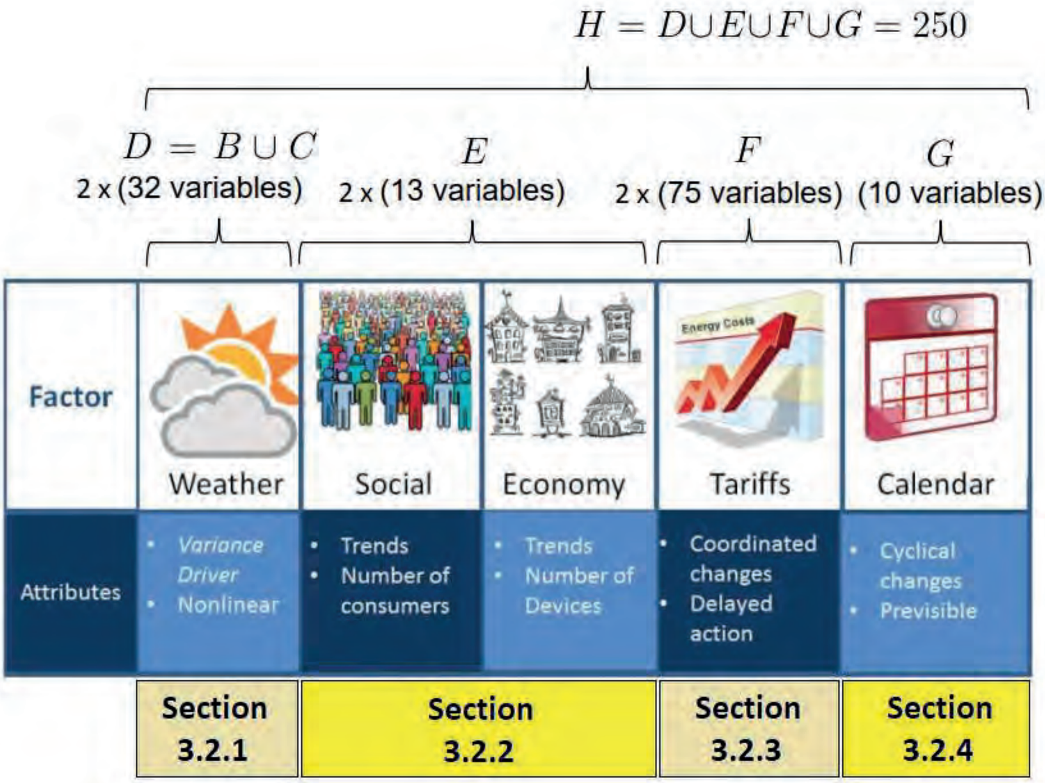


Figure 3.1: The candidate variables sets per types, numbers and sections. Note that $B \supset A$, sets D, E and F contain two data for each prediction day, and that $H = D \cup E \cup F \cup G$, thus comprising all described variables with 250 variables in total.

From the standpoint of load prediction, an innovative aspect of this Chapter is the forecast framework for Brasília that takes into account of a vast set of variables by employing PCA pre-processing and Kalman filters. To the best knowledge of the authors, the proposed prediction system is so far the most comprehensive with regards to the variables that might affect load behavior in Brasília.

3.2.1 Weather Variables

The historical weather data has been collected from the Juscelino Kubistchek International Airport METeorological Aerodrome Reports (METAR)[1], located in Brasília. A METAR presents hourly information about wind speed, wind direction, visibility, relative humidity, dew point, temperature, weather phenomena and altimeter barometric pressure, as measured or observed in the surface. The reports also feature measurements of cloud cover and indicative codes for weather phenomena such as fog, rain, thunderstorms and snow.

Since buildings are modelled as a four sided heat exchanger with distinct heat conduction and convection coefficients at each compass point, wind speed variables are decomposed in four components: north, south, east and west. Such division is performed by means of trigonometric transformations of the average wind speed and average direction as reported in METAR.

In this forecasting system, the maximum, minimum and average temperature compose the classical weather input Set A, with 3 variables. Set A and the remaining METAR variables compose the Set B, with 22 inputs.

3.2.2 Socioeconomic Variables

For the population input, time series of Brasília region are obtained in the Brazilian Institute of Statistical Geography (IBGE[2]) database. Since the time series have only monthly or annual values, daily values are obtained by cubic splines. A similar approach is executed for the GDP input. Dividing GDP by the population, one can obtain the GDP per capita. A modified rolling grey algorithm as described in [87] is applied to simulate forecasts of these candidate input variables, as most of these indicators cannot be collected in real time.

Additional variables are taken into account, such as the fraction of low income households, the relative sales volume index and the energy intensity indicator for industries with low, medium and high specific energy consumption, relative imports and export indexes, life expectancy at birth, basic sanitation at residences and birth rate. Also present are the official price index (IPCA), dollar exchange rate and GDP in US dollars. The 13 socioeconomical variables are designated as input set E in Section 3.5.

3.2.3 Energy Tariffs

In Brazil, low voltage customers only have access to the conventional monomial tariff, in which there is a fixed tariff for energy ($/kWh). High voltage clients must adhere to a

[1] Wunderground Database, retrieved from http://www.wunderground.com/

[2] IBGE Database, retrieved from http://seriesestatisticas.ibge.gov.br/

binomial tariff contract, in which there two rates: one for energy ($/kWh) and another for demand ($/kW). There is a surtax if the demanded power is higher than the contract limit. The tariff type can be conventional, hourly seasonal type green or hourly seasonal type blue, moving from fixed rates for demand, energy and surtaxes to different rates due to seasons and peak hours. In Fig. 3.2 a 60-day moving average representation of tariffs (conventional type) by consumer classes is presented.

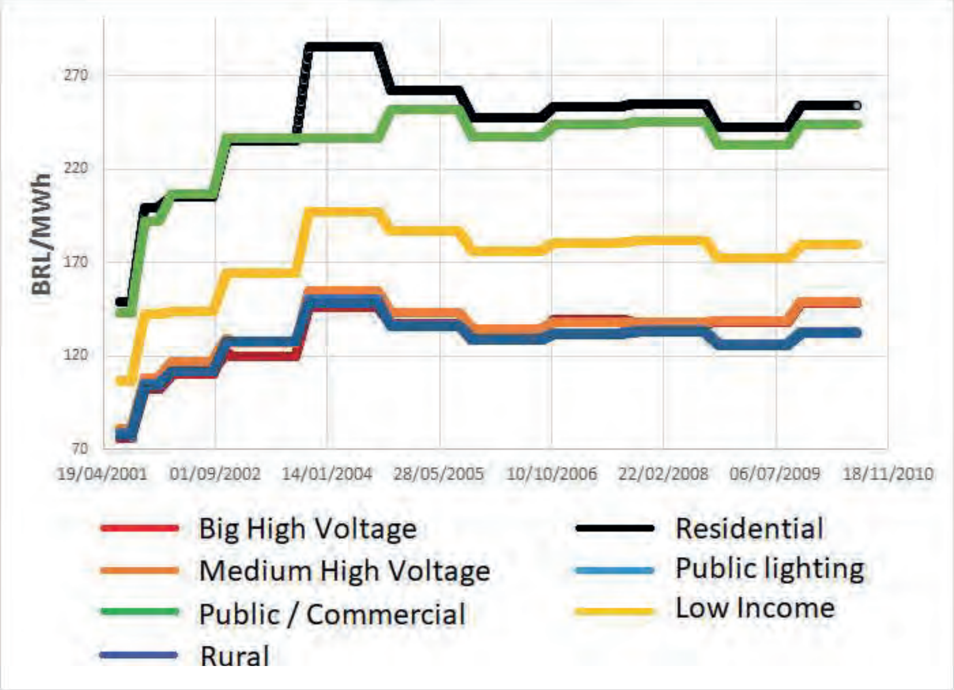


Figure 3.2: 60-day moving average of electricity tariffs in Brasília, in Brazilian Reais (BRL) per MWh, by consumer class

Due to the multitude of classes, types and seasonal periods, the historic tariffs time series is composed of 75 candidate variables, being 11 low voltage conventional, 10 high voltage conventional, 18 hourly seasonal type green and 36 hourly seasonal type blue. It is designated as input set F in Section 3.5.

3.2.4 Calendar Events

The load profiles have markedly distinct behavior in working days, holidays and weekends. There are also atypical days [88] with different load curves, such regular day preceding or following a holiday. Large media and sports events can also lead to uncommon behavior in the electricity demand.

While introducing additional variability to the forecasting problem, calendar events have the advantage of being known *ex-ante*, which can be represented as a binary variable. These can be described as boolean time series that have a true value when the event is expected, being it false otherwise.

Expanding on other papers characterizations, for this load forecasting system there are binary variables for each day of the week, for summer saving time, for holidays and uncommon days. The latter are classified as such due to proximity to other holidays or the occurrence of major media or sports events, such as important soccer matches of 2002, 2006 and 2010 World Cups. These 10 variables are designated as set G in Section 3.5.

3.3 Methodology

Principal Component Analysis (PCA) exploits whether certain variables of a process are correlated. If some degree of redundancy exists, PCA algorithm provides a reduction in dimensionality of the data source. Therefore, PCA aims at data reduction of data dimensionality while preserving the maximum of its variability [89]. In [90], PCA algorithm is used to extract the main features of the original data, removing redundant content. By removing the unnecessary content, PCA naturally enhances Kalman filter performance.

The purpose of a Kalman based load forecast system is to periodically and recursively select the most explanatory set of input variables from a given set of candidates and reliably predict the base, average and peak demand for the next day of operations. Our forecasting system can be divided in three functional blocks, as shown in Fig. 3.3.

3.3.1 Preprocessing

The preprocessing block prepares the candidate variables to be combined and selected in the PCA, either by normalizing mean and variance of the candidate set, as by applying nonlinear operations.

The mean and variance normalization is a simple procedure designed to enforce uniformity in the amplitude scale of the candidate variables, except for those that are boolean. This is done with the goal of minimizing numerical errors. Taking a sample of a given length n of the i-th candidate variable $\hat{U}_{0i}$, which has mean $\overline{U_{0i}}$ and variance σ_{0i}^2, it can be normalized to zero mean and unitary variance by the linear operation as follows:

$$\hat{U}_{1i} = \frac{(\hat{U}_{0i} - U_{0i})}{\sigma_{0i}} \tag{3.1}$$

Nonlinear operations are performed on variables that have well documented relationships with electric load. In the proposed load forecasting system, they comprise temperature

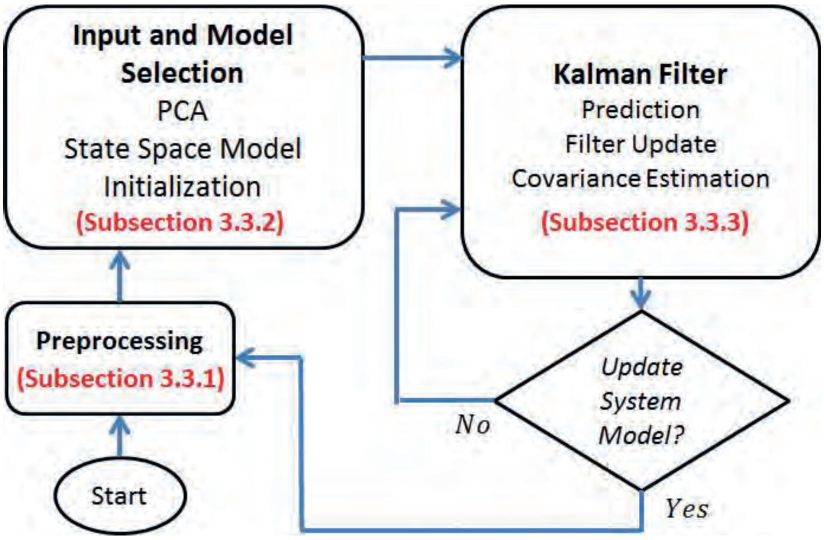


Figure 3.3: Block diagram of the proposed load forecasting system

variables, humidity and wind variables, which are transformed in heating degree-days, cooling degree-days, enthalpy latent days and power law convection and psicometric coefficients. For instance, the load response to temperature is known to behave nonlinearly, specially at cold and hot extremes [75, 77]. This is verified in Fig. 3.4, a scatter plot of maximum daily temperature and peak demand.

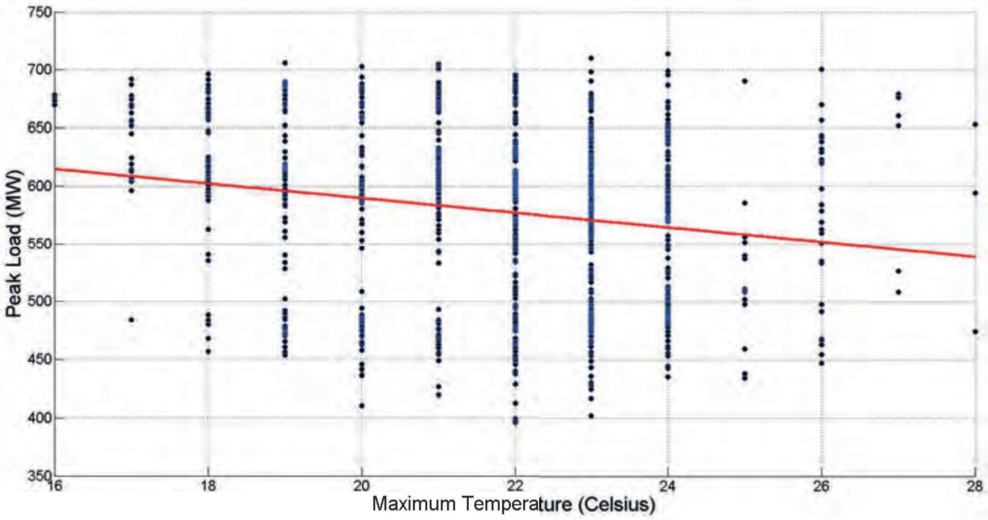


Figure 3.4: Scatter plot of maximum temperature and peak demand in Megawatts (MW). The large number of virtually uncorrelated points suggest the coefficient provides a poor fitting.

The Cooling Degree-Days (CDD) is a measure of the severity and duration of hot weather, defined as the integral sum of the subtraction between a given reference cooling temperature and the ambient temperature over time. The CDD values are estimated by the United Kingdom Meteorological Office (MET Office) method [91], that is simpler yet reasonably accurate and only requires minimum and maximum temperatures. The relationship between the CDD and the electric demand is stronger than the unprocessed temperature value. Fig. 3.5 presents a scatter plot that illustrates the correlation between the two variables. Because the large number of non-correlated points stays at zero degree-days, they do not affect the determination of the CDD coefficient.

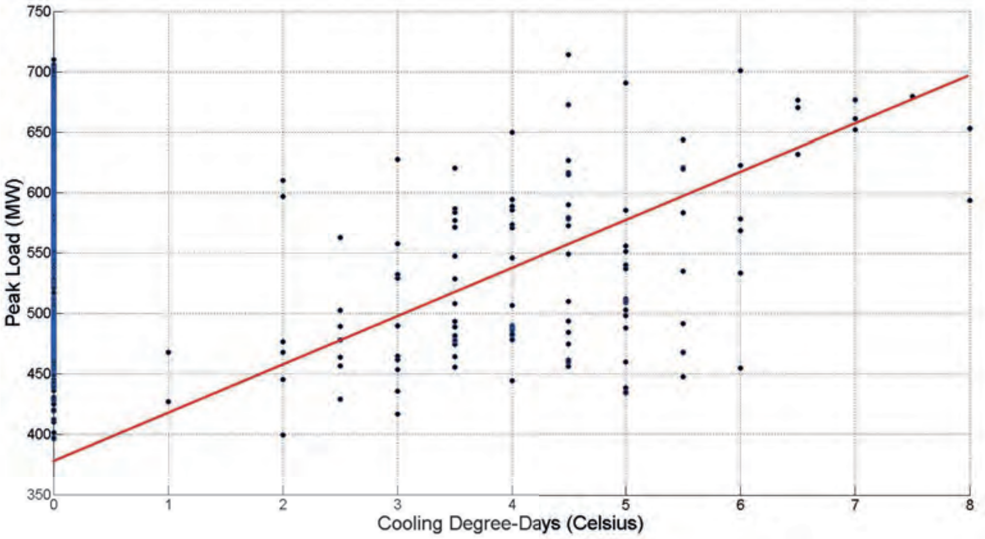


Figure 3.5: Scatter plot of CDD and peak demand in MW. Because the large number of non-correlated points stays at zero degree-days, they do not affect the determination of the CDD coefficient.

Similarly, the Heating Degree-Days (HDD) is a measure of the severity and duration of cold weather that relates to the heating requirements. However, as Brasília has a hot climate in average, HDD requires a different parametrization than the performed to model building heating demand. In Brazil, water heating by electric showers is prevalent. This load behavior requires higher reference temperatures as water has a greater thermal conductivity than air.

In the proposed load forecasting system, three CDD and three HDD variables are created, with reference temperatures of 16, 20 and 24 for the CDD and 22, 26 and 30 degree Celsius for HDD. These variables model the nonlinear relationship of cooling and heating requirements to the electric demand as a piecewise trilinear function.

Enthalpy latent days (ELD) indicates the amount of energy required to remove excessive moisture from the outdoor air without reducing the indoor air temperature, hence lowering the indoor humidity to an acceptable level [92]. Eq. (3.2) defines Enthalpy latent days as the summation of positive enthalpy differences between the outdoor air enthalpy h_0 with relative humidity x_0, and enthalpy h_b with indoor reference relative humidity x_b. For both enthalpies at the outdoor air temperature θ_0, the reference humidity is set in our load forecasting system as 50%:

$$\mathrm{ELD} = \sum_{t=1}^{24} \left[h_0\left(\theta_0, x_0[t]\right) - h_b\left(\theta_0, x_b\right)\right] \tag{3.2}$$

The wind can have observable effects in electricity consumption. To model such dependencies, the directional wind inputs are transformed by means of power laws to provide the convection coefficients h_c for the heat transfer modeling, as in

$$\dot{q} = h_c S(T - T_0) \tag{3.3}$$

where the heat transferred per unit time $\dot{q}$ is a function of the convective heat transfer coefficient h_c, the contact area S and the difference between T and T_0, respectively the temperatures of the object and the fluid, and

$$h_c = 5.14 v_w^{\alpha}, \tag{3.4}$$

where v_w^{α} is the air speed raised to power of α. In [93], a power law approximation in the form of Eq. (3.4) is inferred, relating the coefficient h_c to v_w^{α}. In agreement with [93], the exponent $\alpha = 0.82$ is used to create four additional variables to model heat convection on facades oriented to each compass point. Exponent $\alpha = 2$ is employed in maximum wind and average wind non-directional inputs to model human comfort psychometric functions. These 10 nonlinearly transformed variables produce the set designated as C in Section 3.5.

3.3.2 Input and Model Selection

This load forecasting system employs PCA to search and select the input variable set that better explains the variance in electric demand, by means of linear combination of the candidate variables that generate a set of orthogonal inputs, called principal components. A method to reduce dimensionality is to select the j components with higher variance that explain a given percentage of the candidate set total variance, discarding the other components altogether. Composed of 250 variables, the original set displays high cross-correlation between the input themselves, as presented in Fig. 3.6.

PCA is applied at a training sample of the d_0 candidate variables, assembled in this forecasting system from their 360 previous values. The size of d_0 can be as high as 250, when all candidate inputs are those presented in 3.2. The objective is to reduce the dimensionality of the input set from d_0 to d. By means of a SVD decomposition, the left-singular vectors, the

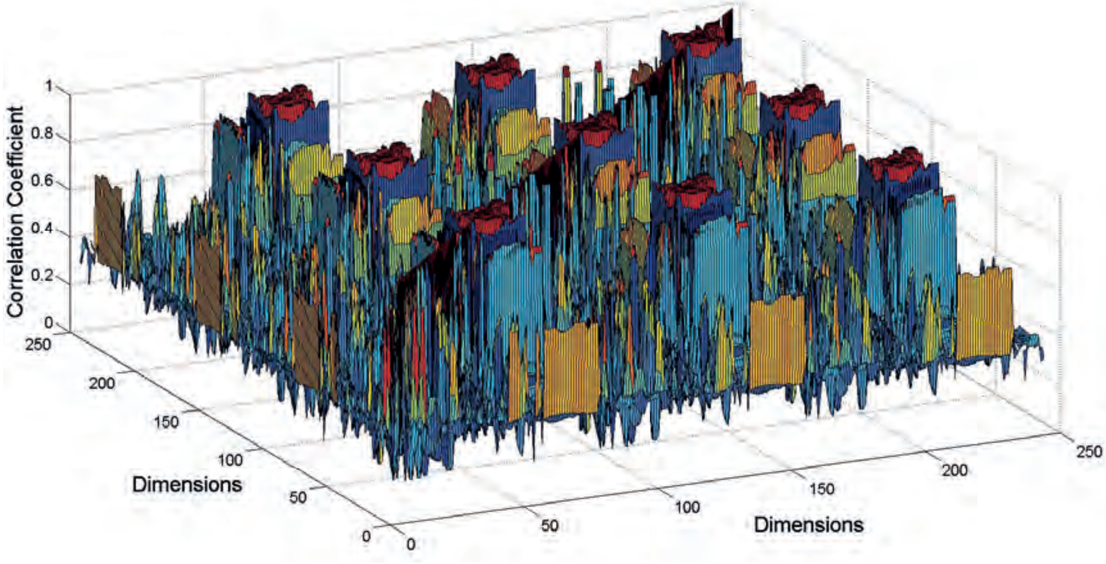


Figure 3.6: Correlation between 245 candidate variables and peak demand, which corresponds to the projection in the planes $x = 0$ or $y = 0$

singular values and the right-singular vectors are obtained. The d singular values that represent 99.999 % of the total variance are selected, their quantity determining the dimension in the selected input set. The left and right-singular vectors are then employed to produce the transformation matrix T.

For prediction, as the next day d_0 values of the candidate variables become available, they are transformed by T in a optimized input of d variables, which are used for the prediction of next day electric load. In order to adapt to seasonal variations, this process is repeated at every 120 days. The cross-correlation of optimized input set with 126 variables is shown in Fig. 3.7, obtained from 250 candidate inputs at the first realization.

The initialization procedure sets the initial parameters that the Kalman based predicting scheme needs in order to operate reliably. The first parameter to be set is the model order, which sets n state variables to be employed and the maximum input delay q. For the training dataset, the scheme needs 120 days of past data, which are the previous electricity demand and the exogenous input time series. A range of candidate model orders is then simulated over the training dataset as to peak the model with the smallest MSE. Note that n stands for the Auto-Regressive (AR) component. After testing the values for $n \in \{1, 2, \ldots, 14\}$, the lowest average error is achieved with $n = 7$, and we assume this as the AR model order for all the prediction periods. The maximum input delay is one day, $q = 1$, i.e., the candidate variables D, E and F give values for the day of prediction and the day before. After selection of the input variable set, we generate a suitable State Space model that can lead to short term load forecast by means of a Kalman predictor algorithm.

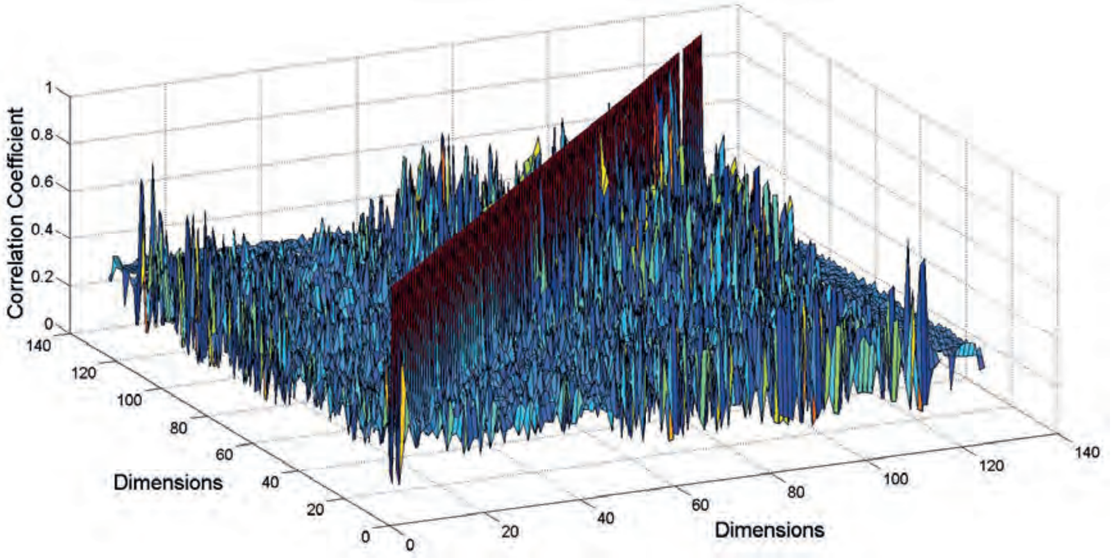


Figure 3.7: Correlation between the 126 selected variables and peak demand, which corresponds to the projection in the planes $x = 0$ or $y = 0$

3.3.3 Kalman Filter

The Kalman filter model [94] assumes that the state of a system at an instant $[k+1]$ evolves from the prior state at instant $[k]$. Mathematically, the State Space model is represented by

$$\mathbf{x}[k+1] = \mathbf{A}\mathbf{x}[k] + \mathbf{B}\mathbf{u}[k] + \mathbf{w}[k] \tag{3.5}$$

and

$$\mathbf{y}[k+1] = \mathbf{C}\mathbf{x}[k+1] + \mathbf{v}[k+1], \tag{3.6}$$

where $\mathbf{x}$ is the vector representative of the state variables, $\mathbf{u}$ denotes the vector of exogenous inputs and $\mathbf{y}$ relates to the output. Vectors $\mathbf{w}$ and $\mathbf{v}$ stand for system and measurement uncertainties modeled as zero mean. In (3.5) and (3.6), matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ are commonly specified in accordance with the characteristics of the process that is modeled, when it is known at prior. This is not the case in the present work, as the content of matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$, which establish the influence of each set of temporal data, is yet to be determined. In this case, an efficient model of coefficients determination is drawn from [95], where the matrices **A**, **B** and **C** assume the forms

$$\mathbf{A} = \begin{bmatrix} a_1 & a_2 & \cdots & a_{n-1} & a_n \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix}; \mathbf{B} = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1m} \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}, \tag{3.7}$$

and

$$\mathbf{C} = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}, \tag{3.8}$$

where $\mathbf{A} \in \mathbb{R}^{n\times n}$, $\mathbf{B} \in \mathbb{R}^{n\times m}$ and $\mathbf{C} \in \mathbb{R}^{1\times n}$. While n stands for the model order of the AR component, m is the set of all exogeneous data considered. As mentioned in Section 3.3.2, $n = 7$. With regards to m, in Section 3.2 the number of 250 variables for set H is explained. By pre-processing the exogeneous data via PCA filters, the number m of exogeneous inputs, as a rule, is reduced since part of the candidate variables is excluded. As shown in Section 3.3.2, the value of m in terms of H is reduced to $m = 126$ after applying PCA. Similar reduction of m occurs to each set from A to G when tested individually.

Due to the specific State Space representation that is employed for the predicting scheme, only the elements in the first row of matrices **A** and **B** matrices must be determined by linear least squares. As a consequence of the large amount of data, the elements are computed via an iterative Generalized Minimum Residual (GMRES) method.

The process noise of **w** and **v** are assumed to be drawn from zero mean multivariate normal distributions with covariances given by the covariance matrices **Q** and **R**, respectively. Thus, the Kalman filter is a time domain technique that relates inputs, output and state variables through (3.5) and (3.6).

In this load forecasting system, the predicting algorithm consists of the recursive repetition of Eqs. (3.9) to (3.14). The matrix **K** is the Kalman gain, **P** is the error covariance matrix for the state estimate **x**, and $\mathbf{I_n}$ denotes the identity matrix of order n.

$$\hat{\mathbf{x}}[k+1] = \mathbf{A}\mathbf{x}[k] + \mathbf{B}\mathbf{u}[k] \tag{3.9}$$

$$\hat{\mathbf{y}}[k+1] = \mathbf{C}\hat{\mathbf{x}}[k+1] \tag{3.10}$$

$$\hat{\mathbf{P}}[k+1] = \mathbf{A}\mathbf{P}[k]\mathbf{A}^T + \mathbf{Q}[k] \tag{3.11}$$

Note that Eqs. (3.9) to (3.11) are calculated before the measurement of the electricity demand, while the remaining filter equations improve the predictions with the information gained by the measurement. In sequel,

$$\mathbf{K}[k+1] = \hat{\mathbf{P}}[k+1]\mathbf{C}^T(\mathbf{C}\hat{\mathbf{P}}[k+1]\mathbf{C}^T + \mathbf{R}[k])^{-1} \tag{3.12}$$

$$\mathbf{x}[k+1] = \hat{\mathbf{x}}[k+1] + \mathbf{K}[k+1](\mathbf{Y}[k+1] - \mathbf{C}\hat{\mathbf{x}}[k+1]) \tag{3.13}$$

$$\mathbf{P}[k+1] = (\mathbf{I} - \mathbf{K}[k+1]\mathbf{C})\hat{\mathbf{P}}[k+1], \tag{3.14}$$

where $\mathbf{K}[k+1]$ is the Kalman gain matrix $\mathbf{K}$ updated at the instant $[k+1]$. Likewise, matrices $\mathbf{P}$, $\mathbf{Q}$ and $\mathbf{R}$ are equally updated at each instant, as the Kalman filter is adaptive. Adding to the original set of Kalman filter equations, the predicting block also employs variance estimation steps, shown in Eqs. (3.15) to (3.19).

$$\mathbf{V}[k] = \mathbf{Y}[k] - \mathbf{C}\mathbf{x}[k] \tag{3.15}$$

$$\mathbf{R}[k+1] = \frac{1}{k}\mathbf{R}[k] + \frac{(k-1)}{k}\mathrm{Var}(\mathbf{V}[k]) \tag{3.16}$$

$$\mathbf{w}[k] = \mathbf{x}[k] - \hat{\mathbf{x}}[k] \tag{3.17}$$

$$\triangle\mathbf{Q} = \sqrt{(\mathrm{Var}(\mathbf{w}[k])^2 - \mathbf{I_N} \cdot \mathrm{Var}(\mathbf{V}[k])^2)} \tag{3.18}$$

$$\mathbf{Q}[k+1] = \frac{1}{k}\mathbf{Q}[k] + \frac{(k-1)}{k}\triangle\mathbf{Q} \tag{3.19}$$

After Eq. (3.19), the algorithm moves ahead to the next time step and repeat the process, starting from Eq. (3.9). The load forecasting system has the input set and State Space model refreshed at every 120 time steps.

3.4 RF Energy Harvesting

In Section 3.4.1, we review the main definitions and results in the state of the art with regards to RF harvesting systems. In Section 3.4.2, we describe the measurement campaign for collecting data of the dBm incidences in four places of Brasília, Brazil. Section 3.4.3

presents the main differences between rectennas and antennas harvesting RF circuits in terms of constructive aspects. Simulations for comparing rectennas with antennas performances for RF harvesting systems in urban environments are presented in Section 3.4.4.

3.4.1 Theory and State of the Art

The main objective of the RF energy harvesting system is to convert the existing RF power from the space into usable direct current (DC) electrical source [96]. The information of the received signals is not relevant, as the signals captured by the antenna are converted into energy in order to recharge the device [97]. Since environmental RF levels are lower than those that can be provided by a dedicated RF source, the efficiency of the harvesting system and its minimum startup power are of critical importance. RF recycled power from TV broadcasts is 100 times weaker than solar power [98]. On the other hand, RF energy from TV broadcasts can obtain power 24/7, except during the maintenance periods.

Even though the transferred power from a broadcast station has the order of more than several kilowatts, the received power has the order of microwatts to milliwatts. Therefore, the main utilization of RF harvesting energy systems is for Wireless Sensor Networks (WSN), since their power consumption is very low. Generally, a sensor node consumes more than 10 mW of power in order to transmit a packet over a wireless link [99]. In many instances, only a few milliwatts are needed to power wireless sensors. More commonly used wireless sensor nodes consume dozens microwatts in sleep mode and hundreds microwatts in active mode [100]. As a consequence, the RF harvesting shows to be a promising solution for WSN as they may operate in isolated areas where recharging is not possible.

Despite advancements in end-to-end circuits (i.e., input RF to output DC), only a few power conversion attempts with low input RF power levels at true ambient RF energy harvesting have been reported. For example, a relatively efficient rectenna utilizing a modified omnidirectional patch antenna has an efficiency of 18 % with a single-tone input RF power of 20 dBm [101], which illustrates how limited harvesting RF waves energy can be. In order to enrich our analysis, the maximum dBm power that is transmitted from the antennas as in [102] is shown in Table III.1.

In terms of GSM signals, we may expect a power density between 0.01 and 1.0 mW/m^2 ($10^{-3} \sim 10^{-1}$ μW/cm^2) indoors everywhere or outdoors on an elevated level, taking into account distances between 25 m and 100 m from a GSM900 base station [103]. Considering the power integrated over the downlink frequency band (935 $\sim$ 960 MHz), we may expect a total power density between 0.1 and 3.0 mW/m^2. The power density received from GSM1800 base stations are, up to 100 m, is in the same order of magnitude as those received from GSM900 base stations at a single frequency or summed for low traffic situations [104].

Table III.1: Frequency bands of energy harvesting and respective maximum dBm power [102]

Standard	Frequency Bands	Band of Interest	Max. Power (dBm)
DTV	470 - 862MHz	470 - 862MHz	70
GSM	900, 1800 MHz	925-960 / 1805.2-1879.8 MHz	40
UMTS	2100 MHz	2110-2170 MHz	40
Wi-Fi, Bluetooth	2.4 GHz	2.4 GHz	30, 20
New Wi-Fi	5 GHz	5 GHz	NA

Multiple-Input Multiple-Output (MIMO) broadcasting systems for simultaneous power and information transfer were studied by [105]. The approach is usually based on adopting an energy receiver and an information receiver as two distinct devices on the receiver side. In [106], an optimal linear beamforming vectors for Simultaneous Wireless Information and Power Transfer (SWIPT) with a two-user MISO system is exploited. Each of two transmitters sends data or transfers energy to its corresponding receiver, behaving either as an information decoder or an energy harvester device, but not as both at the same time. Therefore, techniques dedicated solely to power recycling are still up to date in terms of energetic usage of broadcasting systems, since the information management and energy recycling are always independent.

In [107], a work was carried out at 900 MHz with 50 Ω impedance, using a resonance circuit transformation coupled with a Schottky diode. This scheme yields a DC output voltage of over 0.3 V for an input power level of -26 dBm (2.5 μW). A DC voltage of 0.8 V was achieved from RF input power level of -20 dBm (10 μW) at 868.3 MHz through simulation with no load [108]. A Cockcroft-Walton multiplier circuit was used and produced 1.0 V DC voltage onto a 200 MΩ load for an input power level of 1.0 μW at a fixed frequency of 2.4 GHz [109].

According to [110], experimental results involving the operation frequency of 945 MHz, conversion efficiencies of -5 dB, 0 dB and +5 dB corresponded, respectively, to 3 %, 5 % and 7 %, with R = 100 kΩ, and 20 %, 23 % and 25 % for R = 50 kΩ. This analysis presupposes the employment of a two-stage Dickson voltage multiplier. The efficiency at 0 dBm was 21 %, whereas around -5 dBm the efficiency was 3.2 % [103]. There is still an example of incident power in high buildings from [111]. Measurements within the Department of Electrical and Electronic Engineering building at Imperial College of London were taken on the 11th floor of the south stairwell, which are shown in Fig. 3.8. One must note that it is related to the end-to-end efficiency, which is the quotient between the time-averaged output (i.e., equivalent DC) power into the storage element and the time-averaged input RF power; in other terms, it depends strongly on the efficiency of the harvesting system, but also are relevant as examples for the end-to-end efficiency.

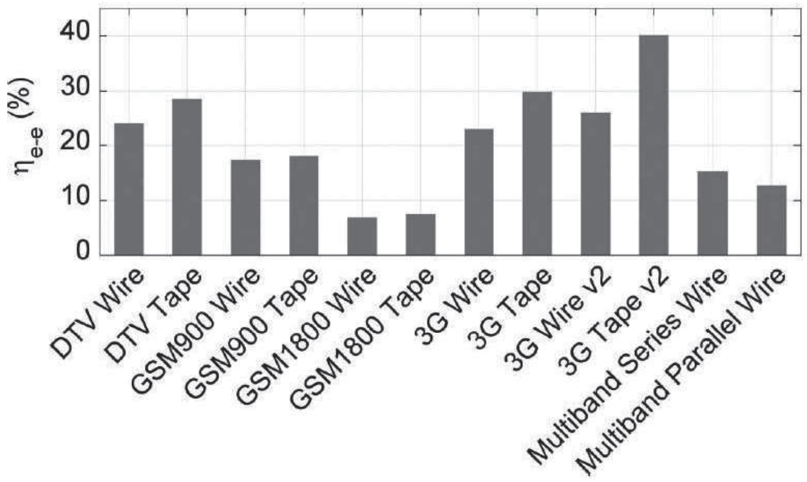


Figure 3.8: End-to-end efficiencies for ambient RF energy harvesting [111].

3.4.2 RF Power Measurement Campaign Performed in Brasília, Brazil

Measurements were undertaken with the aim of evaluating the use of RF waves as an ubiquitous energy source. Two stages of incident dBm measurement took place, first (i) with measurements from four points of Brasilia-DF (Brazil) with the aid of a spectrum analyzer, and (ii) a registers of dBm values in the vicinity of the TV Tower. The measurements were all carried out approximately 1.7 meter above the soil.

3.4.2.1 Stage 1 – dBm survey of the four designated points

The four surveyed points stay near two main targets of research: the Digital Tower and the TV Tower. For each of the two targets, two sites were chosen to the verification of the dBm incident power. The sites selected to take measurements from the Digital Tower were Place 1 (15°41'29.70"S/ 47°51'13.96"O), on the border of the road leading to the Digital Tower, and Place 2 (15°41'59. 46"S/47°49'50.38"O), right in front of this target. In order to take measurements over RF power irradiance arriving on points around the TV Tower, Place 3 (15°47'25.13"S/47°53'28. 88"O), in the parking area of the TV Tower, and Place 4 (15°47'39.61"S/47°53'11.72"O), 300 m away from Place 3 were chosen, as shown in Fig. 3.9.

Each one of the four places was visited three times during that day in order to register morning, afternoon and evening dBm values. The dBm power level took into account the corresponding frequency values, and therefore several measurements over different frequencies were made in order to feature the dBm behavior according to the part of RF spectrum. Fig.

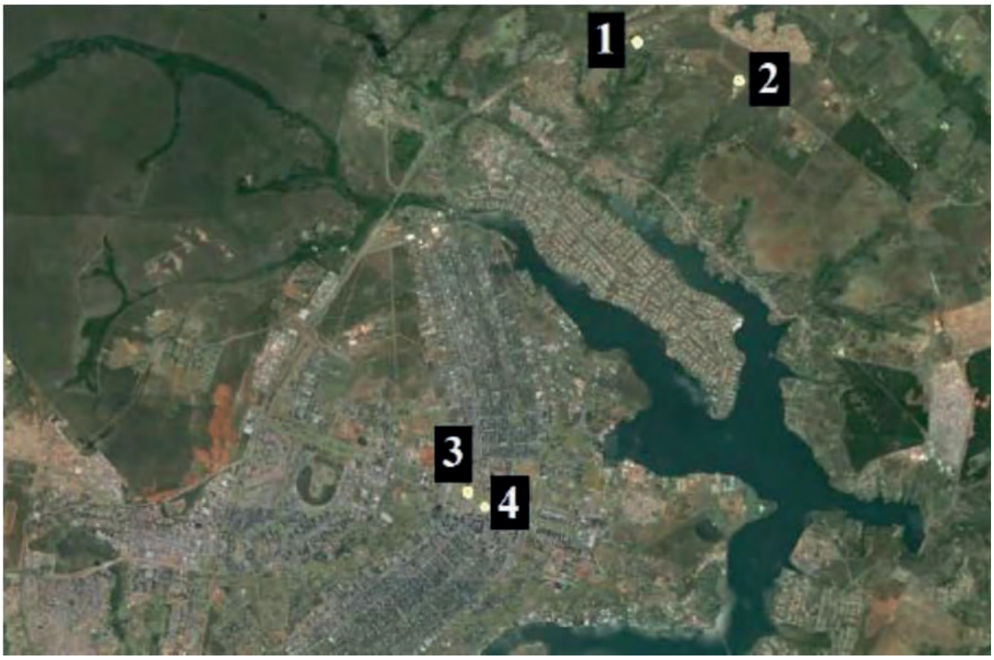


Figure 3.9: Map of Brasilia-DF (Brazil) with the indicated 4 places on which RF spectrum intensities were measured.

3.10 shows the average dBm obtained for the four investigated places. Note that, although Place 2 is nearer to the Digital Tower than Place 1, in the latter higher incident dBm values were identified. Note also that the most advantageous frequency values lie around 90 MHz. The dBm values of Places 3 and 4 achieve positive dBm values, while place 2 has negative dBm near 0 dBm.

3.4.2.2 Stage 2 – dBm of Place 3 with antenna array

Given the fact of the most promising dBm values were observed near spot 3, in this stage we have concentrated the survey on its parking area. Fig. 3.11 depicts the deployment of antennas on top of the vehicle used for measurements.

Measurements were obtained by parking the car in the vicinity of TV Tower while varying the number of antennas, which were fixed onto the roof of the vehicle by means of a strong magnets, enabling them to be installed or removed according to the convenience. The measurements comprised arrays with 1, 2, 3 and 4 antennas. Fig. 3.12 shows the results with the case of 4 antennas installed on top of the vehicle. In this situation, frequency 91.7 MHz registered 11.34 dBm and the channel whose frequency is 93.7 MHz demonstrated to offer 12.45 dBm. By converting into milliwatts, these values are 13.61 mW and 17.59 mW, respectively. The measurements of dBm values with 1, 2, 3 and 4 antennas, the attained results are shown in Table III.2.

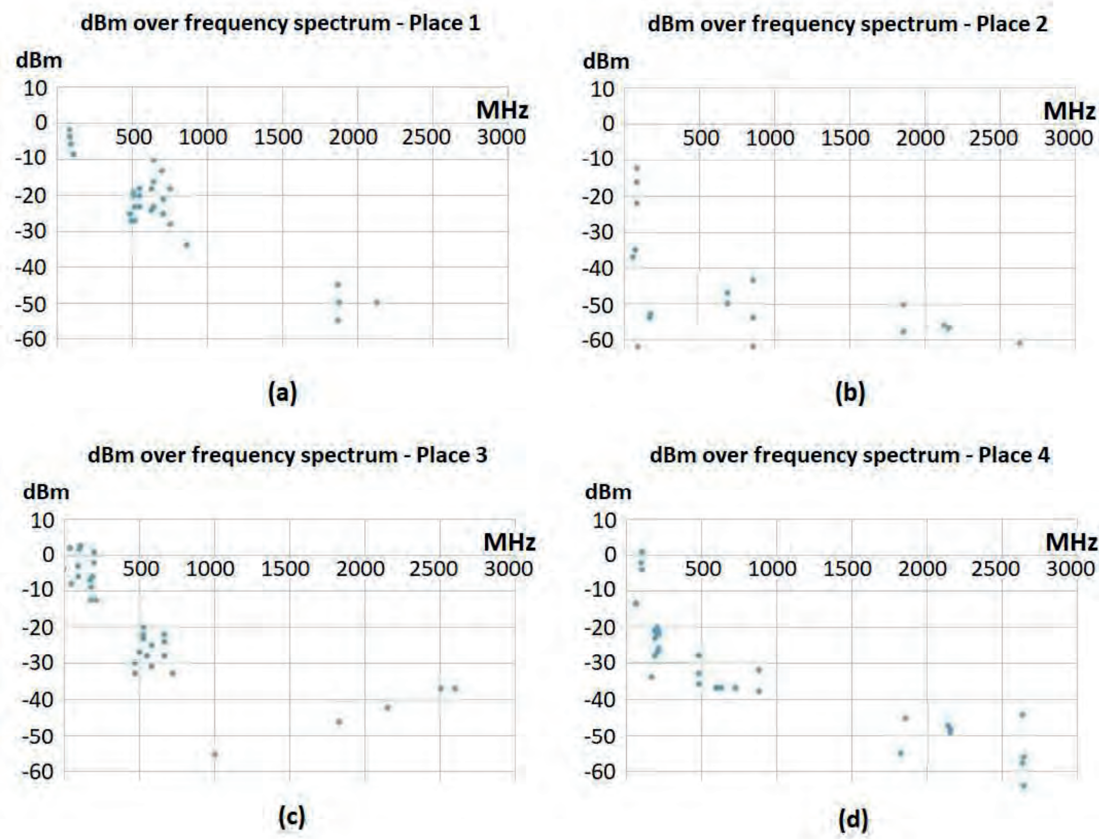


Figure 3.10: Incident dBm results for each one of the four investigated places

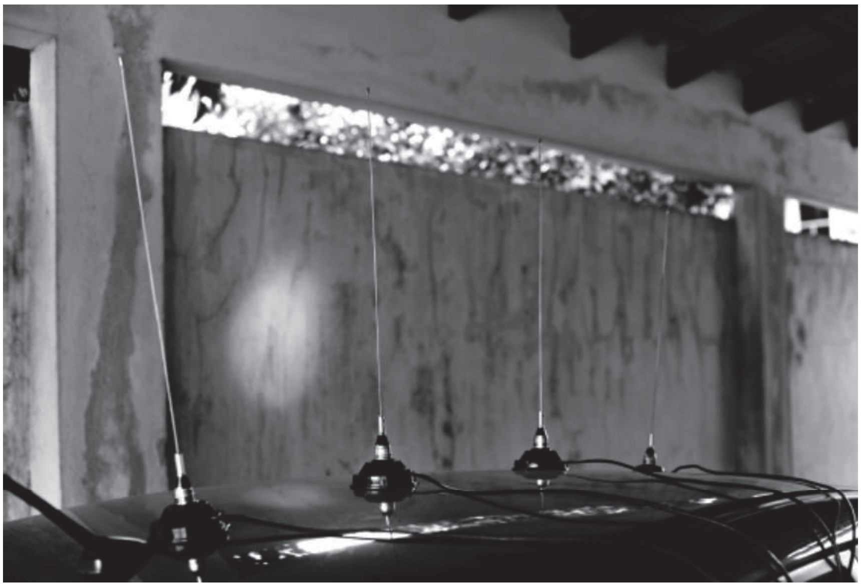


Figure 3.11: Roof of the car with four antennas, fixed in place by mean of magnets at the basis of the antenna

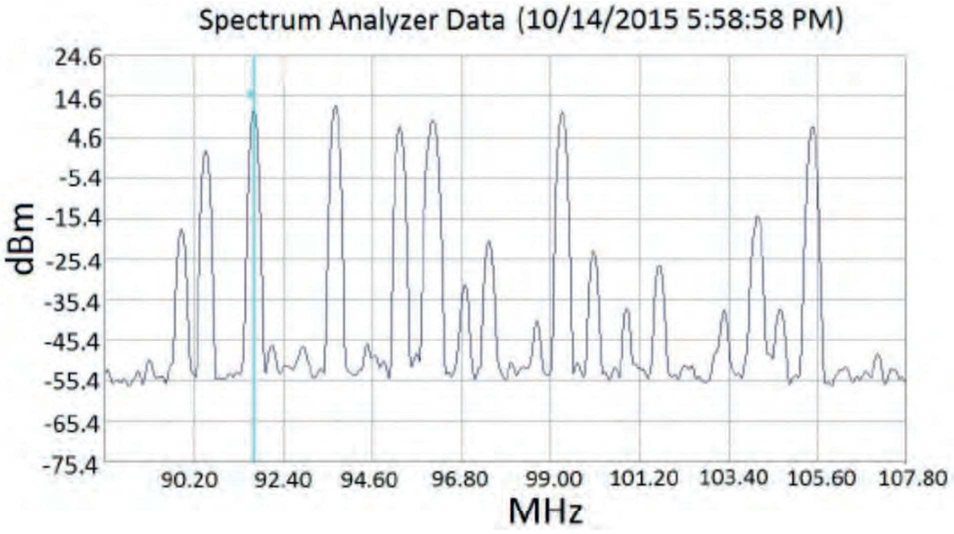


Figure 3.12: Measurements with 4 antennas on top of the vehicle

Table III.2: Values of incident dBm according to the quantity of employed antennas in the parking area of the TV Tower

Channel MHz	Incident dBm			
	1 antenna	2 antennas	3 antennas	4 antennas
91.7	-3.801	1.775	3.265	11.339
93.7	1.652	1.942	6.975	12.453
95.3	-7.894	-8.096	-1.267	7.231
96.0	-15.525	-13.083	-6.817	8.755
96.9	-40.190	-38.385	-34.293	-31.689
99.3	1.458	-0.005	3.199	10.979
104.4	-22.712	-30.298	-28.750	-14.649
105.5	-2.725	1.470	3.914	7.208

Increasing the number of antennas leads not always to improvement of the overall performance. Nonlinear gains arise from the destructive interference or mutual coupling among the antennas. The incident power hereafter adopted is 11 dBm, corresponding to 12.58 mW, value slightly below the best two measurements. Regarding the frequencies 91.7 MHz and 93.7 MHz, this result is considerably above those ones of [4].

3.4.3 Proposed Rectenna Array System for RF Energy Harvesting

Since there are permanent changes in beamforming directions of incident waves in practical situations, the set of the waves arriving on an antenna array immersed in an urban environment is expected to vary continuously, and adaptive beamforming systems should be provided to the harvester system. Any project of antennas array which handles a large number of units should pay heed to prevent problems related to destructive interference, given that it may occur depending on the Direction of Arrival (DOA) of the incident waves.

As shown in Fig. 3.13, when the incident DOA changes, the complex values of the filter weights a_i, $i \in \{1, 2, \dots, n\}$, must be automatically adjusted in order to provide a new optimal gain, maximizing the antennas output at any moment. However, the array of Fig. 3.13 needs an adjacent system monitoring the changes in the incident waves, implying additional power expenditure. Furthermore, the power loss related to commonly commercialized phase shifters varies from 0.5 dB to 14 dB, with average around 5 dB [112]. In the most favorable case, the loss of 0.5 dB admits maximum phase range of $\pi/3$ between each phase shifter. For the observed mean values at 5 dB, the power which is lost amounts up to 68.35 % of the incident one. Such levels of power losses in a system that works with few microwatts are prohibitive, enforcing the adoption of a simpler design that provides less power losses.

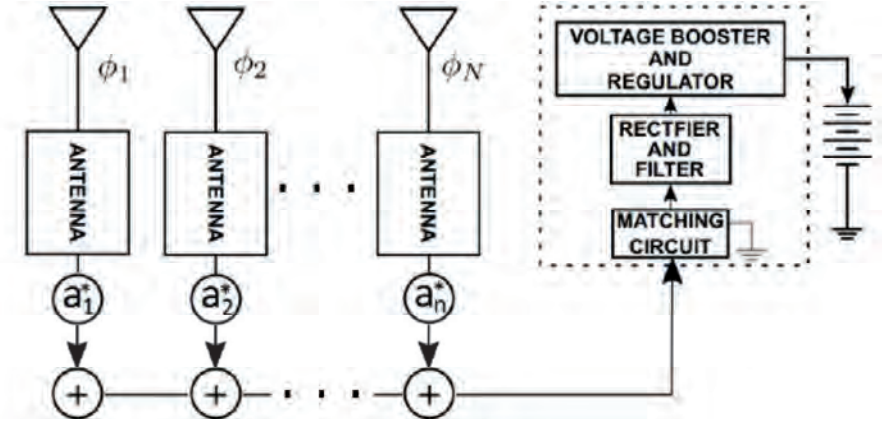


Figure 3.13: Array of antennas with phase shifters

By connecting all antennas in series, we consider that all phases of the incident waves are nearly equal, that is, $\Phi_1 = \Phi_2 = \cdots = \Phi_n$ without loss of generality. In this scenario, all antennas might have their outputs combined as in Fig. 3.14, where the set of receiver antennas is connected in parallel, sharing one single matching circuit. The RF signal is received by each one of the receiver antennas with respective phases $\Phi_1 = \Phi_2 = \cdots = \Phi_n$. This is how the antenna arrays of the following simulations are designed.

The problem with the assumption of $\Phi_1 = \Phi_2 = \cdots = \Phi_n$ is its infeasibility in realistic scenarios, as different DOA vary permanently. In order to avoid such drawbacks, a different scheme of a system with several stages is required. The first element is the RF source, which

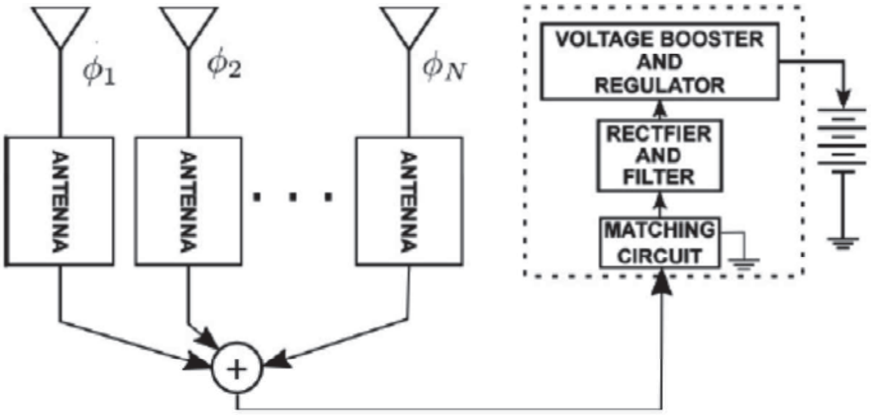


Figure 3.14: State-of-the-art block diagram based on antenna arrays considering a single matching circuit

is the TV and Radio broadcast transmitters in the cities. The next necessary element is our RF receiver with its *matching circuit*. While specifying it, one must pay heed to the selected band, since the management circuit is placed at the input stage to equalize the impedance between the antenna and the next component of the circuit. Next, the voltage booster and the rectifier are the items of our highest concern, since the circuit for energy management may be a point of power losses. A rectenna is an assemblage of the RF receiver, the matching circuit, the voltage booster and the rectifier. Fig. 3.15 shows the rectenna components inside a dashed box.

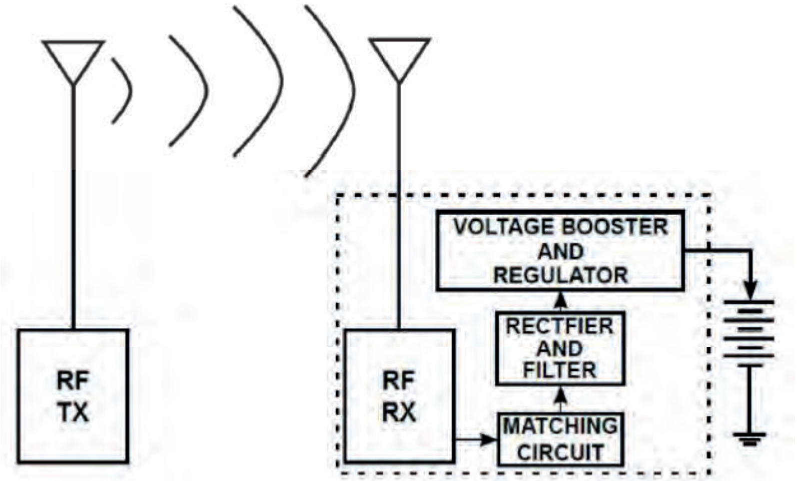


Figure 3.15: Components of a RF energy harvesting system: the rectenna is an antenna with a RF-DC interface [96]. Therefore, the receiver must be integrated to a matching circuit, a voltage booster and the rectifier, whose output is often connected to a battery.

Each rectenna of Fig. 3.16 is therefore an assemblage of the same elements shown in Fig. 3.15, in which the battery is not part of the rectenna, being rather an external element.

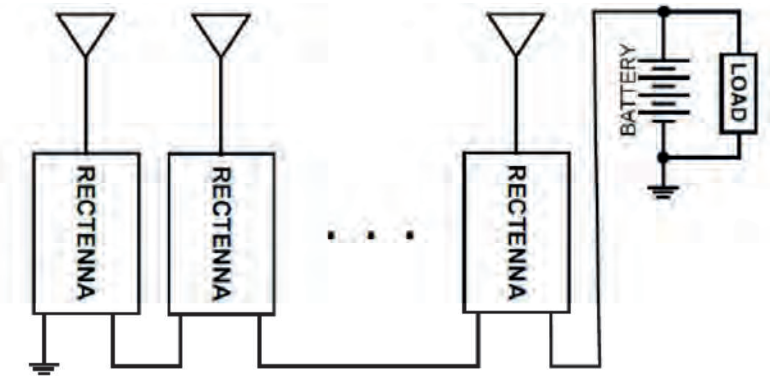


Figure 3.16: Proposed rectenna array system for high efficiency RF energy recycling, with rectennas connected in series

Several rectennas might be installed in series, and the reason for that is the expected low voltage brought for each unit, even with the aid of the booster inside them. Hence, one can provide a not so low voltage for the external load, while preventing variable DOA from influencing the overall power which is drawn from the incident waves.

3.4.4 Simulation of Rectennas and Antennas Performances

We perform the comparison between antennas and rectennas in terms of the net amount of recycled energy by means of simulations. Antennas and rectennas are hereafter named as receiver terminals.

A set of assumptions are made in order to feature the best possible scenario for antennas connected in series. One single hypothetical transmitter antenna emits its signals only with the frequency ω. This hypothetical transmitter is the source of the impinging frequency waveforms of our experiment. We also consider that there are n receiver terminals, projected to the same frequency ω. All receiver terminals are equally distant from the single transmit antenna, receiving line-of-sight (LOS) waves with amplitude A_t. Likewise, we assume only one reflective object introducing diffuse pathways to the transmitted signals. The reflected waves here generated perform the non-light-of-sight (NLOS) component and have random phases under a Gaussian distribution. Furthermore, all reflected waves achieve every receiver terminal with the same amplitude A.

All these conditions are established to assign relative advantages in recycling incident power to the antenna array. The goal is to prove that, even relying on them, antennas connected in series might have problems with shift in phases due to NLOS components. The insertion of a reflective object is adopted as a way of assigning a minimum realism to the model. The end-to-end efficiency is not considered as a decision factor since the need for the matching circuit is a common factor for antennas and rectennas arrays.

The assumption of one single transmitter antenna is made with all receivers being located on points that are supposed to be struck by the incoming waves at the same phase, considering the incident waveform. Fig. 3.17 shows the physical idea based on the set of points in space bearing the same phase given a signal of wavelength λ. In Fig. 3.17(a), several candidate points are given by the highlighted spots, which are deployed along lines of equal amplitudes of the wave F1. In Fig. 3.17(b), given the existence of two waves F1 and F2, the set of spots of equal phases offer fewer options than in Fig. 3.17(a). Finally, in Fig. 3.17(c), with three incident waves F1, F2 and F3, the possibilities remain even scarcer.

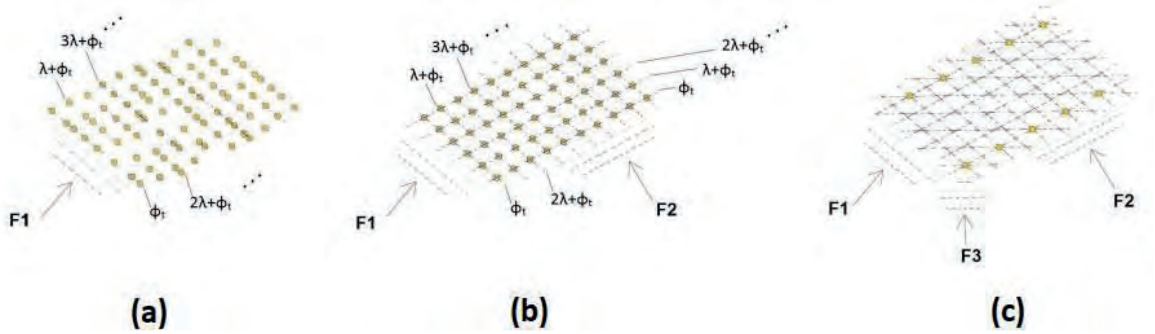


Figure 3.17: Spatial interpretation for same phase regions in which (a) one single wave F1 offers a plenty of points, (b) two incident waves F1 and F2 generate a more difficult environment for achieving such points, and (c) the number of spots is even scarcer with three crossing waves.

The above described single fictional transmitter is adopted in the further analysis, theoretically replacing the set of all actual transmitters that would exist in real urban environments and providing the set of incidences of Fig. 3.17(c). In the same way, all receiver terminals are supposed to lie on equal phase spots considering this fictional transmitter. The additional waveforms F2 and F3 of Fig. 3.17(c) stand for reflections of the original waveform. The idea that Fig. 3.17 conveys is that our experiment theoretically locates the receiver antennas on the spots of Fig. 3.17(c) as likely as possible.

In Fig. 3.18, receivers RX1 and RX2 are equally distant from the transmitter TX, and solely LOS signals are present. NLOS transmission shows up in the model in consequence of the appearance of reflective object, for instance, a truck. Diffuse pathways are therefore created. Along with the additional trajectories, the incident waves are supposed to arrive on the receiver terminals with random phases and the same amplitude A, for the sake of simplicity. The side view is depicted in Fig. 3.18(a) and the upper view is in Fig. 3.18(b).

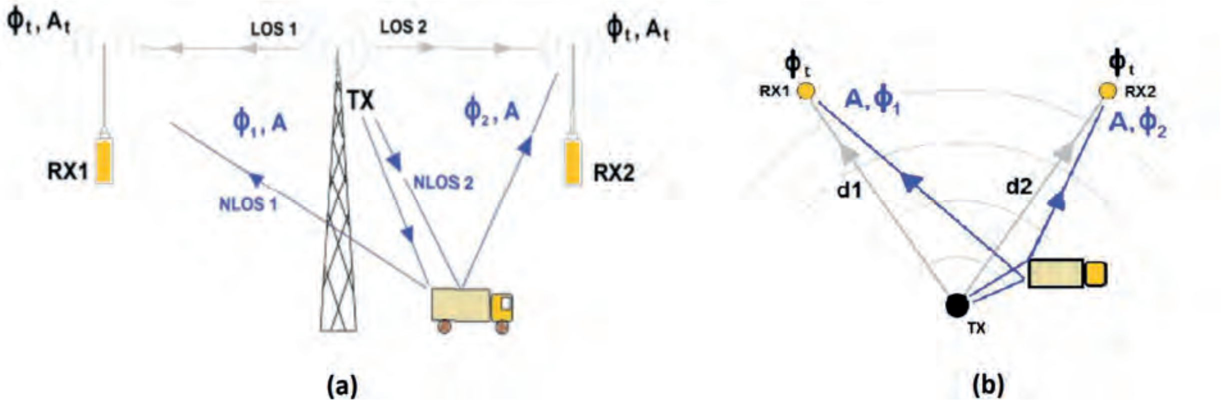


Figure 3.18: Reflective object, for instance, a truck, reflecting waves over NLOS pathways. Incident reflected waves strike the antennas with amplitude A and different phases Φ_1 and Φ_2. In (a) we have the side view and (b) shows the upper view.

A generalization of the scenario of Fig. 3.18 involves several receiver terminals RX1, RX2, ..., RXn, as depicted in Fig. 3.19. All phases of the incident waves that arrive on each receiver are considered random under a Gaussian distribution.

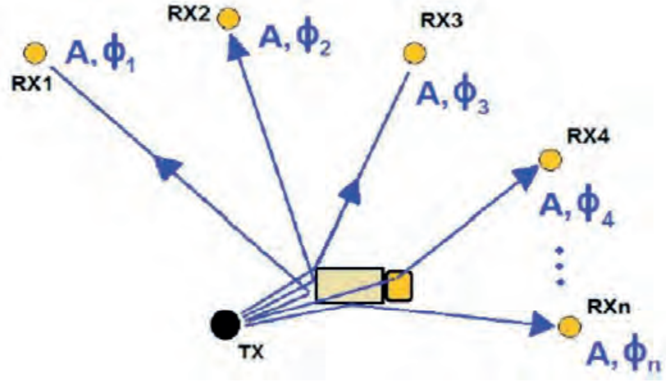


Figure 3.19: Several NLOS waves achieve receiver terminals, as a generalization of Fig. 3.18.

Defining the total power irradiated by the transmitter antenna as a sum of A_{LOS} and A_{NLOS}, A_{LOS} is the amount of RF power that is irradiated by the transmitter and is assigned to LOS pathways, whereas A_{NLOS} is the amount of power coming out the transmitter TX that ends up irradiating through NLOS pathways. According to our model, LOS power A_{LOS} is set to reflect directly the value of the component A_t that effectively reaches the receiver through LOS. The coefficient K_t, the efficiency factor concerning LOS pathways, assumes four defined values: 0, 0.3, 0.6 and 1. The power value A_{NLOS} multiplies the coefficient K, the efficiency factor concerning NLOS pathways, which varies under a Gaussian distribution, to produce the NLOS irradiated power A that achieves the receiver.

$$A_t = K_t A_{LOS} \tag{3.20}$$

$$A = K A_{NLOS} \tag{3.21}$$

For instance, $K = 0.3$ implies that on average the amplitude of the signal wave coming from NLOS pathways is attenuated in 70 % between transmitter and receiver. Alternatively, whether $K_t = 1$, the LOS signal achieves the receiver with zero losses, whereas $K_t = 0$ means that there is a complete shadowing over the LOS pathway. It is occasionally allowed $K > K_t$ in order to investigate the real impact of NLOS components over the performance of an array of antennas, connected as in Fig. 3.14.

In our simulations, each antenna or rectenna is struck by two incident waves, the LOS and NLOS components, over 1500 realizations. The LOS component provides incident waves with the same phase Φ_t on every receiver, whereas the NLOS component has phases Φ_n randomly varying over the n-th receiver, and therefore $\Phi_1 \neq \Phi_2 \neq \cdots \neq \Phi_n$. The randomly distributed phases caused by NLOS incident waves are a key factor for our comparison, since the probability distribution function of phases $\Phi_1, \Phi_2, \ldots, \Phi_n$ is Gaussian. Because the instant power value at each antenna or rectenna is the sum of the LOS and NLOS parcels, the instant power at each receiver device has a Gaussian distribution as well.

3.5 Results

In Section 3.5.1, simulations and results from PCA Kalman for load forecast are shown. A brief analysis of the energetic sustainability of sensors is drawn in Section 3.5.2. In Section 3.5.3 we present the results showing that rectenna based energy harvesting outperform the antenna based systems in terms of power harvesting under variable DOA.

3.5.1 PCA Kalman for Load Forecast - Simulations

In order to validate the proposed PCA-Kalman load forecasting system (PKF) performance, the load time series have been forecast by concurrent methods of linear and nonlinear natures. A classical Kalman Filter (KF) without PCA and variance estimation represent the linear approaches, while a classical BP double layer Artificial Neural Network (BP) and a PCA enhanced BP ANN (PBP) are employed to showcase the performance of these nonlinear methods.

The above described benchmark models are used to forecast base, average and peak demand. For each prediction the Mean Squared Error (MSE), Mean Average Percentual Error (MAPE) and Maximum Percentual Error (MPE) error metrics are calculated. Eight input sets are tested, each designated by a capital letter. Sets A and B are described in Section 3.2.1, set C in 3.3.1, set E as explained in Section 3.2.2, set F in 3.2.3 and set G is described in Section 3.2.4. Set D is formed by the union of B and C, while set H is the union of all the previous input sets.

The forecasting period starts at 29 October 2001 and comprises 914 days. Tables III.3, III.4 and III.5, respectively, summarize results for base, average and peak load forescasting.

Table III.3: Error metrics for Base load

Metric	Method	A	B	C	D	E	F	G	H
MSE	PKF	162,8	91,4	114,9	76,4	234,3	189,3	92,1	37,5
	KF	155,1	92,3	137,3	88,1	251,4	218,7	92,1	80,2
	PBP	177,8	336,4	517,0	474,0	130,6	110,5	321,4	395,7
	BP	137,4	348,6	353,4	500,4	165,3	94,3	252,6	468,4
MAPE	PKF	3,05	2,33	2,63	2,12	3,47	3,29	2,26	1,40
	KF	2,99	2,30	2,86	2,24	3,53	3,07	2,26	2,10
	PBP	3,32	4,49	5,64	5,14	2,90	2,69	4,23	4,79
	BP	2,92	4,57	4,47	5,55	3,31	2,39	3,65	5,16
MPE	PKF	20,8	11,8	13,1	11,1	24,6	19,6	15,1	9,8
	KF	20,3	14,5	16,6	14,3	30,4	35,9	15,1	11,8
	PBP	19,3	18,4	21,6	37,4	13,2	13,2	18,9	23,4
	BP	11,8	18,8	20,7	23,0	13,0	11,4	17,7	26,6

Note that all input sets provide reasonable forecasting performance. For the state space approaches, set C slightly outperforms input set A, as D also outperforms B, giving evidence that the performed preprocessing is beneficial to linear predicting algorithms. The ANN methods, however, are negatively affected. Input set F works well with the neural networks. Input set H combined with the load forecasting system provide the best performance.

Overall, the prediction of average load displays the largest error metrics, probably due to the larger quantity of outliers in this particular time series. The only exception is the PCA-Kalman system, as it shows smaller relative errors at the cost of increased maximum error, as compared with the base load prediction problem. ANN do not seem to perform well in this scenario, displaying large error metrics.

The proposed PCA-Kalman based approach vastly outperforms the other methods for peak load prediction. The KF achieves a MSE almost three times larger, yet forecasting with good accuracy. ANN methods produce better results when employing input set F.

Overall, the proposed system displays good forecasting performance, being capable of

Table III.4: Error metrics for Average load

Metric	Method	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*
MSE	PKF	579,0	339,2	454,4	302,2	1403	631,6	204,2	83,7
	KF	554,2	343,4	484,8	338,7	851,8	640,2	204,2	263,2
	PBP	1247	1780	1306	1999	1544	1152	1945	2389
	BP	1567	2597	1685	1988	1170	1220	1648	2052
MAPE	PKF	3,62	2,90	3,26	2,59	4,71	3,87	2,14	1,32
	KF	3,51	2,89	3,37	2,81	4,16	3,57	2,14	2,46
	PBP	5,79	7,48	6,23	7,72	6,71	5,91	7,33	8,11
	BP	6,69	9,13	7,09	7,54	6,06	6,02	6,75	7,85
MPE	PKF	40,8	18,3	23,4	22,8	54,9	37,0	22,1	12,4
	KF	40,4	18,4	33,6	19,3	39,9	37,0	22,1	19,6
	PBP	28,2	28,8	21,3	33,7	35,2	29,9	30,0	40,1
	BP	36,1	26,4	31,0	40,5	30,9	30,8	25,2	32,1

Table III.5: Error metrics for Peak load

Metric	Method	*A*	*B*	*C*	*D*	*E*	*F*	*G*	*H*
MSE	PKF	501,8	294,2	389,2	263,3	1046,3	529,3	189,9	71,4
	KF	491,0	289,5	413,4	275,4	786,4	494,4	189,9	212,2
	PBP	646,0	1365	1079	1476	813,3	550,0	1627	1632
	BP	666,1	1988	808,2	1630	587,2	561,5	1205	1388
MAPE	PKF	2,63	2,06	2,30	1,88	3,11	2,72	1,69	1,01
	KF	2,60	2,06	2,38	1,98	2,98	2,43	1,69	1,70
	PBP	3,22	4,68	4,20	4,75	3,50	2,99	5,20	4,99
	BP	3,15	5,74	3,63	5,15	3,06	2,79	4,13	4,87
MPE	PKF	27,0	15,4	18,3	14,2	39,4	28,0	8,8	6,4
	KF	27,0	14,8	22,5	14,3	31,5	28,0	8,8	14,5
	PBP	22,4	25,5	16,5	23,5	19,5	15,2	21,6	24,8
	BP	24,5	23,0	17,0	26,7	15,2	25,2	25,6	30,5

daily predictions with MAPE lower than 2 % in all scenarios. In comparison, the linear and nonlinear benchmark predictors could only achieve MAPE lower than 2.5%, at best.

3.5.2 Sensors – Energetic Sustainability

Most commonly used sensors consume dozens microwatts in sleep mode and hundreds microwatts in active mode [100]. We can assume without loss of generality that a typical sensor consumes around 0.1 mW for steady operation. Considering the three best dBm values of each place in stage 1 of measurement campaign (Section 3.4.2.1) and taking their average dBm values, Table III.6 shows the number of points with the same dBm incidence that are needed to feed one sensor. The average efficiency of 18 % is achieved in [101]. Note that the incident dBm here are considered as with one single antenna, and that place 2 was not considered due to the low dBm values. The results show the feasibility of exploiting RF power harvesting for feeding sensors in the assessed urban environment.

Table III.6: Average dBm values obtained in stage 1 of measurement campaign and number of points for feeding a 0.1 mW sensor

Place	1	3	4
Average dBm	-3.2	1.7	-1.9
Power (mW) with efficiency of 18 %	0.086	0.266	0.116
Number of points needed for a 0.1 mW sensor	2	1	1

With regards to stage 2 (Section 3.4.2.1), crossing the average efficiency of 18 % with an incident power of 11 dBm, we achieve the average power of 2.27 mW. The array of 4 antennas occupy circa 0.9 m^2 in space, considering the high of each antenna of 60 cm and the distance of approximately 50 cm between two units. Therefore, power per area achieves 2.5 mW/m^2. Note that this result is in accordance with the final power density between 0.1 and 3.0 mW/m^2 of [103].

3.5.3 Energy Harvesting with Antennas and Rectennas – Simulations

In this Section, the simulations aim solely at evaluating the superiority of rectennas over antennas when dealing with impinging RF waveforms whose phase of arrival varies randomly.

Specifying the ratio between the power drawn by a rectenna array P_{rec} and the power provided by an antenna array P_{ant} as η_A, we have

$$\eta_A = \frac{P_{rec}}{P_{ant}}. \tag{3.22}$$

The variable η_A is therefore a measurement of the superiority of rectenna array power in relation to that delivered by the respective set of antennas. Whether $\eta_A < 1$, the antenna array is advantageous in detriment of the rectennas.

We test η_A profile for the arbitrary values $K_t \in \{0, 0.3, 0.6, 1\}$ of A_{LOS}. In sequel, we vary the number of receiver terminals from 2 until 50 units in simulations. Fig. 3.20 comprises results of η_A regarding the number of receiver terminals and the LOS incident power A_t according to K values.

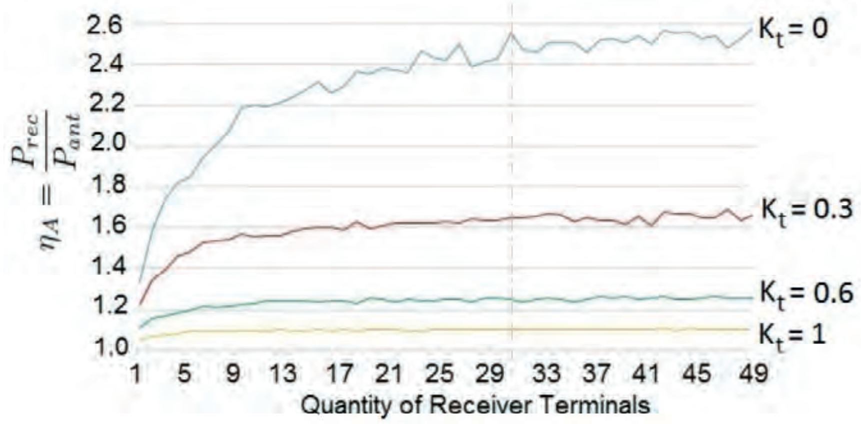


Figure 3.20: Profile of η_A as a function of the LOS incident power $A_t = K_t A_{LOS}$ and the number of receiver terminals

For instance, in the absence of LOS waves and existing around 11 receivers, the power recycled by the array of rectennas is over twice the power delivered by antennas. Considering the situation of purely NLOS transmission, i. e., $K_t = 0$, we vary the quantity of the existent reflective objects in four cases, with 2, 5, 10 and 15 reflective objects. We analyze each of these cases while varying the number of receiver antennas from 2 until 50 as previously. The values for η_A behaved as illustrated in Fig. 3.21. All curves follow approximately an unique, defined logarithmic pattern. However, smoother lines are achieved whether the number of reflective objects rise.

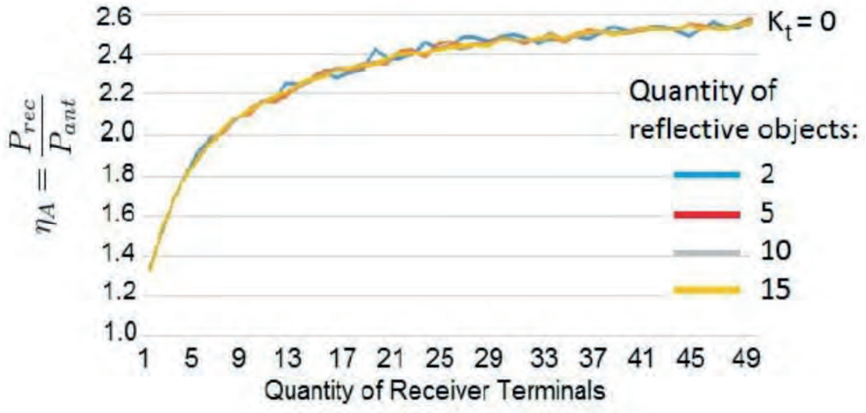


Figure 3.21: Profile of η_A as a function of the quantity of reflective objects only with NLOS incident power, i. e. $K_t = 0$

Regarding all cases of tested η_A values and according to the different number of receivers, arrays of rectennas overcome the ones assembled by antennas in series. Moreover, given a certain quantity of receivers, the lower the average amplitude of LOS signal, the higher is the advantage on using rectennas, as observed in Fig. 3.20.

A major drawback involving antenna arrays is the lack of available spots on which to install them whether the goal is to obtain all incident waves with the same phase. On one hand, higher model orders contribute with larger available wireless power given the higher number of incident waves; on the other hand, encountering the intersections of sinusoidal maximum values becomes increasingly difficult concerning the spatial criterion.

3.6 Conclusions

This chapter proposes a short term load forecasting system. The envisaged method has been validated by forecasting daily values of base, average and peak load in Brasília, Brazil. Linear and nonlinear transformations, as well as data mining techniques were employed to produce optimized input sets.

The candidate socioeconomical variables and the tariff time series are shown to be capable of improving the performance of the evaluated methods, remarking the good performance given by the tariff input set to the ANNs. The nonlinear transformations performed over weather variables are effective at improving the forecasting performance of the State Space predictors. Having built-in nonlinear capabilities, the ANN do not benefit from these transformations, nevertheless.

With regards to the PCA Kalman prediction system, further research must focus on expanding the sets of candidate variables and investigate the possibility to develop universal types of nonlinear transformations, applicable to the full set of candidate variables. The effect of rapidly growing distributed generation, grid storage and demand response over the performance of this load forecasting system might also be a topic for future work.

In parallel, we have evaluated the feasibility of contribution from RF energy harvesting systems by measuring their energy in dBm over four points into Brasília, Brazil. Counting on the global percentage efficiency of 18 %, we have have achieved that, over places 1, 3 and 4, as few as 4 points are required to feeding sensors with energy for steady operation. Moreover, we have proposed a rectenna array device. By means of simulations, we have shown that rectennas can significantly increase the efficiency of RF harversting systems. Our results allow us to conclude that RF harvesting systems are feasible to provide power to sensors that are located in areas of difficult access to people and where conventional power systems are not available.

IV

DATA SECURITY AND TRADING SYSTEM FOR SMART GRID WITH PERFECT SECRET SHARING

4.1 Introduction

The power grid is a crucial large scale infrastructure. In order to allow high level of automation, information security, distributed energy control and robust load fluctuation management of the power grids, smart grids (SGs) are essential. Several interdisciplinary aspects are involved in SG, such as interoperability, information security, scalability, reliability, energy efficiency, re-usability, communication backbone, electrical actuators, sensor and control technologies [113]. Although the SGs are constantly improved, they are still vulnerable to cyber attacks. Hence, current power grids should be further improved to fit the demands regarding to data security [26] and energy trade between prosumers [27] [28].

Historically in the electricity market, the flow of electrical power as well as the corresponding consumption measurements and prices have been imposed in a vertical frame, from the companies/producers to the consumers. Nowadays, power, information and prices flow bi-directionally. Ubiquitously produced data in end-points flow upwards in the grid, reporting to the utility company about all sort of actions [11] [12]. Power is generated inside the boundaries of final user real states — integrating the micro scale power generation — and exported to the company or other consumers [13] [14]. As an illustration, the total worldwide PV panels installations have reached 300 GW by 2016, of which about 28 % are decentralized grid connected worldwide [15]. From those, a great fraction consist of captive clients of the power utility company, which until recently were not allowed to export power. The capability of these clients to trade energy is a matter of increasing interest in the state-of-the-art SG. As a reflex of these phenomena, prices are undergoing a process of decentralization [16] [17], with the possibility of cooperative or non-cooperative frameworks, including auctions [18].

To the best of our knowledge, there are open issues in the literature with respect to secrecy of prices and identities of prosumers in energy trading systems. For instance, the main efforts towards privacy consists of obfuscating the instantaneous consumption pattern of each consumer [29] [30]. This is generally accomplished by hiding the instantaneous power consumption of the client as fine-grained consumption data can reveal in details the life style of the consumer [114] [115]. However, the profile of traded energy also delivers relevant information about the prosumers to his neighbors. As in [116], the ability to link the bids to individual consumers allows for an untrusted entity to build up a profile of the consumer's behavior. In particular, the quantities of traded energy can be very informative about the economical welfare of the owner [117]. Privacy requirements dictate that prosumers cannot gain information regarding other prosumers' consumption and production — not even if they are trade partners [118]. Models dealing with energy trade directly among prosumers [27] [31] limit themselves to exploit the trade environment without discussing in details data security aspects related to the identities of the traders in relation to their neighbors. As a consequence, several topics related to privacy requirements are still open in SGs, such as power production and bidding in trading systems.

In a broader aspect, protecting sensitive information in cyber-physical systems such as SG is currently an increasingly difficult task. According to [119], large quantities of data are collected from various applications in SG, such as smart metering substation state monitoring, electric energy data acquisition and smart home with Smart Meteres (SM), all of them gathering sensitive data. Various attacks aim at imposing harm against secure and stable SG operation [120]. Inviolability of consumption data is at the center of discussions in the realm of SG data secrecy protection. For instance, in [115], a study about the impact of data granularity on edge detection methods, which are the common first step in non-intrusive load monitoring algorithms, shows that devices whose consumption is above 50 W can be detected. Moreover, data protection is specifically difficult due to the low capacity of the SM in terms of data aggregation and data handling [121].

Among the most common threats to data confidentiality, brute-force attacks (BFA) occur when an attacker tests sequentially all possible values of a protected information until the correct one is discovered. BFA has shown to be one of the major threats to network security despite the computational burden on the attacker side [122]. A machine compromised by a brute force attack can cause serious damages such as distributing sensitive information and participating in distributed attacks [123]. Protection against BFA includes enhancing the set of possible values of the protected information [124] or limiting the number of queries [122]. The latter is circumvented by using massive botnets, each bot querying potential passwords. On the other hand, when the secret range increases, BFA becomes a harder task.

One of the most important techniques for protecting a sensitive information is the Secret Share Scheme (SSS) [38] [39]. An SSS protects a secret $S \in \mathbb{Z}$ by dividing it into shares that are distributed by an honest dealer to n participants. Only when a coalition of at least t participants occurs, the secret is revealed to them. In these terms, the SSS is said to be a threshold scheme of the type (n, t) [39]. An SSS-(n, t) is said perfect if, when $(t - 1)$ or less participants combine their shares, it is not possible to extract any information about the secret [125]. The state-of-the-art SSS in the literature are the Shamir SSS [126] and the Asmuth-Bloom's SSS [127], the latter based on the CRT. The Asmuth-Bloom's SSS improves the Mignotte SSS [128], which is not perfect. Note that the computational complexity of the secret retrieval from t shares in the Asmuth-Bloom's SSS behaves as $\mathcal{O}(t + 1)$, while in the Shamir's SSS it behaves as $\mathcal{O}((t + 1)log^2(t + 1))$ [129].

However, although the Asmuth-Bloom's system offers a perfect SSS, the problem with this approach is that the range of the secret values is very limited, i.e., the size in bits of the shares is usually greater than the size of the secret itself. As a consequence, increasing the range of secret values contributes to improve CRT based SSS as a data security technique.

In this chapter, we consider the problem of providing data privacy for self-interested players that trade energy in the context of a Neighborhood Area Network (NAN). The energy is sold by local prosumers and purchased by their neighbors in a competitive market, with the support of a Trusted Third Party (TTP), a central operator that is reputed inviolable. Our framework deals simultaneously with SG data security requirements and energy trade systems as follows. As a first contribution, the proposed framework has a privacy-preserving model which has a low computational complexity and avoids completely an unauthorized party to identify the bidders, the number or types of them, and even if the bids achieve or not a deal. As a second contribution, all the bids are made clear to the authorized NAN participants, with all SM owners having access to how many bids are proposed, their types, prices and quantities. Nevertheless, the proposed framework avoids totally any access to the bidders identities. As a third contribution, we present a perfect CRT based SSS that has a better ratio between the sizes of the shares and the secret than in the Asmuth-Bloom's SSS. The secret is an integer-valued number $S \in \{0, 1, \ldots, (k - 1)!\}$ that is associated with each permutation of k elements of a vector $\mathbf{s} \in \mathbb{Z}^k$ obtained from the shares, which are sparse matrices. We prove mathematically that our system is not only is perfect but also presents a potentially unlimited gain in terms of secret range. Considering for instance a set of the highest 6 co-prime numbers under 100, the gain in secret range per bits used is leveraged in the order of 10^{103} in comparison with the Asmuth-Bloom's SSS. We use the proposed CRT based perfect SSS in the SG trading system as a confirmation that the agents effectively know one of the keys used in XOR operations.

The remainder of this chapter is organized as follows. Section 4.2 surveys the state of the art in terms of SG security requirements and energy trading schemes for electricity markets, along with the revision of the Asmuth-Bloom's CRT based SSS. In Section 4.3, we use the technical requirements observed in the literature to present a detailed problem description and describe the proposed framework for energy trade in a NAN, showing how the proposed SSS provides a data integrity survey. Section 4.4 shows the results in terms of data secrecy, trading system features and gain in secret sizes for the proposed SSS. Section 4.5 concludes the chapter.

4.2 SG Data Security – State of the Art

In this section, we provide an overview of the literature in terms of data security, SG trading systems and the CRT for cryptographic applications. In Subsection 4.2.1, the state of the art for SG in terms of data security and privacy requirements is exploited. In Subsections 4.2.2 and 4.2.3, data security towards privacy and cryptographic systems applied to Smart Grids are revised, respectively. Further, in Subsection 4.2.4 we present basic characteristics of trading systems for connected prosumers. The Asmuth-Bloom's SSS has a review in Subsection 4.2.5, and a tensorial approach for CRT systems is revisited in Subsection 4.2.6.

4.2.1 Data Security and Privacy in the Smart Grid

According to [11], availability is defined as the capability of information to be accessible reliably and on time. Integrity consists of the protection of data against unauthorized modification or destruction, while confidentiality means preventing unauthorized disclosure of information that is not open to the public and individuals. Authentication is based on recognizing and validating the true identity of the communicating parties. Authorization consists of the permission of access to a system, also known as access control, and non-repudiation assures that a certain action performed or message sent cannot be later denied by their author. Although availability is important to provide network access for end users, data integrity and confidentiality are more critical in the Advanced Metering Structure (AMI) network near the final consumers. For example, stringent timing requirements are typically related to distribution and transmission networks, whereas in low voltage the message latency for meter reading may vary from minutes to hours [11] [130].

IV DATA SECURITY AND TRADING SYSTEM FOR SMART GRID WITH PERFECT SECRET SHARING

Inviolability of consumption data is at the center of discussions in the realm of SG data secrecy protection. For instance, data from off-the-shelf SM are sufficient to identify the TV movies viewed [131] due to unique fluctuations in the brightness of the movies influence the energy consumption of the TV set. In [115], a study about the impact of data granularity on edge detection methods, which are the common first step in non-intrusive load monitoring algorithms, shows that devices whose consumption is above 50 W can be detected. Moreover, data protection is specifically difficult due to the low capacity of the SM in terms of data aggregation and data handling [121].

Masking the identity of each user is the dominant strategy in order to provide user privacy, which is achieved by means of the assignment of false IP data to each SM [132] [118]. This technique is however sensitive to de-anonymization, which consists of the re-identification of nodes' identities behind their false IP. According to the probabilistic frameworks of [12], reported consumed energy at a 10 kWh scale can reduce the percent of re-identified SMs to between 10 % and 30 %. One should note that it may not be applicable in regions where the law requires that energy reporting should be done with kWh accuracy. In [131], 68 % of all consumption data can be re-identified as they have found unique combinations of feature values in the energy-consumption data of 122 households. Adopting false IP has also a weakness related to updating pseudonyms, which is frequently required for false IP nodes. These updates include revocation of current pseudonym and registration of the new one, such that, in order to avoid linkage of the two pseudonyms, after revocating the old one, the customer waits a certain period before registering the new pseudonym [132]. This time interval can be used by a malicious observer for leveraging their re-identification capability. Another problem related to identity protection is the impersonation attack, in which the adversary manages to intercept previous login messages from the user that it intends to impersonate [133]. By succeeding to login with the credentials of the attacked node, the adversary assumes its identity. Common solutions for the impersonation attacks include mutable field of messages, as by means of hash functions [134].

Although identity can be masked, all devices must know who they are communicating with, and who they are supposed to communicate with [135]. Hence, even when the identity of the SM is not disclosed, a list of the IP numbers should be made available to each node, since a basic principle is that consumers have the right to know where their information is being shared [136]. There are cases in which direct interactions between final nodes are expected, for instance in Secret Sharing techniques, as used in [30] [137]. In the case of Demand-Response programs, the Open Automated Demand Response (OpenADR), which is used as a tool for directly controlling load in order to manage power demand, is proposed [138]. However, OpenADR does not permit peer-to-peer communication between Virtual End-Nodes (VENs), so that most of protocols that rely on direct communication between end-users such as in [14, 27, 29, 113, 121], [139]- [140] remain uncovered by this technology.

The use of Internet Protocol (IP) and commercial off-the-shelf hardware and software is one of the most serious vulnerabilities of SG [141]. The Internet as part of the Wide Area Network (WAN) is considered undesirable [136] – such integration entails cyber threats since the SG is based on ethernet, TCP/IP and other operating systems, thus making the grid more susceptible to attacks. In [30], a privacy preserving power usage protocols which allow computations to be completely outsourced to cloud servers is presented. However, since the power usage data are sensitive information, they split the sensitive data through different clouds by means of a Secret Sharing technique. Another examples are in [11] [141], which claim that the SG imposes much more strict security requirements than the Internet in order to fully achieve efficient and secure information delivery for critical power infrastructures. Hence, in this chapter, we assume that the Internet and off-the-shelf protocols such as TCP/IP are not to be integrated to the SG trading infrastructure. Therefore, with regards to the communication links in the NAN, the systems in use for data transmission are supposed fully dedicated to the SG environment and independent of the Internet protocols.

As an additional guide to achieving a secure data exchange protocol, [142] provides a List of Requirements that aids to characterize to whom any information should be available in SG, as well as the potential of harm that data leakage can cause, as shown in Table IV.1.

Table IV.1: Security Requirements for SG Data in [142]

Requirement	**What Each End-User Must Be Able To**
Information Awareness	To be aware of the information that can be extracted from various data sets
Data Control	To control which data sets are released to which stakeholder
Data-Information link	To be aware of which information can be extracted from a specific data set
Data Aggregation	To be aware which information can be extracted from various combinations of specific data sets
Risk Awareness	To be aware which risk could arise by misuse of a specific information
Situation Dependency	To be able to release my data situation specific to dedicated stakeholders
Data Overview	To have an overview about which data is released to which stakeholder
Data Minimization	To release only as little data as necessary
Release Expiry	To have the expiry of specific data releases properly handled

4.2.2 Cryptographic Solutions for the SG

Cryptography is a central aspect in SG data security. The several devices that embed cryptographic applications execute their routines using symmetric or asymmetric keys. Symmetric keys use the same key to encrypt and decrypt the message, while asymmetric keys use a public key for encryption and a private key for decryption [143]. Each of these keys needs different resources and, in practice, both types of encryption are used [144]. In fact, the state of art presents a very division in terms of symmetric and asymmetric keys for SG when the application is related to the SM itself. For instance, the homomorphic Paillier cryptosystem, which is based on the Discrete Logarithm Problem (DLP), is a type of asymmetric key, and is proposed as the solution for SM in the solutions presented in [145], while other DLP-based algorithms are also proposed for a SM application in [132] [146]. However, in [114] they are described as not desirable for SG, which typically has limited resources. The Paillier encryption is also mentioned as not computationally efficient due to its expensive operations [29] [136].

Lightweight keys can serve to protect data, as long as the key is inaccessible. In [147], a comparison of computational overheads among XOR, Shamir's Secret Sharing and homomorphic encryption is presented. If C_1 is the cost associated to the XOR operation, C_2 is the cost associated with the Shamir's Secret Sharing Scheme and C_3 with homomorphic encryption, then $C_1 < C_2 \ll C_3$. Due to its extremely low overhead, XOR keys are used in AES, E-DES and Blowfish Encryptions [148] and utilized also as encryption method in [30].

In the realm of SG, [137] presents an SSS-based distributed communication architecture that guarantees the privacy of fine-grained users data while enabling the energy supplier to access aggregate energy measurements and per-user coarse-grained data for billing purposes. Their privacy preserving communication architecture is based on sharing encryption keys rather than the energy consumption values themselves, considering a semi-honest adversary.

4.2.3 Blockchain in the Smart Grid

In blockchains, a problem relates to privacy, as all transactions are public [118]. Blockchains are designed to achieve peer-to-peer electronic payments directly, without participation of a trusted third party [113] and, as such, they presuppose the lack of a central authority or coordinator from having access to all registers and actions of a network. This assumption collides with the role of the power utility companies, which are held accountable by the local regulators about the electricity assets in their area, i.e., they are responsible for billing costumers, surveying the use of power grid assets, and so further. Furthermore, as largely adopted in the state of the art [137, 139, 146], [149, 150], a TTP can be adopted for the data exchange system in the SG. The role of a TTP is frequently assigned to the utility company due to its natural position in the respective SG network. At least in such cases, the financial compensation between traders demands an entity managing the energy exchanges and the respective financial transfers.

In [13], a new currency, the NRGcoin, is proposed, however without a detailed description of the compensation mechanism mentioned here, i.e., how the financial transfer occurs and who is in charge of surveying it. In [118], a novel blockchain-based transactive energy system is described for energy trade between final prosumers. However, the Distribution System Operator (DSO) is set to ensure the safe operation of the microgrid and regulate its total load. In order to achieve so, the DSO can limit the energy and financial assets that the prosumers' withdraw for trading. The DSO can also set a price policy for the microgrid, i.e., the DSO operates as a TTP. For all these limitations, we envisage great challenges to employing blockchains for the specific case of SG trading systems.

4.2.4 Trading Systems for SG prosumers

In a NAN, each household unit is represented as a Home Area Network (HAN) and is equipped with a SM [151]. Some of the HANs are prosumers that in some occasions export energy to the grid from their Distributed Energy Resources (DER).

From the standpoint of energy trade, there are two predominant perspectives in the literature. When energy is seen as a public service, the tendency is the proposal of cooperative games as in [152] [153] [154], where the underlying goal is market control and the achievement of social fairness. Alternatively, when the main goals are market efficiency and decentralization [150] [155] [156], energy is seen as commodity to be traded, and a free market model is sought. In this work, we align our analysis to the latter group. Table IV.2 summarizes the division of approaches in the literature with regards to the commercial treatment of energy.

Table IV.2: Commercial treatment of energy and the respective approaches in the literature

Criterion	Cooperative games	Competitive games
How energy is predominantly seen	as a public good	as a commodity
Main goals	social fairness, market control	decentralization, market efficiency
Main references	[139] [149] [153] [154] [157] [158] [159] [160] [161] [162] [163] [164] [165] [166]	[14] [27] [28] [150] [155] [156] [167] [168] [169] [170] [171] [172] [173] [174]

There are three different compensation mechanisms in [175] associated to topologies of DER installation and billing regimes. These mechanisms are related to the relationship between prosumer and utility company only. The first one is the Net Energy Metering (NEM), depicted in Fig. 4.1, that allows a DER that is generating more electricity than the consumed by the household to export the excess of energy to the grid, earning a corresponding credit in kilowatt-hours (kWh). In order to correctly effectuate the measurements, in the NEM, the SM is bi-directional, i.e., it spins backwards when production surpasses consumption. The Buy All, Sell All system, shown in Fig. 4.2, is an arrangement that provides a standard sell rate to a DER system for all of the electricity they generate. In Buy All, Sell All schemes, the HAN cannot consume the energy that the DER produces, exporting it entirely. The third arrangement, as illustrated in Fig. 4.3, is the Net Billing, in which the DER system owner can consume electricity generated by the DER in real time and export any generation in excess of on-site consumption to the grid, so that all net energy exports are metered and credited at a predetermined sell rate, without kWh banking [175]. The main features of the three compensation mechanisms are summarized in Table IV.3.

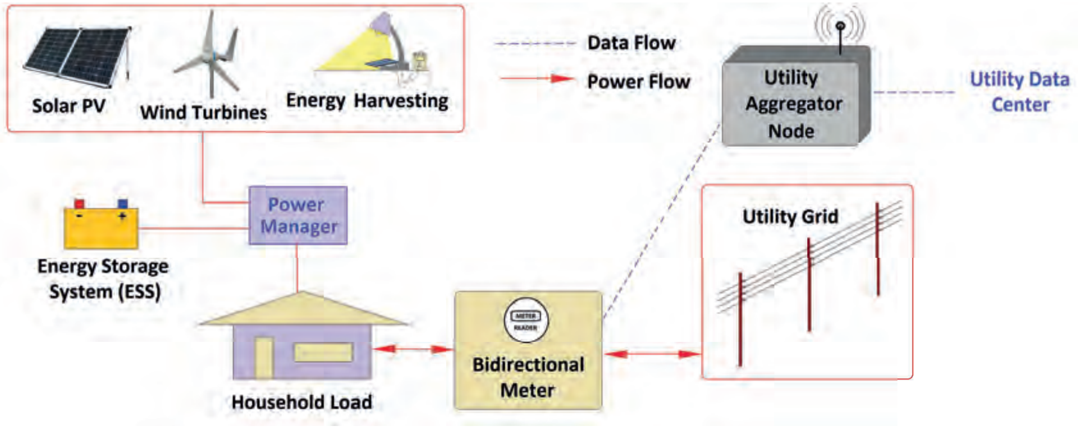


Figure 4.1: In the Net Energy Metering topology, the SM is bi-directional and exceeding kWh produced in a billing cycle can be banked as kWh credits.

Since the prosumers in a NAN are able to trade their energy not only with his neighbors but also with the utility company, the aforementioned mechanisms must consider consumer-to-consumer trade designs. The NEM characteristic of not allowing financial reward for the exceeding energy makes it unable to the proposed frameworks, which are applicable to the Buy All, Sell All and Net Billing systems.

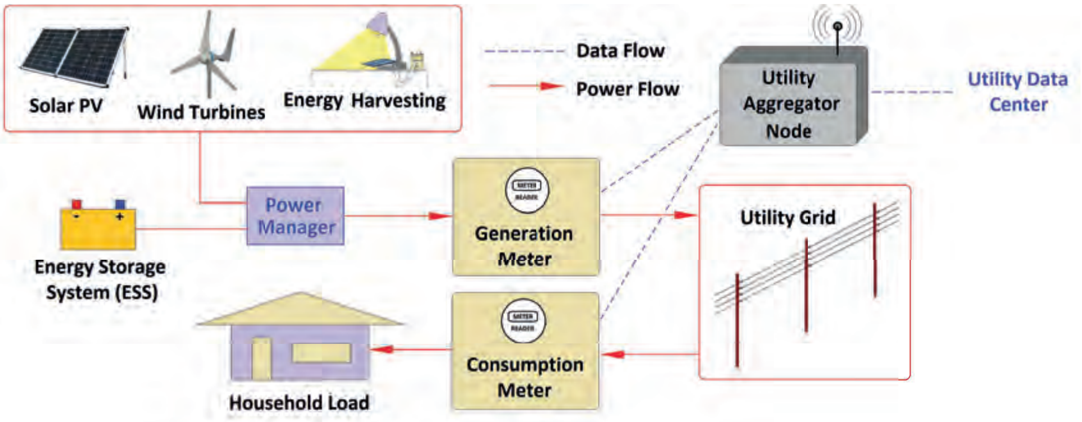


Figure 4.2: In the Buy All, Sell All topology, all produced energy is exported to the grid, being rewarded under a previously defined sell-rate. Note that the generation and consumption infrastructures are completely separated.

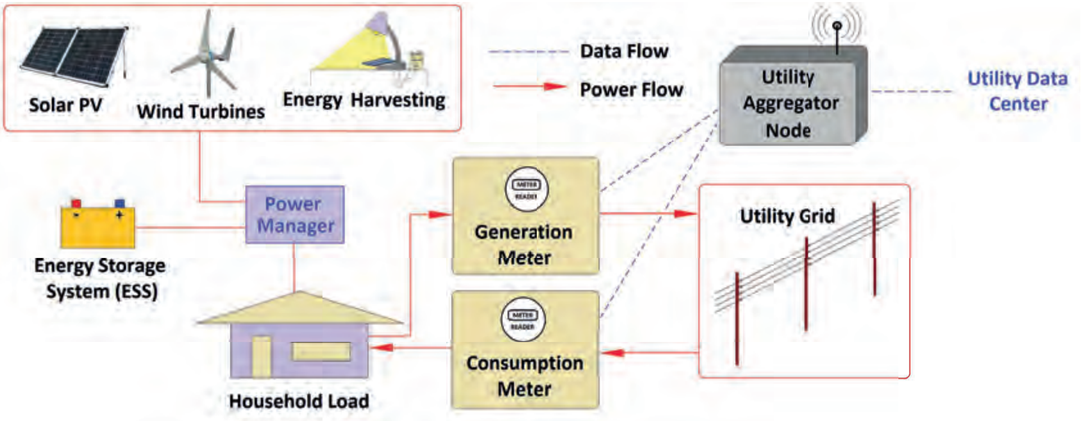


Figure 4.3: In the Net Billing topology, a HAN can consume electricity generated by their DER system in real time and export any generation in excess to the grid. The exported energy is also rewarded by means of a sell rate, without kWh banking.

With regards to energy prices in the SG, the consumers can purchase energy either from the power grid or from other prosumers. In the first case, a key aspect is that the utility company, which is the company responsible for managing all power grid assets and its operation, usually sells the energy for a unitary price P_u to its consumers and purchases the energy from the prosumers by a smaller price, P_l [152] [175] [176]. Therefore, an interval of energy unitary prices that enable trades between prosumers and consumers is given by $P_l < P_b < P_u$, where P_b is the price of local traders with which all final users obtain profit with regards to P_u and P_l, since every purchaser is supposed to prefer paying $P_b < P_u$ for the kWh, with a symmetric interpretation by the side of the sellers.

Table IV.3: Metering and Billing Arrangements for Compensation Mechanisms [175]

Compensation Mechanism	Net Energy Metering	Buy All, Sell All	Net Billing
Self-Consumption Allowed	Yes	No	Yes
Netting Frequency	Billing Cycle	Billing Cycle	Instantaneous
Intra-Billing Cycle Banking of kWh	Yes	No	No
Key Benefits	Simplicity	No reduced sales for the utility company. Potential for more precise compensation for DG production	Encourages self-consumption
Challenges	Reduced utility company sales	Customers may illegally wire for self-consumption if more economically desirable and utility enforcement unlikely	Reduced utility company sales
Proposed Framework	No	Yes	Yes

4.2.5 The Asmuth-Bloom's CRT Based SSS

The Asmuth-Bloom's SSS is based on the Chinese Remainder Theorem (CRT). The CRT explains how to solve an algebra problem in which an integer-valued N is determined from its remainders, as in (2.1). The theoretical interval in which the number N is uniquely determinable is the so-called dynamic range

$$d = \mathrm{LCM}(M_1, M_2, \ldots, M_L), \tag{4.1}$$

where M_i, for $i \in \{1, 2, \ldots, L\}$ are the co-prime moduli, LCM denotes the least common multiple of a set of numbers [60] [61], and $0 \leq N < d$ is the range for unique values of N. From the knowledge of M_i and r_i, the CRT offers the straightforward calculation [43, 49].

In the Asmuth-Bloom's (n, t) SSS, there are n participants and it is necessary a minimal coalition of t players to reconstruct the secret $S \in \mathbb{Z}$. The dealer creates co-prime integers $M_0 < M_1 < M_2 < \cdots < M_n$, subject to

$$\prod_{i=1}^{t} M_i > M_0 \prod_{i=1}^{t-1} M_{n-i+1} \tag{4.2}$$

and

$$M_0 > S. \tag{4.3}$$

The dynamic range is defined as the interval where the secret lies and is given by

$$D = \prod_{i=1}^{t} M_i. \tag{4.4}$$

Eq. (4.4) shows that D must be covered by the shares of the t lowest moduli, ensuring that the results are the same for any possible coalition of at least t players. The dealer computes

$$y = S + aM_0, \tag{4.5}$$

where a is any random positive integer that preserves the condition $0 \leq y < D$. Instead of transmitting y, which is a linear combination of the secret S, the dealer distributes the j-th specific share of y to j-th player as follows:

$$y_j = y \bmod M_j, \tag{4.6}$$

where $j \in \{1, 2, \ldots, n\}$.

Eqs. (4.2)–(4.6) delimit the *Sharing Phase*. In sequel, the *Construction Phase* is carried out with t or more shareholders exchange mutually their shares. Note that gathering n shares as of (4.6) yields the system of (2.1) for $n = L$. The underlying algebraic process of Asmooth-Bloom's SSS ensures that, whenever $(t-1)$ shares are collected, the range of possible values of S in function of possible values for the k-th share always spans entirely the set $\{0, 1, 2, \ldots, M_0 - 1\}$. For a deeper analysis, see [125, 127].

Example 1. Let a SSS of the type $(5, 3)$ where $M_1 = 11$, $M_2 = 13$, $M_3 = 16$, $M_4 = 17$ and $M_5 = 19$. Since $t = 3$ shares are needed to reconstruct S, the dealer in an Asmuth-Bloom's SSS obtains via (4.2) the maximum value for M_0,

$$M_0 < \frac{11 \cdot 13 \cdot 16}{19 \cdot 17}, \tag{4.7}$$

from which he achieves $M_0 = 7$ since it is the maximum integer-valued number that satisfies (4.7), co-prime with every M_j, for $j \in \{1, 2, \ldots, n\}$, and since $M_0 < M_1$. As a consequence of (4.3), the possible values for the secret are $S \in \{0, 1, 2, \ldots, 6\}$. Now the dealer chooses an arbitrary value for a which copes with $0 \leq S + 7a < 2288$ and calculates the shares y_j, $j \in \{1, 2, 3, 4, 5\}$, distributing them individually over all the 5 participants.

In order to illustrate Asmuth-Bloom's CRT perfection, let the owners of the moduli $M_4 = 17$ and $M_5 = 19$ exchange their shares, obtaining the system

$$\begin{cases} y \bmod 17 = 0 \\ y \bmod 19 = 14, \end{cases} \tag{4.8}$$

and that both shareholders intend to infer the possible values for S by speculating about the remainder of another participant, which they still do not know. If they have success in obtaining a narrower set of values for S only by knowing $(t-1)$ shares, this means that the system is not perfect. As any other participant, they know in advance that $M_0 = 7$. Supposing that they have learned the value of for instance $M_1 = 11$ from previous interactions, they try to speculate about the values of $y_1 \in \{0, 1, 2, \ldots, 10\}$ which belongs to the first shareholder. In each of these values, if they calculate y and next apply $S = y \bmod M_0$, the sequence of outputs is shown in the vector

$$\mathbf{s} = [\,5 \quad 1 \quad 4 \quad 3 \quad 6 \quad 2 \quad 1 \quad 4 \quad 0 \quad 3 \quad 2\,]. \tag{4.9}$$

What [125] and [127] prove is that, whenever less than t shareholders try to envisage the value of S, all the numbers in the set $\{0, 1, 2, \ldots, M_0 - 1\}$ always appear in the vector $\mathbf{s}$ shown in (4.9).

According to [129], the sizes of the share spaces and the secret space are not equal, meaning that the values of the shares are higher than the secret itself. In order to see this, it suffices to note that $S < M_0 < M_1 < \cdots < M_n$, and that each share contains the value of M_j, for $j \in \{1, 2, \ldots, n\}$. Hence, despite its perfection, the Asmuth-Bloom's SSS has as a drawback the range of values for S, which is confined to M_0 as shown in (4.3).

4.2.6 Background of the Tensorial Approach for the CRT

Considering a CRT system of L rows as shown in (2.1), the remainders and respective moduli $N \bmod M_j = r_j$ can be described by row vectors $\mathbf{e}_j \in \mathbb{Z}^{1\times M_j}$, where all entries are zero, except the r_j-th entry, which has value 1. When $r_j = 0$, then the M_j-th entry has value 1. This equivalence is proven in Chapter II [1] and represents an alternative form of solving the CRT in the error-free case. For instance, if the second row of a CRT system is $N \bmod 9 = 6$, the vector

$$\mathbf{e}_2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \tag{4.10}$$

is assembled as a representation of this row. Using D as obtained in (4.4), denoting

$$w_j = D/M_j, \tag{4.11}$$

for $j \in \{1, 2, \ldots, n\}$, and setting up the row vector $\mathbf{u}_p \in \mathbb{Z}^{1\times p}$ with 1 in all entries, the value of $0 \leq N < D$ in such system is the cardinality of the only column in the matrix

$$\mathbf{V}_e = \begin{bmatrix} \mathbf{u}_{w_1} \otimes \mathbf{e}_1 \\ \mathbf{u}_{w_2} \otimes \mathbf{e}_2 \\ \vdots \quad \vdots \\ \mathbf{u}_{w_L} \otimes \mathbf{e}_L \end{bmatrix} \tag{4.12}$$

that satisfies

$$\mathbf{V}_e(j, N) = 1, \tag{4.13}$$

for $j \in \{1, 2, \ldots, L\}$ as the CRT of (2.1) comprises L rows. Note that the vectors $\mathbf{e}_i$ and $\mathbf{u}_p$ are row vectors, differently of Eqs. (2.29)-(2.33) in Chapter II, where they were adopted as column vectors. Note still that determining the cardinality of the column with all-one entries of $\mathbf{V}_e$ as in (4.13) corresponds to finding the only non-zero entry in a row vector $\mathbf{v}_e \in \mathbb{Z}^{1\times D}$ assembled by means of

$$\mathbf{v}_e = \left[\mathbf{u}_{w_1} \otimes \mathbf{e}_1\right] \odot \left[\mathbf{u}_{w_2} \otimes \mathbf{e}_2\right] \odot \cdots \odot \left[\mathbf{u}_{w_L} \otimes \mathbf{e}_L\right]. \tag{4.14}$$

Given that each row of $\mathbf{V}_e$ in (4.12) is an input value for the Schur product in (4.14), the cardinality of the only column in $\mathbf{V}_e$ with all entries 1 is the same of the only entry 1 of $\mathbf{v}_e$ in (4.14). Therefore, N can be calculated by the Kronecker and Schur products of (4.14). The routine comprised by Eqs. (4.10)-(4.14) can be seen as an alternative version of the expression (2.25) as it refers to determining the value of N in terms of sets intersection.

4.3 Proposed Framework for the NAN Electricity Trading System

We start this section by specifying the technical requirements with which our system is supposed to cope. These requirements derive from the state-of-the-art remarks made in Section 4.2. In sequel, we present in Section 4.3.2 the novel perfect CRT based SSS, which is used as a security feature in the proposed privacy-oriented data security system shown in Section 4.3.3 and the trading system in Section 4.3.4.

4.3.1 Requirements for the NAN Architecture and Security Framework

As technical requirements for data security in the proposed framework, the number of interactions between a node/unit and the central controller, which is TTP, as well as between final nodes should be minimal. Each node must have a different AES 128 bit key, and the encrypted messages can be combined with XOR encryption, as its main feature is low cost and simplicity. In our framework, time-stamps are also to be used as an additional way of ensuring the correctness of the sender identity, constituting an extra argument for symmetric keys.

The system must be de-anonymization proof. External observers are not allowed to know the identities of the bidders, the quantity of them, if there are bids of which type, nor if any deal is achieved. An example of NAN is sketched Fig. 4.4, where there are two spies, namely house 5 and a vehicule that drives through the area. As house 5 has access to the system, it learns the quantity of bids and which of them are actually converted into energy trade. However, for an external observer, the system must be altogether blind. In any case, none of them not allowed to discover who are the bidders. They must be also prevented from impairing communication data in terms of integrity and confidentiality without being discovered. Even houses 3 and 10, which respectively sells and buy the amount of energy e, also must ignore mutually their roles in the trading session.

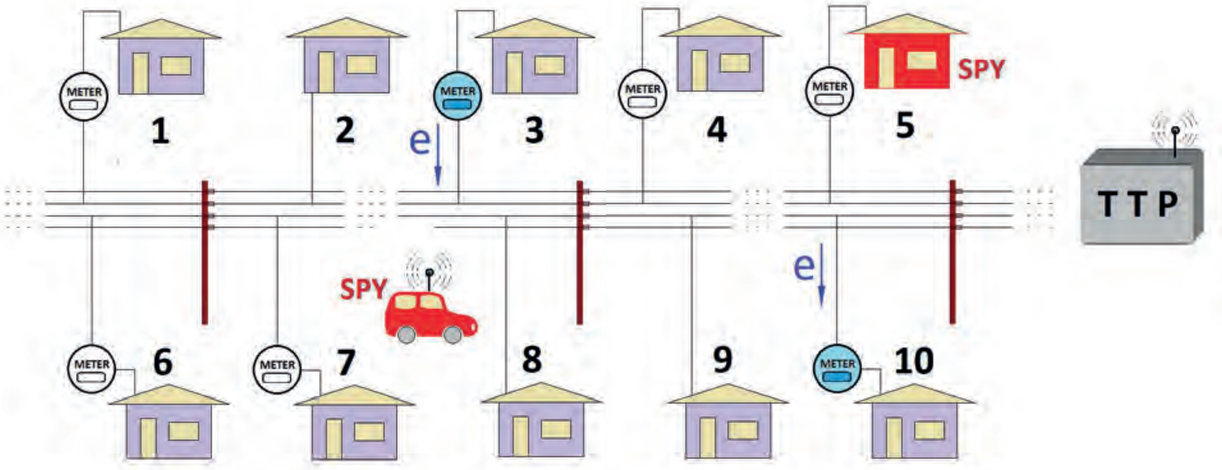


Figure 4.4: As house 5 has access to the system, it learns the quantity of bids and which of them are actually converted into energy trade. However, for an external observer, the system must be altogether blind, and none of them not allowed to discover who are the bidders. Houses 3 and 10 also must ignore mutually their roles in the trading session.

The system must resist traffic analysis and impersonation attacks. The latter is considered achieved if the AES key of the node is kept undiscovered by the attacker. The adversary is supposed to be malicious and powerful, counting on an virtually infinite computational capability.

Although AES 128 bit encryption key is reputed secure, an active attacker can infer recurrent data and identify patterns if ciphertexts are repeated. Note that the repetition of ciphertexts in the case of SG trades is likely to occur since prices and quantities of energy can lie around typical values, easing the task of an adversary of identifying the occurrence of offers with similar characteristics. In order to prevent this drawback, we adopt a Linear-Feedback Shift Register (LFSR) [177], which provides a linear function of the previous state of a sequence of bits according to the value of the most left-sited one at each iteration. The initial value of the LFSR is called seed and the bits that influence the next values of the LFSR are called taps. The period of a LFSR is the minimal number of different outputs before repeating its seed and is given by $p = 2^n - 1$, where n is the highest position of the tap that makes the feedback polynomial achieve the maximum possible period. Tables of the taps for maximum-length LFSR in function of each n up to 168 bits are given in [178].

In the considered NAN, the Aggregator, or central operator, plays the role of a TTP. We use the communication system shown in [179], which provides a reliable wireless intra-battery management system and handles low values of signal-to-noise-plus-interference ratio (SNIR) by varying the length of direct sequences of bits. This is achieved by means of code division multiplexing of several decentralized controllers with a central controller. In doing so, the proposed patent provides a reliable and adaptive link for communication between the TTP, which is the central controller, and the consumers in a NAN. The patent in [179] can incorporate different families of codes, including for instance Walsh, Gold, M-sequence, Kasami and Chaos, as well as different modulation schemes, such as Phase Shift Keying (PSK), Quadrature PSK (QPSK) and Chirp Spread Spectrum (CSS). The system in [179] outperforms systems such as ZigBee, Bluetooth and LoRa in terms of bit error rate (BER) and latency for critical safety applications.

4.3.2 Proposed CRT Based SSS

For an SSS of the type (n, t), our proposal consists in a CRT that is solved with aid of some tensorial algebra operations such as the Kronecker and Schur products. The secret $S \in \mathbb{Z}$ has a bijective relationship with the possible permutations of k numbers in a row vector $\mathbf{s} \in \mathbb{Z}^{1 \times k}$. First, we see how to produce the sequence in the vector $\mathbf{s}$, and next we exploit the bijective relationship between the vector $\mathbf{s}$ and the scalar S.

We adapt the system of (4.12), which is used to encounter the value of a single scalar N in a CRT system as that of (2.1), in order to generate k pairs of integers (x_i, N_i), with $x_p < x_q$ for $p < q$ and $N_p \neq N_q$ for $p \neq q$, over the range $0 \leq N_i < D$, for $i \in \{1, 2, \ldots, k\}$. The values x_i are taken from the vector $\mathbf{x} = \begin{bmatrix} 1 & 2 & \ldots & k \end{bmatrix}$, such that $x_i = \mathbf{x}(i)$. In order to produce the k pairs, the dealer endows the vectors of (4.10) with a second dimension whose size is k, obtaining the share matrices

$$\mathbf{E}_j(i, q) = \begin{cases} 1, \text{ if } \mathrm{mod}(N_i, M_i) = q \text{ and } q \in \{1, \ldots, M_i - 1\}, \\ 1, \text{ if } \mathrm{mod}(N_i, M_i) = 0 \text{ and } q = M_i, \\ 0, \text{ otherwise}, \end{cases} \tag{4.15}$$

where $\mathbf{E}_j \in \mathbb{Z}^{k \times M_j}$ for $j \in \{1, 2, \ldots, n\}$ and $i \in \{1, 2, \ldots, k\}$. As a consequence, each row of $\mathbf{E}_j$ has exactly one entry 1, exactly because each row of $\mathbf{E}_j$ reproduces a vector as $\mathbf{e}_j$ given by (4.10). At least t shares as of (4.15) are combined via

$$\mathbf{V} = \left[\mathbf{u}_{w_1} \otimes \mathbf{E}_1\right] \odot \left[\mathbf{u}_{w_2} \otimes \mathbf{E}_2\right] \odot \cdots \odot \left[\mathbf{u}_{w_t} \otimes \mathbf{E}_t\right], \tag{4.16}$$

where the similarity with the vector $\mathbf{v}_e$ given by (4.14) is evident. In fact, $\mathbf{V}$ in (4.16) can be seen as k stacked vectors $\mathbf{v}_e$ of (4.14), each one with a different N_i, $i \in \{1, 2, \ldots, k\}$. Note that if one desires to compare matrix $\mathbf{V}$ to the matrix $\mathbf{V}_e$, two operations are required. First, adapting the notation of $\mathbf{V}_e$ and considering t shares of the form $\mathbf{V}_{e,i}$, for $i \in \{1, 2, \ldots, t\}$, the tensor

$$\boldsymbol{\mathcal{V}} = \mathbf{V}_{e,1} \sqcup_3 \mathbf{V}_{e,2} \sqcup_3 \cdots \sqcup_3 \mathbf{V}_{e,t}, \tag{4.17}$$

is created, where $\boldsymbol{\mathcal{V}} \in \mathbb{Z}^{k \times M_i \times t}$ and $\sqcup_r$ denotes the concatenation of two tensors along the r-the mode. Next, for all the $(k \cdot M_i)$ 3rd-dimension fibers of $\boldsymbol{\mathcal{V}}$, we make

$$\mathbf{V}(i, q) = \begin{cases} 1, \text{ if } \boldsymbol{\mathcal{V}}(i, q, :) = \mathbf{u}_t, \\ 0, \text{ otherwise}, \end{cases} \tag{4.18}$$

where $\mathbf{u}_t \in \mathbb{Z}^{1 \times t}$ is an all-one entries vector. Therefore, the value 1 is assigned to each entry of $\mathbf{V}$ that corresponds to each 3rd-dimension fiber of tensor $\boldsymbol{\mathcal{V}}$ with all entries 1. Note that applying either (4.16) or (4.18) delivers the same matrix $\mathbf{V} \in \mathbf{Z}^{k \times D}$, which bears a unique entry whose value is 1 in each row, and that there is not a column of $\mathbf{V}$ that has more than one entry 1.

After processing (4.16), we set up the vector

$$\mathbf{v} = \mathbf{x}\mathbf{V}, \tag{4.19}$$

for obtaining the secret vector

$$\mathbf{s} = \mathbf{v}(\mathbf{v} \neq 0), \tag{4.20}$$

meaning that the row vector $\mathbf{s} \in \mathbb{Z}^k$ receives only the k nonzero values of $\mathbf{v}$ in the sequence that they appear in $\mathbf{v}$.

In the proposed SSS of the type (n, t), the secret is a number S that is associated to a permutation of the sequence $\{1, 2, \ldots, k\}$ that is encountered in the vector $\mathbf{s}$ of (4.20). As a consequence, $k!$ possible values for S exist, such that $S \in \{0, 1, \ldots, (k! - 1)\}$. A table of correspondence could be assembled with the aim of linking a value of S with a specific permutation of elements in $\mathbf{s}$; however, with sufficiently large values of k, such table can be very memory expensive. A preferable solution is employing a mapping algorithm that assigns a number to each permutation. One of the most efficient codes of this type is the Lehmer Code [180], which is shown in Appendix A. With aid of Algorithm 6, once the value of S is determined, the dealer converts it to the corresponding permutation in $\mathbf{s}$, spreading the x_i values over the range D according to pairs (x_i, N_i) and following $\mathbf{v}(N_i) = x_i$, for $i \in \{1, 2, \ldots, k\}$. Likewise, using the same code, each shareholder can calculate S once the sequence in $\mathbf{s}$ is revealed. Note that the Lehmer Code establishes a bijective relationship between S and a given permutation in $\mathbf{s}$.

Theorem 1. *The maximum allowed number k of nonzero elements in each share $\mathbf{X}_j$ is the maximum value that satisfies*

$$D > k \prod_{i=1}^{t-1} M_{n-i+1}. \tag{4.21}$$

The proof is presented in Appendix B. A geometric interpretation of what the violation of (4.21) entails is shown in Fig. 4.5. Defining

$$v_{-1} = \prod_{i=1}^{t-1} M_{n-i+1}, \tag{4.22}$$

and $\mathbf{v}_{-1}$ as the resulting matrix of the highest $(t-1)$ shares $\mathbf{E}_j$, as shown in Fig. 4.5(a), so that $\mathbf{v}_{-1} \in \mathbb{R}^{k \times v_{-1}}$. Suppose that $k = 4$, hence making matrix to be graphically repeated $k = 4$ times laterally in an equivalent visualization of the Kronecker product $\mathbf{u}_4 \otimes \mathbf{v}_{-1}$. This example presupposes that $3v_{-1} < D < 4v_{-1}$, and that each share $\mathbf{E}_j$ has $k = 4$ nonzero values. The entries a, b, c, d are used to denote the sequence of N_i that show up in $\mathbf{v}_{-1}$, each one referring to a row of $\mathbf{v}_{-1}$. Note that any vector of k nonzero values must be repeated laterally at least k times in order to provide any permutation of the k entries, including their palindrome, which starts by its last character and ends with its first one. In Fig. 4.5(b), the dynamic range D does not include the first entry a at the fourth repetition of $\mathbf{v}_{-1}$ within the length D, which is a possibility as the entries can be elsewhere located in $\mathbf{v}_{-1}$. As a consequence, given the dimension of $\mathbf{v}_{-1}$, the maximum number of nonzero elements in each share that ensures the palindrome production is $k = 3$, and at least one of the entries a, b, c, d must be dismissed from each share, implying that the final vector obtained with at least k shares comprises $k! = 6$ possible permutations of the entries.

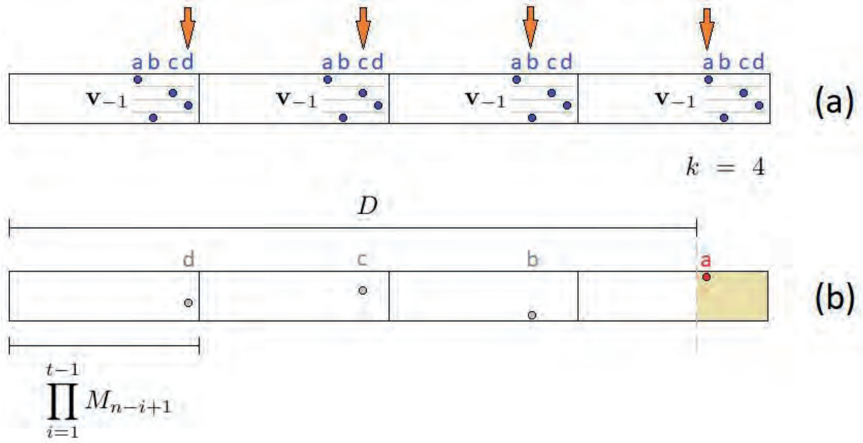


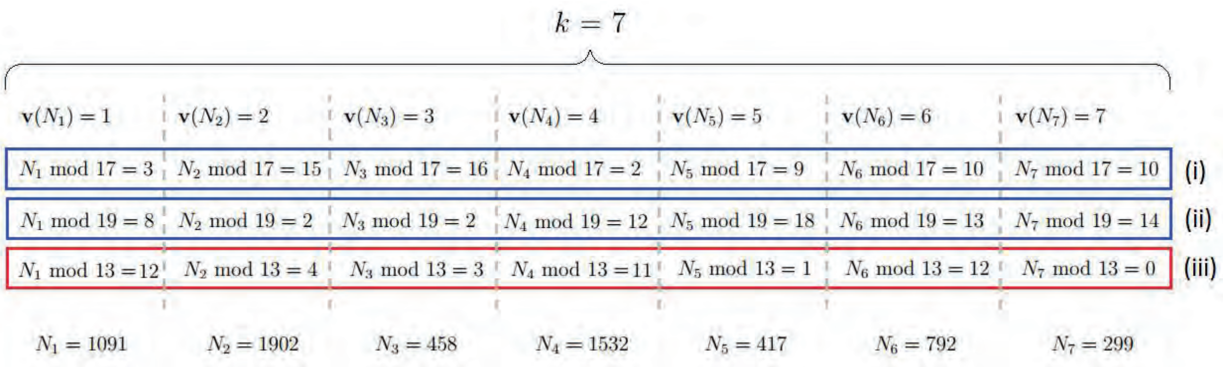
Figure 4.5: Example of what the violation of (4.21) entails. In (a), a hypothetical vector $\mathbf{v}_{-1}$ has its nonzero entries in such a position that, after the Kronecker product $u_4 \otimes \mathbf{v}_{-1}$, the first entry remains out of the dynamic range D, as shown in (b). The maximum number k of nonzero values in each share hence must be the number of repetitions of $\mathbf{v}_{-1}$ in D, which in this example is 3.

From the above, the dealer can share an integer-valued secret $S \in \{0, 1, \ldots, (k! - 1)\}$ among n players so that at least t of them exchange their shares mutually by undertaking the following steps:

- The dealer selects a number $S \in \{0, 1, \ldots, k! - 1\}$ and next converts it into a sequence $\mathbf{s} \in \mathbb{Z}^k$ in accordance to the Lehmer Code;
- The dealer assigns to each entry in $\mathbf{s}$ a number $N_i \in \{1, 2, \ldots, D\}$, so that $\mathbf{v}(N_i) = \mathbf{s}(i)$ for $i \in \{1, 2, \ldots, k\}$, subject to $N_1 < N_2 < \cdots < N_k$, forming the pairs (x_i, N_i);
- For $j \in \{1, 2, \ldots, n\}$, the dealer assembles matrices $\mathbf{X}_j$ in accordance with (4.15) and secretly distributes them to the shareholders.

Example 2. Suppose the same SSS moduli of Example 1, i.e., $M_1 = 11$, $M_2 = 13$, $M_3 = 16$, $M_4 = 17$ and $M_5 = 19$. Applying (4.21), $k = 7$ is the maximum nonzero numbers in each share $\mathbf{E}_j \in \mathbf{Z}^{7 \times M_j}$, for $ij \in \{1, 2, 3, 4, 5\}$. As in the former example, the owners of the shares (4.8) $M_4 = 17$ and $M_5 = 19$ intend to know the secret by speculating about the share of the owner of $M_1 = 11$. Given that $k = 7$, the shares of the fourth and fifth shareholders are sparse matrices, respectively $\mathbf{E}_4 \in \mathbb{Z}^{7\times 17}$ and $\mathbf{E}_5 \in \mathbb{Z}^{7\times 19}$. Let the entries of nonzero values in the share $\mathbf{E}_4$ be $(1,3)$,$(2,15)$,$(3,16)$,$(4,2)$,$(5,9)$,$(6,10)$,$(7,10)$ and in the share $\mathbf{E}_5$ be $(1,8)$,$(2,2)$,$(3,2)$,$(4,12)$,$(5,18)$,$(6,13)$,$(7,14)$, with 1 in all of these entries and zero elsewhere. The shareholders try to infer the secret by testing $\mathbf{E}_1$ with 1 in each entry of each row, subject to 7 non-coincident nonzero outputs in vector $\mathbf{v}$ of (4.19), from which they obtain 823543 possible values for S by means of (4.20) and the Lehmer Code. All the values $S \in \{0, 1, 2, \ldots, D-1\}$ appear in the 823543 tests, as expected in a perfect system.

An alternative representation of each share in (4.15) is a vector of remainders and the modulus M_i. The content of $\mathbf{E}_j$, for $j \in \{1, 2, \ldots, n\}$, is the equivalent to the set of remainders shown in Fig. 4.6. Maintaining the Example 2 with the (5,3) SSS, let $N_1 = 1091, N_2 = 1902, N_3 = 458, N_4 = 1532, N_5 = 417, N_6 = 792$ and $N_7 = 299$. Note that $N_j \in \{1, 2, \ldots, D\}$, for $j \in \{1, 2, \ldots, n\}$, as $D = 11 \cdot 13 \cdot 16 = 2288$, and that $\mathbf{v}(N_i) = i$, for $i \in \{1, 2, \ldots, k\}$. The key aspect of the proposed method is that, with the first two shares (i) and (ii), $k!$ permutations of N_i along the entries of $\mathbf{v}$ can be produced. Only when the t-th share (iii) is given, all N_i are determined and the sequence of values in $\mathbf{v}$ is cleared.



$k = 7$

$\mathbf{v}(N_1) = 1$	$\mathbf{v}(N_2) = 2$	$\mathbf{v}(N_3) = 3$	$\mathbf{v}(N_4) = 4$	$\mathbf{v}(N_5) = 5$	$\mathbf{v}(N_6) = 6$	$\mathbf{v}(N_7) = 7$	
$N_1 \bmod 17 = 3$	$N_2 \bmod 17 = 15$	$N_3 \bmod 17 = 16$	$N_4 \bmod 17 = 2$	$N_5 \bmod 17 = 9$	$N_6 \bmod 17 = 10$	$N_7 \bmod 17 = 10$	(i)
$N_1 \bmod 19 = 8$	$N_2 \bmod 19 = 2$	$N_3 \bmod 19 = 2$	$N_4 \bmod 19 = 12$	$N_5 \bmod 19 = 18$	$N_6 \bmod 19 = 13$	$N_7 \bmod 19 = 14$	(ii)
$N_1 \bmod 13 = 12$	$N_2 \bmod 13 = 4$	$N_3 \bmod 13 = 3$	$N_4 \bmod 13 = 11$	$N_5 \bmod 13 = 1$	$N_6 \bmod 13 = 12$	$N_7 \bmod 13 = 0$	(iii)
$N_1 = 1091$	$N_2 = 1902$	$N_3 = 458$	$N_4 = 1532$	$N_5 = 417$	$N_6 = 792$	$N_7 = 299$	

Figure 4.6: With the first two shares (i) and (ii), $k!$ permutations of N_i can be produced. Only when the t-th share (iii) is given, all N_i are determined and therefore the sequence of values in $\mathbf{v}$ is cleared.

Following Fig. 4.6, the vector $\mathbf{v}$ produced has $k = 7$ non-zero entries with $\mathbf{v}(1091) = 1$, $\mathbf{v}(1902) = 2$, and so forth, yielding

$$\mathbf{v} = [\,0\ \ldots\ 0\ 7\ 0\ \ldots\ 0\ 5\ 0\ \ldots\ 0\ 3\ 0\ \ldots\ 0\ 6\ 0\ \ldots\ 0\ 1\ 0\ \ldots\ 0\ 4\ 0\ \ldots\ 0\ 2\ 0\ \ldots\ 0\,], \tag{4.23}$$

from which, applying (4.20), we obtain

$$\mathbf{s} = [\ 7\ \ 5\ \ 3\ \ 6\ \ 1\ \ 4\ \ 2\ ]. \tag{4.24}$$

Finally, by means of the Lehmer code, (4.24) informs the secret

$$S = 3549. \tag{4.25}$$

Note that the proposed SSS can be adapted in terms of the secret range. It has been shown that $S = k! - 1$ defines the maximum value of the secret in the range of $k!$ values, as zero is included. If a range different from $k!$ values is desired, an adjustment is to establish a new maximum value σ for the secret, as in

$$S' = S \bmod \sigma, \tag{4.26}$$

where $k! > \sigma$ and S' denotes the new secret maximum value. This is useful in cases where S' must be an integer that does not correspond to any possible $(k! - 1)$. Suppose for instance that a secret range of exactly $\sigma = 10^{15}$ distinct values is needed. There is not a $k \in \mathbb{Z}$ for which $k! - 1 = 10^{15}$. However, for $k = 18$, $k! - 1 = 6.4 \cdot 10^{15}$, and if (4.26) is applied, the new secret range is produced without impairing the perfection of the SSS.

4.3.3 Privacy-Oriented Data Security System

The privacy-oriented system is explained with aid of Fig. 4.7. As a requirement presented in Section 4.3.1, each node has a different AES key with the TTP that is created when the node purchases and installs a SM. Therefore, the creation of the AES key between each user and the TTP is a previous step to the following framework.

During a day, regular intervals in which a trading session can happen are called time slots. We adopt 15 minutes for each slot, as in [118]. The initial time slot is called the slot zero, which corresponds to Step 1 in Fig. 4.7. In this slot, each SM receives from the TTP a ciphertext on AES 128 bit encompassing as contents the XOR keys $\mathbf{K}_1$ and $\mathbf{K}_2$ along with the LFSR seeds and taps, and the "SM Schedule". The latter is a list of the time slots in which the respective node must act as a data confirmation agent if a trading session takes place at that time slot. The nodes do not learn the SM Schedule of the other nodes. Since 95 slots are specified over a day, the tap of highest order in each LFSR used must be $n \geq 7$, as in this case, $p = 2^7 - 1 = 127$ different keys. The keys $\mathbf{K}_1$ and $\mathbf{K}_2$ are bit matrices with dimensions $\mathbf{K}_1, \mathbf{K}_2 \in \mathbb{Z}^{n \times u}$, where n is the number of houses with a SM in the NAN and u is the length of each bid. In this work, we adopt $u = 32$. Each row of $\mathbf{K}_1$ and $\mathbf{K}_2$ is a different

LFSR with its own seed and taps. Note that the rows $\mathbf{K}_1$ and $\mathbf{K}_2$ are updated by the SM itself between two subsequent time slots.

Before a time slot ends, any authenticated node that desires to trade energy forwards to the TTP a purchase or a sell bid, which in Fig. 4.7 occurs in Step 2, i.e., the bid submission. The plaintext of the bid, which can only be accessed by the AES key owners, must encompass the bid itself and the IP of the bidder. Only one of the existing AES keys enables the TTP to decrypt this ciphertext successfully, as there is a different AES key per node. The TTP decrypts the ciphertext by using all the existing AES keys until one of them delivers a plaintext that encompasses one of the bidders IP, msking the system imune to impersonation attack as described in Section 4.2.1. At this point, the TTP validates the bid if the AES key used to attain the plaintext corresponds to the IP of the respective AES key owner. After validating the bid, the TTP uses the same AES key to broadcast a ciphertext of a content that comprises the IP of the bidder, which is the only node able to decrypt this message properly. The bidder thus obtains the confirmation of its order registration. The trading session is open when, at any instant between two time slots, at least one valid bid is decrypted by the TTP.

Hence, in every trading session a subset v of the n households forward a bid for purchasing or selling energy such that $v \in \{1, 2, \ldots, n\}$. After receiving the v offers, the TTP assembles the bit matrix $\mathbf{P} \in \mathbb{Z}^{n \times u}$, which is the plaintext of all offers. Given that there are only v bids, the TTP creates $(n - v)$ false offers and include all of them in $\mathbf{P}$, observing that the v true bids are inserted in random rows. Thus, with the key $\mathbf{K}_1$, the TTP computes

$$\mathbf{M}_1 = \mathbf{K}_1 \oplus \mathbf{P}, \tag{4.27}$$

where $\oplus$ stands for the XOR operator. Hence, in Step 3, the bids disclosure occurs when the TTP broadcasts the cipher matrix $\mathbf{M}_1$ to all nodes after the end of the time slot in which the bids came up, since matrix $\mathbf{P}$, which contains all raw data, cannot be published. When the nodes receive $\mathbf{M}_1$, they learn that a trading session has been created. The SM-owners can easily compute $\mathbf{P}$ since they have $\mathbf{K}_1$. They can distinguish the true bids from the false ones as the latter present inconsistencies in their bit structures, which infringe the rule of bids assemblage, as shown in details in Section 4.3.4.

Several problems can affect a trading session based solely on (4.27), since packet losses, collision or unfavorable SNIR conditions might prevent some nodes of receiving the cipher matrix $\mathbf{M}_1$. Therefore, a confirmation step is needed, which is provided by Step 4, with bids verification. Each node that receives $\mathbf{M}_1$ computes a second cipher matrix,

$$\mathbf{M}_2 = \mathbf{K}_2 \oplus \mathbf{P}, \tag{4.28}$$

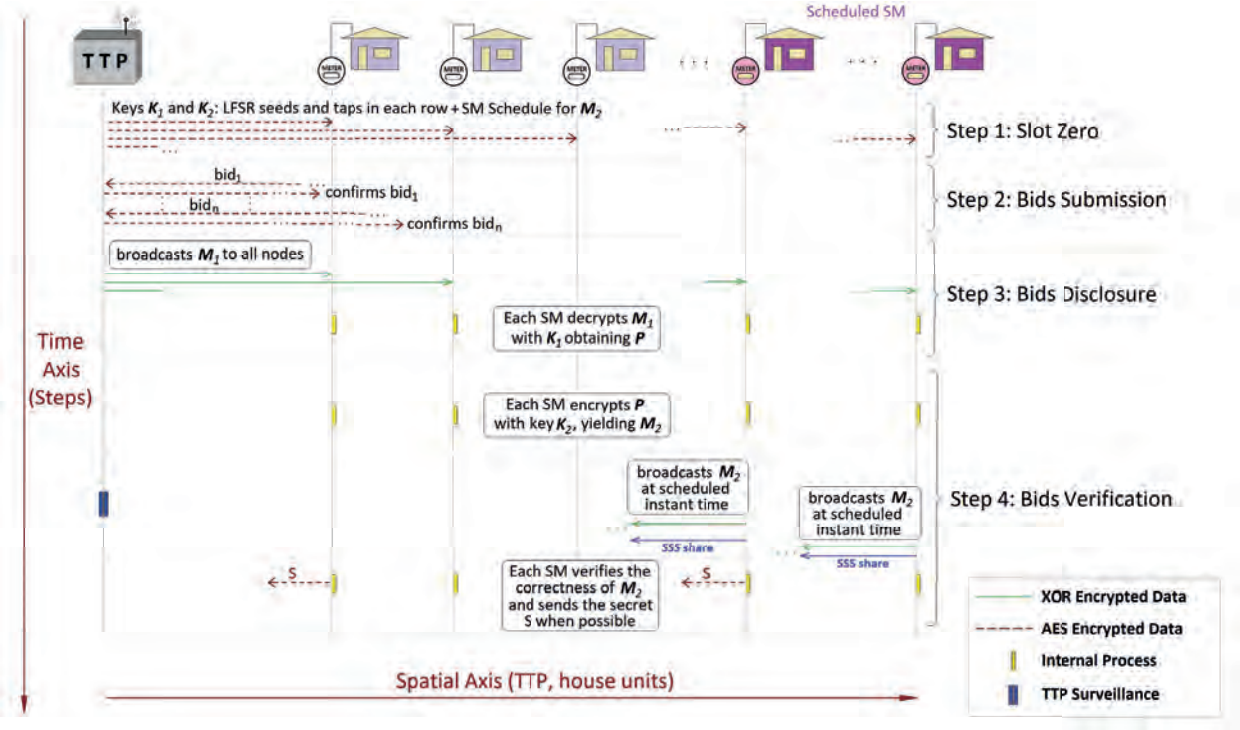

TTP
Scheduled SM
Keys K_1 and K_2: LFSR seeds and taps in each row + SM Schedule for M_2
Step 1: Slot Zero
bid$_1$
confirms bid$_1$
bid$_n$
confirms bid$_n$
Step 2: Bids Submission
broadcasts M_1 to all nodes
Step 3: Bids Disclosure
Time Axis (Steps)
Each SM decrypts M_1 with K_1 obtaining P
Each SM encrypts P with key K_2, yielding M_2
broadcasts M_2 at scheduled instant time
broadcasts M_2 at scheduled instant time
Step 4: Bids Verification
SSS share
SSS share
S
Each SM verifies the correctness of M_2 and sends the secret S when possible
S
XOR Encrypted Data
AES Encrypted Data
Internal Process
TTP Surveillance
Spatial Axis (TTP, house units)

Figure 4.7: Overall sequence of steps in the proposed framework, with the slot zero and the stages of bids submission, bids disclosure and bids verification. Note that only in the slot zero and in the bids submission stage the ciphertext is obtained via AES. The nodes designed to act as confirmation agents in a given time slot are depicted on the right side of the figure.

which is the matrix that is used as confirmation data. The matrix $\mathbf{K}_2 \in \mathbb{Z}^{n \times u}$ differs of $\mathbf{K}_1$ as the participants must prove to know the content of $\mathbf{P}$ without retransmitting $\mathbf{M}_1$. Since the nodes are not reputed trustworthy, they have to prove that they know the plaintext $\mathbf{P}$ by producing a different ciphertext, i.e., $\mathbf{M}_2$. Note that the increase in memory due to this second matrix key is irrelevant as the product nu bits reaches approximately 3 kB for each 100 house units. Note also that an external observer cannot learn how many offers are posed by the bidders as the size of $\mathbf{M}_1$ and $\mathbf{M}_2$ is always $n \times u$. Moreover, $\mathbf{M}_1$ and $\mathbf{M}_2$ are cipher matrices that do not deliver any useful information for an external observer that does not know $\mathbf{K}_1$ and $\mathbf{K}_2$. Recall that, in Step 1, along with the SM Schedule, the TTP also informs in which second of the slot the node must confirm the data. Thus, Step 4 consists of broadcasting $\mathbf{M}_2$ to all nodes of the NAN during the second specified by the TTP, addressing the requirement of using time-stamps to ensure the correctness of the sender identity.

Here, the role of the proposed CRT based perfect SSS shown in Section 4.3.2 in our framework is explained. With exception of the first node that broadcasts the cipher matrix $\mathbf{M}_2$, the nodes comprised in the SM Schedule for a given time slot can theoretically copy the content of the first broadcast $\mathbf{M}_2$. In this situation, they can emit an information without knowing its content, and hence they are not enforced to prove that they effectively know the content of matrix $\mathbf{P}$. Such situation must be prevented as the matrix $\mathbf{M}_2$ in this case loses its data confirmation capability.

As a solution, each node chosen as an agent by the SM Schedule in the given time slot also receives during the slot zero a share $\mathbf{E}_j$ as in (4.15) and broadcasts it along with matrix $\mathbf{M}_2$. The shares are selected by the TTP in order to set up an SSS (n, t), where $t \in \{2, 3, \ldots, n-1\}$ is the minimal number of agents in a successful coalition. The value of t is chosen as a function of the necessity of redundancy to the SSS, as at least one node can fail to broadcast $\mathbf{M}_2$ and $\mathbf{E}_j$. Hence, even when some of the n nodes fail to broadcast their shares, at least t shares must be provided, enabling other SM-nodes to calculate S. Hence, the subset of nodes of the NAN that receive at least t shares calculate the scalar S and transmit this number to the TTP encrypted via AES, as in the scheme of Fig. 4.8. In this example, all SM Scheduled SM are honest and only the nodes 1, 2, 5 and 6 calculate S properly. The rest of nodes must avoid sending any data to the TTP, by which the TTP learns that they did not succeed to receive at least t shares. If node 3, which did not receive correctly at least t shares from the confirmation agents, nevertheless sends an arbitrary and wrong value S' to the TTP, it includes node 3 in the Revocation List due to its false information.

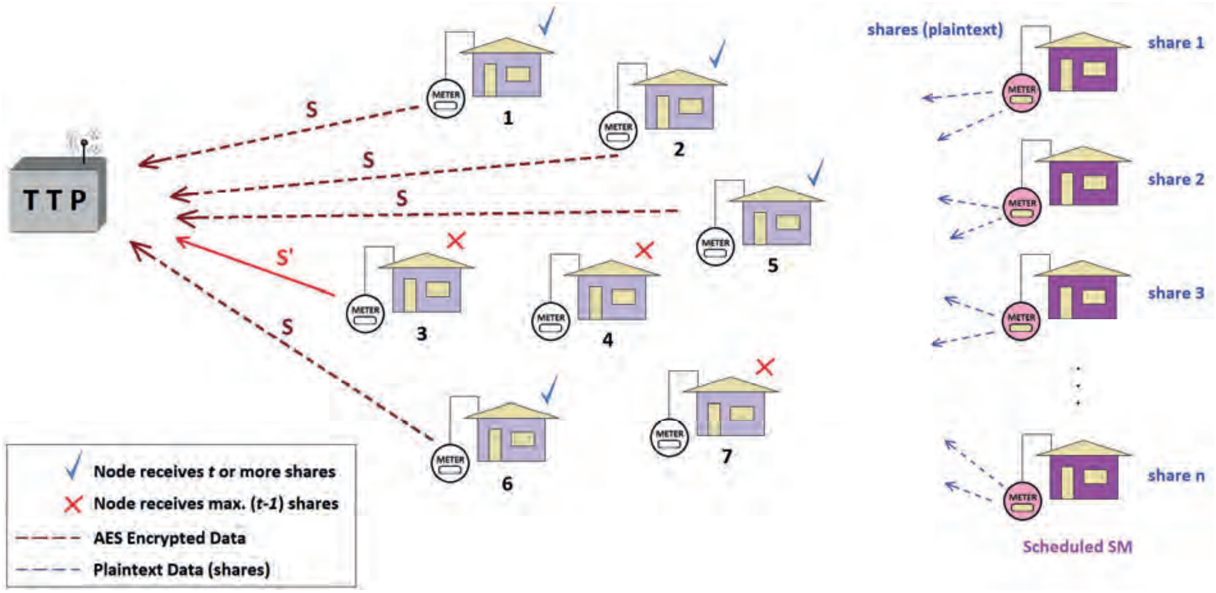


Figure 4.8: The subset of nodes of the NAN that receive at least t shares calculate the scalar S and transmit this number to the TTP encrypted via AES. In this example, only the nodes 1, 2, 5 and 6 calculate S properly. Although node 3 does not receive its share appropriately, it sends an arbitrary and wrong value S' to the TTP, which therefore includes node 3 in the Revocation List.

When the number of t broadcasts is not achieved, the TTP revocates the missing nodes, without canceling the trading session.

4.3.4 Trading System Framework

The prices transmitted by each bidder are $P_{r,i}$ in

$$P_l < P_{r,i} < P_u, \tag{4.29}$$

where r indicates the round, with $r \in \{1, 2\}$ as the proposed system has two rounds, i denotes the node that submits the offer, and $P_{r,i}$ is the actual unitary price of the kWh offered by the i-th node. Given that v out of the existing n SM-owners forward offers, $i \in \{1, 2, \ldots, v\}$. The prices $P_{1,i}$, $P_{2,i}$ are expressed in tenths of cents in order to reduce the probability of two offers to have exactly the same bit sequence. Two bit strings $p_{1,i}$ and $p_{2,i}$ express the values of $P_{1,i}$ and $P_{2,i}$. Likewise, the quantity Q_i of kWh in each offer is constrained to an interval $Q_l < Q_i < Q_u$, and thus Q_i is also denoted with an auxiliary bit string q_i. The bit strings $p_{1,i}$, $p_{2,i}$ and q_i comprise 10 bits each. Instead of 1024 possible values, for simplicity we reduce them to 999 values from 0000000001 up to 1111100111. Two bits complete the entire sequence, namely the type t_1 of the offer, with $t_1 = 0$ for sell and $t_1 = 1$ for purchase offers, and the status of the order in terms of time, with $t_2 = 0$ when the bid is valid only in the next trading session and $t_2 = 1$ to orders that stay valid throughout the day until a bid matches it. The length of 32 bits of the bid is complete with t_1, t_2, $p_{1,i}$, $p_{2,i}$ and q_i gathered sequentially, as in Fig. 4.9.

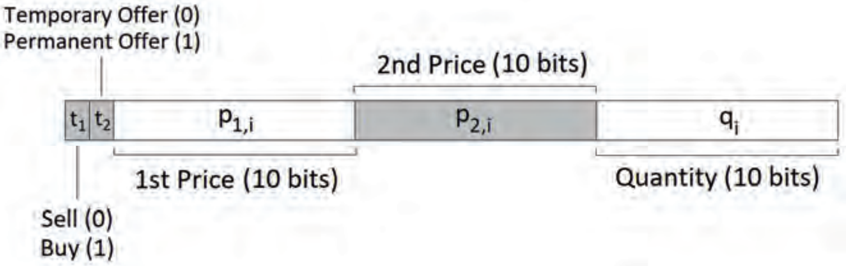


Figure 4.9: Bits of a bid string in terms of the bid content. The length of 32 bits comprise in this order t_1, t_2, $p_{1,i}$, $p_{2,i}$ and q_i.

Purchasers must offer prices $P_{1,i} < P_{2,i}$ while sellers must set $P_{1,i} > P_{2,i}$. All prices of $r = 1$ interact producing deals as long as purchasers' prices are equal or over sellers' prices. In sequel, the offers of $r = 2$ are combined in order to achieve further deals. The mechanism of prices interaction is beyond the scope of this chapter. A suggestion for the proposed system is the McAfee's Trade Reduction Mechanism (TRM) [18].

When $t_2 = 1$, price uncertainty is eliminated to all participants. This may cause that several bids in the next time slot have the same price values. In the case of this event, the time of arrival establish the priority of all the incoming bids. Therefore, if two bids have exactly the same prices, the first to arrive at the TTP has the preference of the match.

4.4 Results

In Section 4.4.1, we illustrate the performance of our framework in terms of security requirements and in Section 4.4.2 we undertake a comparison with the state-of-the-art privacy and trading systems. In Section 4.4.3 we show the gain in secret size in terms of the shares.

4.4.1 Security Analysis

4.4.1.1 Proposed framework in its steps

In the D-Y adversary model, an attacker is assumed unable to break cryptographic primitives [116]. Thus, the slot zero and the bids submission stage are considered data leakage proof as the required time for breaking AES keys is in the order of 10^{37} seconds [148]. Note that a node that forwards an offer in the bids submission stage includes its own IP in the plaintext which is encrypted, and the TTP confirms the arrival of any valid offer to its author. Furthermore, the exact instant of this message forward is uncertain over the entire duration of the time slot. A successful attack in this case demands continuous and explicitly intrusive actions. This collides with the notion in [181], according to which malicious data attacker are supposed to compromise as few data as possible in order to inject undetectable attacks with the lowest cost and effort. Hence, the probability of successful attacks in the slot zero and bids submission stage is assumed significantly unlikely.

In terms of data protection, it is useful to see that, for an external observer, the messages from the SMs and the TTP can be of any content as all nodes are supposed to exchange data with the TTP informing for instance the node availability, voltage measurements, etc. In Fig. 4.10, the adversary is represented by the red vehicle. It receives also dummy packets that can be exchanged between TTP and the SMs in order to thwart traffic analysis attacks. As a consequence, it is not possible for the adversary (spy) to infer the purpose of such messages. For internal attackers, i.e., those who possess an SM and are authenticated, it is equally not possible to devise the contents of the messages exchanged by the TTP and the other nodes, given that all the AES keys are different. As a consequence, in the slot zero and in the bids submission stages, the external observer can infer no useful information. In the bids disclosure stage, the bit matrix $\mathbf{M}_1$ is broadcast to all SMs of the NAN. The nodes are now more vulnerable to attacks on the data content since the adversary can see the TTP broadcasting data of nu bits, which is always the dimension of $\mathbf{M}_1$. The spy might decide to try to alter the ciphertext bits deliberately, however, in order to compromise the entirety of broadcast data, the attacker must access all links between the TTP and final nodes, a very problematic task if the network is sufficiently spread spatially. In the bids verification stage, the matrix $\mathbf{M}_2$ has, in comparison to the bit matrix $\mathbf{M}_1$, two additional protections: specified nodes are programmed to broadcast it and in specific time windows – in Fig. 4.10, house units 6 and 4 are the scheduled nodes and have the time windows $t1$ and $t2$, respectively –, as a case of time-stamps application. Only when the attacker knows in advance which are the nodes that are scheduled to broadcast $\mathbf{M}_2$ in the respective time slot, it can harm the broadcasts content. However, it is taken as impossible due to the lack of access to the AES keys between the TTP and the other nodes, which derives from the D-Y adversary model. Note that, due to the sequence of stages, the spy cannot infer how many orders are posed, nor of which type are those.

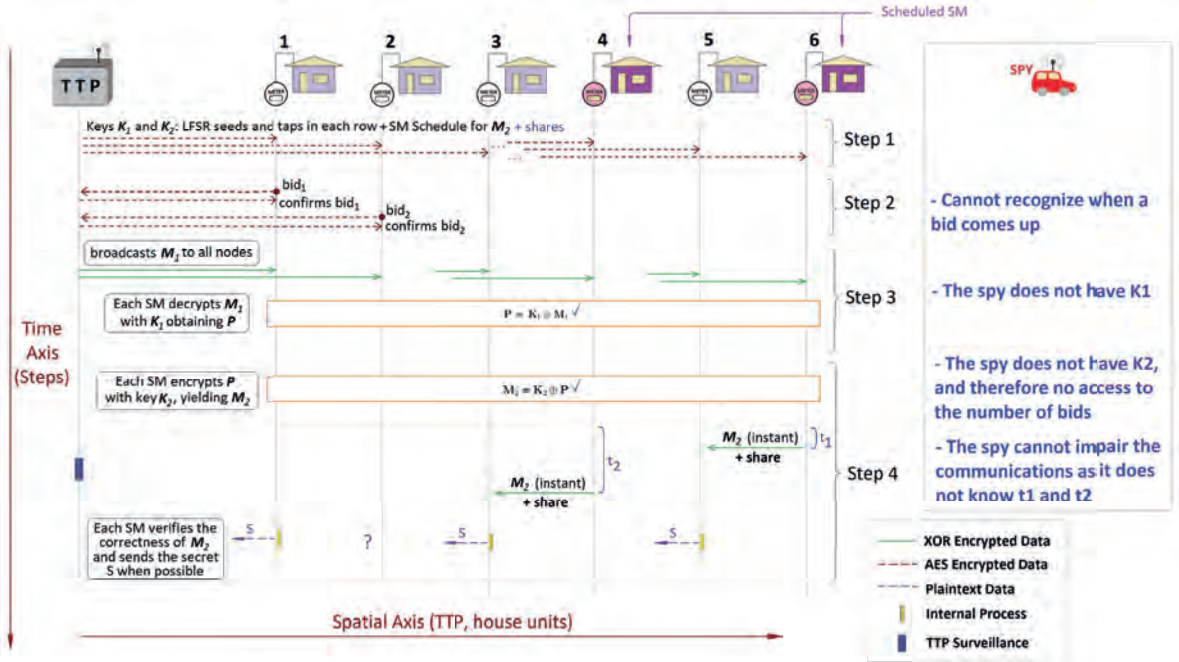

Scheduled SM
TTP
1
2
3
4
5
6
SPY
Keys K_1 and K_2: LFSR seeds and taps in each row + SM Schedule for M_2 + shares
Step 1
bid$_1$
confirms bid$_1$
bid$_2$
confirms bid$_2$
Step 2
- Cannot recognize when a bid comes up
broadcasts M_1 to all nodes
Step 3
- The spy does not have K1
Each SM decrypts M_1 with K_1 obtaining P
$P = K_1 \oplus M_1$ ✓
Time Axis (Steps)
Each SM encrypts P with key K_2, yielding M_2
$M_2 = K_2 \oplus P$ ✓
- The spy does not have K2, and therefore no access to the number of bids
M_2 (instant) + share
t_1
t_2
M_2 (instant) + share
Step 4
- The spy cannot impair the communications as it does not know t1 and t2
Each SM verifies the correctness of M_2 and sends the secret S when possible
S
?
S
S
Spatial Axis (TTP, house units)
XOR Encrypted Data
AES Encrypted Data
Plaintext Data
Internal Process
TTP Surveillance

Figure 4.10: Sequence of stages in the proposed framework from the standpoint of a spy. Due to the sequence of stages, the spy cannot infer how many orders are posed, nor of which type are those.

Suppose that the attacker is an internal node and it is scheduled to broadcast $\mathbf{M}_2$. It can broadcast a different matrix, say $\mathbf{M}_2'$, trying to induce all other nodes in error. However, such an attack cannot avoid the nodes of receiving the correct $\mathbf{M}_2$ from other scheduled nodes. Moreover, the TTP undertakes strict surveillance over the broadcast $\mathbf{M}_2$ contents. A node that broadcasts a false version $\mathbf{M}_2'$ is included in the Revocation List, even when it transmits the correct share $\mathbf{E}_j$. The protocol admits that a scheduled node does not broadcast $\mathbf{M}_2$, since it might not have received $\mathbf{M}_1$ due to package losses. In comparison to an external observer, which is not authenticated, an internal node can learn the number and the types of the bids, however not being able to link them with the respective authors.

4.4.1.2 Security and Privacy Features

We classify our framework under the requirements of Anonymity, Untraceability, No Impersonation, Unforgeability, Non-Repudiation, Verifiability, Non-Linkability, Linkability within a Single Bidding Round, Privacy, Forward Security, Authenticity and Integrity as in [138].

Anonymity is achieved when no unauthorized entity is able to identify the bidder during the bidding. Our system accomplishes this goal via AES keys between each bidder and the TTP. The constant size of matrices $\mathbf{M}_1$ and $\mathbf{M}_2$ avoid identification by means of traffic analysis, overall because such nodes must broadcast these cipher matrices even when they are not bidders. Untraceability is attained when the bid winner cannot be identified at the end of the bidding by untrusted entities. However, the winning bidder's legitimacy should be verifiable. Furthermore, no individual should be traceable during a bidding round. Our framework meets this requirement as the matrices $\mathbf{M}_1$ and $\mathbf{M}_2$ deliver and confirm all informations about the bids, all nodes can know the winner bid. Note that the winner identity is never accessible for any node, but completely recognized by the TTP.

When no one participates in the bidding with the identity of another bidder, the No Impersonation is achieved. Since all nodes are only admitted when their ciphertexts include their IP into the AES encrypted message, the TTP cannot accept false participants. Unforgeability is fulfilled when no one is able to falsify a valid bidding price. In the proposed framework, it derives from the No Impersonation requirement.

Non-repudiation is observed when the bidders cannot deny their bid after the winning bidder determination. After the TTP accepting the bid, the node that sent its offer cannot deny it, which makes part of matrix $\mathbf{P}$, since the TTP is assumed inviolable. Verifiability by its turn is achieved when anyone can verify the validity of the bids. This relies on the fact that matrix $\mathbf{M}_2$ reproduces the content of matrix $\mathbf{P}$ from what is informed by $\mathbf{M}_1$.

Non-linkability among various bidding rounds consists of not a participant being able to access results that enable a bidder to be identified in various bidding rounds. Due to the invariable sizes of $\mathbf{M}_1$ and $\mathbf{M}_2$, it is impossible even for authenticated nodes to know when a node submit bids. Linkability within a single bidding round is achieved when anyone can determine the number of times a bidder has bid, which in the case of the proposed framework is straightforward as it imposes that each node bids in maximum only once per time slot. In terms of Privacy, untrusted entities must not be able to link bids to individual consumers. Moreover, they must not be able to infer private information about individual consumers. It is applicable even for internal nodes.

Even if the current bidding key is compromised, no information about the previous keys should be leaked, which is the concept of Forward Security. The LFSR ensures that the XOR keys are not repeated from one time slot to another, making it infeasible to access the previously submitted bids. Authenticity and Integrity of all bid notifications occurs when all bids are be verifiable. The TTP verifies it with regards to each ciphertext received.

The Single Registration requirement consists of a bidder to be required to register in the system only once, and then can participate in all future bid sessions. In our framework, this is provided with the secrecy of the XOR keys with their seeds and taps. Using matrices $\mathbf{K}_1$ and $\mathbf{K}_2$, every node can forward intelligible bids, when it is automatically admitted as a participant. Easy Revocation is defined as the ease for the Registration Manager of revoking a bidder. In case of errors, the TTP in our system easily revocates a scheduled node that broadcasts false contents or a bidder that tries to submit an impossible bid – for instance, a purchasing bid in a household that has not generation assets. Incentive Allocation consists of the bid winner to be able to claim the incentive without revealing its identity and no other entity should be able to impersonate the winner. The winner can claim the incentive by messages between it and the TTP. Furthermore, in case of the same bid prices, the TTP chooses the winning bid with a temporal criterion, i.e., that is forwarded first, and can inform this fact to a bidder that claims a deal.

The requirements presented in Table IV.1 are now proven to be met in Table IV.4.

4.4.2 Comparison with the state-of-the-Art Systems

Comparing the proposed framework with the state-of-the-art ones, Table IV.5 shows the results in terms of privacy in the SG, with 1 for existing and 0 for non-existing features. As privacy enhancement method, [113], [146] and [138] use DLP-based encryption systems, which have a comparatively high computational complexity. In [182], nodes embedded with distributed controllers coordinate with neighboring peers in order to find the optimal operating data, such as instantaneous consumption power. In doing so, they transmit plaintext information, and therefore none of the common cryptographic techniques are applied. In [118], employing a large number of anonymous addresses is part of the solution for privacy attainment, contrary to our framework and to the discussion of Section 4.2.1. In [146], a privacy-preserving data aggregation (P2DA) scheme that aggregates electricity consumption

Table IV.4: Results of the Proposed System in terms of the Security Requirements of Table IV.1 [142]

Requirement	**What Each End-User Are Able To in Our Protocol**
Information Awareness	All participants are aware of the information that can be extracted from the data sets, which are bid prices and quantities
Data Control	Each bidder controls when it makes a bid. Identities of the bidders are disclosed only to the TTP due to the D-Y adversary model
Data-Information link	The false IP and the bid values are the only data that can be extracted from their forwarded messages
Data Aggregation	All nodes are aware that the data that can be extracted from various combinations of data sets are only the bid contents
Risk Awareness	The risk arising by misuse of a specific information is a higher capability of speculating about the bidder's financial status
Situation Dependency	Each node controls the data that it delivers, and knows the set of the possible destination nodes
Data Overview	The data released are the bid values, and they are released to a specific set of stakeholders, i.e., the other SM owners
Data Minimization	The bid data are designed to result in stringently little message sizes
Release Expiry	Expiry of bids are guaranteed by means of the LFSR based XOR keys in each row of $\mathbf{M}_1$ and $\mathbf{M}_2$

data focusing on consumer's privacy is presented, while [138] presents a demand-response demand bidding (DR-DB) protocol. The address of each SM is disclosed in [138] and in the proposed framework only to a central controller, while in [182] it is disclosed only to one-hop nodes. Interested parties can possibly be identified as such by a malicious observer in all compared systems by means of traffic analysis or address de-anonymization, i.e., an observer can identify the role of a participant by such attacks, while in our system it is completely avoided. The proposed framework is de-anonymization proof since, even when the IP of a node is identified, an observer cannot conclude if the node is a bidder or not. For instance, note that the SM-schedule establishes that a node that does not take part in the bids forward the matrix $\mathbf{M}_2$.

In Table IV.6, the proposed framework and the state-of-the-art schemes are compared in terms of pricing systems. Our framework allows for free-price formation and a variety of different auction systems, as for instance the suggested one [18]. The proposed framework also dismisses previous information about energy consumption profile. Such data are a requisite

Table IV.5: Comparison between the proposed framework and the state-of-the-art approaches in terms of privacy

Reference	[113]	[146]	[138]	[182]	[118]	Proposed
Cost of cryptography	High	High	High	None	Medium	Low
Disclosure of SM address	1	1	1	1	0	1
Dismissal of secure communication channel	1	1	0	0	1	1
ID de-anonymization proof	0	0	1	0	0	1
Absence of need for connection between each pair of nodes	0	1	1	1	0	1
Impossibility of interested parties identification	0	0	0	0	0	1

for the systems in [167], which characterizes the operation for the benchmark scenario of a DR market where the operator has full information of all DR-related parameters, such as the utility-function of the consumers, which is representative of their consumption profile and decision-making process. The proposed framework also allows the inclusion of storage elements, which are excluded from the systems such as [156], which develops an energy trading system of a community energy storage (CES) device for demand-side load management within a NAN. The energy users that have their own photovoltaic power generation are allowed to trade energy from their personal surplus with the grid and the CES device in a competitive game framework. Pricing freedom is fully guaranteed in [27] and [173]. In [183], a power market scheduling center (PMSC) is proposed, which manages all the energy providers and makes them provide a unified price to the subscribers, and the energy providers generate the optimal quantity of electricity to get the maximum utilities. In [167], aggregators provide DR services to the operator and guarantee a reduced electricity bill to the end-users, negotiating with both sides in order to maximize its own profit. In these examples, the final consumers are limited to play the role of price-takers.

4.4.3 Secret size gain in the proposed SSS

As shown in Fig. 4.6, the proposed SSS can be seen as k different CRT systems to solve, implying that, if a is the amount of data processed via Asmuth-Bloom's, the proposed SSS

Table IV.6: Comparison with the state-of-the-art approaches in terms of pricing systems

Reference	[183]	[149]	[167]	[184]	[154]	[158]	[156]	[160]	[27]	[152]	[173]	Prop.
Dismissal of consumption profiles	0	0	0	0	0	1	0	0	1	1	1	1
Admissibility of storage elements	0	1	0	1	0	1	0	1	0	1	1	1
Pricing freedom	0	0	0	0	0	0	0	0	1	0	1	1
Inclusion of a TTP	1	1	1	0	1	1	0	1	1	1	0	1
Competitive market	0	0	1	1	1	0	1	0	1	0	1	1

entails dealing with data of size ak. We determine the enhancement of the secret in terms of data size by means of the relationship

$$G = \frac{G_r}{G_d}, \tag{4.30}$$

where G_r is the gain in terms of secret range and G_d indicates the enlargement of data size that is needed to provide G_r. The gain G is therefore

$$G = \frac{k!}{M_0} \cdot \frac{a}{ak} = \frac{(k-1)!}{M_0}, \tag{4.31}$$

from which it is clear that the gain G with regards to the Asmuth-Bloom's SSS is potentially unlimited.

From Eqs. (4.2) and (4.21), note that M_0 and k can differ, as M_0 must be co-prime to any other M_j, $j \in \{1, 2, \dots, n\}$, whereas k does not need to cope with co-primality. In Tables IV.7, IV.8 and IV.9, different combinations of M_i and t show the gain G, respectively for $n = 4$, $n = 5$ and $n = 6$. Note that, as n increases, also the gain G is enhanced. Through the observation of Tables IV.7, IV.8 and IV.9, it is possible to perceive that the gain increases more rapidly with the moduli M_i when n is higher. The worst G among all the results is seen in Table IV.9, where the moduli $\{11, 13, 15, 17, 19, 23\}$ deliver a gain of 50% in the cases of $t = 3$ and $t = 4$. On the other hand, with the moduli $\{83, 87, 89, 91, 95, 97\}$ and $t = 2$ or $t = 5$, there are 6.04 $\times 10^{103}$ possible values of secrets per bits used.

Table IV.7: Examples of gain G given by (4.31) for secret sizes in terms of memory for sets $n = 4$

M_i	(n, t)	M_0 (4.2)	k (4.21)	G (4.31)
$\{11, 13, 17, 19\}$	(4,2)	7	7	1.03 $\times 10^{2}$
	(4,3)	7	7	1.03 $\times 10^{2}$
$\{23, 29, 31, 37\}$	(4,2)	18	18	1.98 $\times 10^{13}$
	(4,3)	18	18	1.98 $\times 10^{13}$
$\{47, 53, 59, 61\}$	(4,2)	40	40	5.10 $\times 10^{44}$
	(4,3)	40	40	5.10 $\times 10^{44}$
$\{79, 83, 87, 97\}$	(4,2)	67	67	8.13 $\times 10^{90}$
	(4,3)	67	67	8.13 $\times 10^{90}$

Table IV.8: Examples of gain G given by (4.31) for secret sizes in terms of memory for sets $n = 5$

M_i	(n, t)	M_0 (4.2)	k (4.21)	G (4.31)
$\{11, 13, 17, 19, 23\}$	(5,2)	6	6	20
	(5,3)	5	5	4.8
	(5,4)	6	6	20
$\{29, 31, 37, 41, 43\}$	(5,2)	20	20	6.08 $\times 10^{15}$
	(5,3)	18	18	1.98 $\times 10^{13}$
	(5,4)	20	20	6.08 $\times 10^{15}$
$\{47, 53, 59, 61, 67\}$	(5,2)	37	37	1.01 $\times 10^{40}$
	(5,3)	35	35	8.44 $\times 10^{36}$
	(5,4)	37	37	1.01 $\times 10^{40}$
$\{79, 83, 87, 91, 97\}$	(5,2)	67	67	8.13 $\times 10^{90}$
	(5,3)	64	64	3.10 $\times 10^{85}$
	(5,4)	67	67	8.13 $\times 10^{90}$

Moreover, we simulate the performance of our SSS in and evaluate it in terms of perfection. We select five moduli sets from $\{11, 13, 16, 17\}$, etc., until $\{11, 13, 16, 17, 19\}$, until $\{11, 13, 16, 17, 19, 23, 25, 27\}$, adding the last modulus at each set. For each of these groups, n is the number of moduli, t is the threshold or the minimal number of players that must provide a coalition, which is varied in the range $t \in \{2, \ldots, n - 1\}$, and k is calculated by means of (4.21). In each set, v_{-1} is defined as in (4.22) and the shares $\mathbf{E}_j$ of (4.15) are given. At this point, any of the other $(n - t + 1)$ moduli is chosen arbitrarily and its share is varied to reproduce all the $k!$ possible sequences of $\mathbf{s}$. A realization is considered successful when

Table IV.9: Examples of gain G given by (4.31) for secret sizes in terms of memory for sets $n = 6$

M_i	(n, t)	M_0 (4.2)	k (4.21)	G (4.31)
$\{11, 13, 15, 17, 19, 23\}$	(6,2)	4	6	30
	(6,3)	4	4	1.5
	(6,4)	4	4	1.5
	(6,5)	4	6	30
$\{25, 27, 29, 31, 37, 38\}$	(6,2)	17	17	1.23×10^{12}
	(6,3)	13	13	3.69×10^{7}
	(6,4)	13	13	3.69×10^{7}
	(6,5)	17	17	1.23×10^{12}
$\{37, 39, 41, 43, 44, 47\}$	(6,2)	31	33	8.49×10^{33}
	(6,3)	29	30	3.05×10^{29}
	(6,4)	29	30	3.05×10^{29}
	(6,5)	31	33	8.49×10^{33}
$\{83, 87, 89, 91, 95, 97\}$	(6,2)	74	74	6.04×10^{103}
	(6,3)	67	69	3.70×10^{94}
	(6,4)	29	30	3.70×10^{94}
	(6,5)	31	33	6.04×10^{103}

the owner of the t-th share can reproduce all possible values $S \in \{0, 1, \ldots, k! - 1\}$ by means of the Lehmer code by varying its share content in the face of the given first $(t - 1)$ shares. For each combination (n, t), 10^4 realizations are made. The successful ones are shown in Table IV.10.

Table IV.10: Percentage of successful realizations for testing the perfection of the proposed SSS - 10^4 realizations

Moduli set	t					
	2	3	4	5	6	7
$\{11, 13, 16, 17\}$	100%	100%	-	-	-	-
$\{11, 13, 16, 17, 19\}$	100%	100%	100%	-	-	-
$\{11, 13, 16, 17, 19, 23\}$	100%	100%	100%	100%	-	-
$\{11, 13, 16, 17, 19, 23, 25\}$	100%	100%	100%	100%	100%	-
$\{11, 13, 16, 17, 19, 23, 25, 27\}$	100%	100%	100%	100%	100%	100%

4.5 Conclusion

As an emerging cyber-physical system, the SG is attractive for enabling distributed energy control, allowing for high level of automation and security of the power system. Power is produced inside the boundaries of final user real states and exported to the company or other consumers. As a consequence, the old grid structures must be reformulated. With regards specifically to the prices in the new electrical systems, a key aspect is that the utility company usually sells the energy for a unitary price P_u to his consumers and buys back the energy of prosumers for a different unitary price, P_l, yielding the range for dealing prices between final users.

In the proposed framework, we provide an effective approach for privacy protection of prosumers in a NAN that takes into account the problem of self-interested players intending to trade energy. Our results show higher consistency when compared to the state-of-the-art models, specially in what concerns to privacy protection against IP de-anonymization, traffic analysis and impersonation attacks. In order to achieve these objectives, we use AES 128 bit associated with LFSR based XOR matrices, which have constant sizes, independently of the number of bidders. In doing so, our cryptographic framework has a considerable low computational cost.

Furthermore, although the Asmuth-Bloom's system offers a perfect SSS, its problem consists of the limited range of the secret values, i.e., the size in bits of the shares are usually greater than the size of the secret itself. Large ranges of secret values are desirable in an SSS because they provide more difficulty to an attacker that uses BFA to succeed in discovering the secret. Our CRT based SSS is not only is perfect but also presents a potentially unlimited gain in terms of secret range. Considering for instance a set of the highest 6 co-prime numbers under 100, the gain in secret range per bits used is leveraged in the order of 10^{103} in comparison with the Asmuth-Bloom's SSS. We integrate this SSS in one of the steps of our trading framework for a NAN, enhancing its security levels.

V

CONCLUSIONS

Smart Grid arises as a new form of the electric grid that is predominantly automated, bearing bi-directional data exchange features. In the task of providing data communication, signal processing applied techniques as well as data protection towards users privacy are expected to integrate the basis of the SG as a whole. In this sense, we exploit functionalities of the SG with the CRT that allow for data transmission and cryptographic applications.

In Chapter II, a novel method for estimating a real number using the CRT was presented. The method is based on an ME scheme that is optimized by means of a mapping vector that indicates in which parts of the dynamic range the search for the real number should occur. This mapping vector is assembled via Kronecker products of previously defined vectors. We also provide a version of the mapping vectors based on tensorial n-mode products, delivering in the end the same information of the original method. For its characteristics, it is suitable overall for CRT systems with few moduli, which in the case of WSN corresponds to low quantity of sensors. According to results tested over 10^5 realizations, in the case of equal variances, the proposed KME-CRT is consistently superior to the state-of-the-art methods CFR-CRT and MLE-CRT in terms of percentage of correct estimations. Our proposed technique enhances the probability of estimating an unknown number accurately even when the errors in the remainders surpass $1/4$ of the greatest common divisor of all moduli.

In Chapter III, we address the problem of predicting power consumption series data in Brasília-DF, Brazil. Candidate socio-economical variables and tariff time series are shown to be capable of improving the performance of the evaluated methods. The nonlinear transformations performed over weather variables are effective for improving the forecasting performance of the state space predictors. Furthermore, we propose a system for providing energy to the sensors in use. First, we undertake a measurement campaign over four places in Brasília and demonstrate the feasibility of RF energy harvesting systems in urban environments due to relatively high levels of dBm incidences. Furthermore, we show the superiority of rectennas for use in the RF systems is shown by means of simulations. In our system, the RF harvesters are to be used close to the deployed sensors, as they are installed normally in places of difficult access.

Finally, in Chapter IV we revise the main demands for SG data protection in accordance to the state of the art and provide an effective approach for privacy protection for prosumers in a NAN. Our approach takes into account the problem of self-interested players intending to trade energy with full privacy. We further present a perfect CRT based SSS that has a better ratio between the sizes of the shares and the secret than in the Asmuth-Bloom's SSS. Our SSS is not only perfect but also presents a potentially unlimited gain in terms of secret range. Considering for instance a set of the highest 6 co-prime numbers under 100, the gain in secret range per bits used is leveraged in the order of 10^{103} in comparison with the Asmuth-Bloom's SSS. We employ our SSS as a way of enhancing data security levels in the proposed framework, whose results show higher consistency when compared to the state-of-the-art models, specially in terms of privacy protection against IP de-anonymization and traffic analysis attack. In order to achieve these goals, we use AES 128 bit associated with LFSR based XOR matrices of constant sizes, independently of the number of bidders, which makes our cryptographic framework of a very low computational cost.

5.1 Recommendations for Future Research

For future works concerning CRT, we envisage that errors that follow distributions different of the Gaussian one should be investigated. Tensor based mapping vector routines of Algorithms 3 and 5 are still under development and are also a matter of concern for future studies. In terms of CRT techniques, applying the mapping vector to the Multi-Stage Robust CRT and studies involving CRT in a probabilistic way, as in the case of unrestricted errors of [59], is supposed to be very promising.

From the standpoint of SG, there is a relative lack of real experiments that reproduce the frameworks and protocols proposed in the revised literature. In the case of the framework proposed in Chapter IV, SMs are expected to undergo problems of package loss and other flaws in communication links. In this sense, real experiments with hardware communicating in a NAN are seen as a key factor for knowing the real performance and the potential problems that might occur in SG applications.

A

THE LEHMER CODE

This appendix reviews the Lehmer Code, which is used to assign a random variable to a permutation in a bijective relationship.

The Lehmer Code allows for achieving a bijective relationship between an integer-valued number S and a sequence of k elements in a vector $\mathbf{s}$. Since $k!$ permutations of the k elements are possible, the range of values of S is $\{0, 1, \ldots, (k! - 1)\}$. As a consequence, the value of $k!$ delimits the range for the values of S. The Lehmer Code [180] shown in Algorithm 6 provides the conversion of a permutation in $\mathbf{s}$ to the number S. The input of the algorithm is a sequence of numbers $n_i \in \mathbf{s}$, for $i \in \{1, 2, \ldots, k\}$, which is compared to the original sequence of increasing values $n_1 < n_2 < \cdots < n_k$ in order to produce the output S. The code assigns weights to the displacements of the first $(k-1)$ values in terms of their expected increasing order.

Algorithm 6 Lehmer Code from the Permutation to the Number

```
procedure LEHMER CODE 1 ($\mathbf{s}$)
    $\mathbf{a} \leftarrow \text{zeros}(1, k-1)$
    $\mathbf{b} \leftarrow \begin{bmatrix}(k-1)! & (k-2)! & \ldots & 1!\end{bmatrix}$
    for $i = 1 : k-1$ do
        $\begin{bmatrix}c & j\end{bmatrix} \leftarrow \min(\mathbf{s})$ % $c$ receives the minimum value $n_i$ and $j$ receives the position of $n_i$ in $P$
        $\mathbf{a}(i) \leftarrow j - 1$
        $\mathbf{s} \leftarrow \mathbf{s}(\mathbf{s} \neq n_i)$ % Removing $n_i$ from the vector $\mathbf{s}$
    $S \leftarrow \mathbf{a}\mathbf{b}^T$
return $S$
```

As examples, the vector $\mathbf{s} = \begin{bmatrix}1 & 2 & 3 & 4\end{bmatrix}$ is the reference vector for increasing order of the elements, without any permutation, so that $S = 0$. If, for instance, $\mathbf{s} = \begin{bmatrix}3 & 4 & 2 & 1\end{bmatrix}$, there are three displacements of the number 1 and two displacements of the number 2, which in comparison to the increasing order $\begin{bmatrix}1 & 2 & 3 & 4\end{bmatrix}$ yields $S = 3 \times 3! + 2 \times 2! + 0 \times 1! = 22$. Note that the entry of highest value in $\mathbf{s}$ is not relevant for calculating S.

In order to do the reverse, i.e., obtaining the sequence from a given S, as a mere inversion of the logic of Algorithm 6, one must calculate the folding integers $\mathbf{a}(i)$ that deliver $S = \sum_{i=1}^{k-1} \mathbf{a}(i) \times (k-i)!$. As an illustration, when $S = 11$, the folding integers are $\mathbf{a}(1) = 1$, $\mathbf{a}(2) = 2$, and $\mathbf{a}(3) = 1$. As a consequence, beginning with the sequence in the vector $\mathbf{s} = \begin{bmatrix}1 & 2 & 3 & 4\end{bmatrix}$, the number 3 is displaced one slot rightwards, yielding $\mathbf{s} = \begin{bmatrix}1 & 2 & 4 & 3\end{bmatrix}$; then, 2 is displaced two slots, yielding $\mathbf{s} = \begin{bmatrix}1 & 4 & 3 & 2\end{bmatrix}$ and 1 is displaced one slot, resulting in the final permutation $\mathbf{s} = \begin{bmatrix}4 & 1 & 3 & 2\end{bmatrix}$.

B

PROOF OF THEOREM 1 FROM CHAPTER IV

Let a sequence of symbols $\{a_1, a_2, \ldots, a_k\}$ be repeated k times laterally, as if in a Kronecker product. If the palindrome of the original sequence, i.e., $\{a_k, a_{k-1}, \ldots, a_1\}$ must be selected from the series of symbols blocks, it is easy to prove that only one symbol per group can be chosen in order to produce it, as suggested by Fig. 2.1.

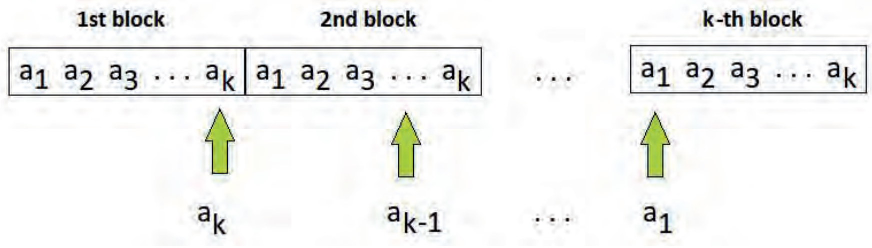


Figure 2.1: With the k blocks of the original sequence of symbols, only one symbol per symbol group can be chosen in order to produce the palindrome.

The highest possible cardinality of a_1 in the scheme of Fig. 2.1 with the $(t-1)$ highest moduli, i.e., a vector with length v_{-1} as defined in (4.22). It is clear that a_1 can theoretically appear in the $(k\prod_{i=1}^{t-1} M_{n-i+1} - k)$-th position inside the range D,

$$k\prod_{i=1}^{t-1} M_{n-i+1} - k \leq D. \tag{B.1}$$

Since in any case $M_1 \prod_{i=1}^{t-1} M_{n-i+1} > \prod_{i=1}^{t} M_i$, $M_1 \prod_{i=1}^{t-1} M_{n-i+1} > D$. As a consequence,

$k < M_1$, and (B.1) can be reduced to

$$k \prod_{i=1}^{t-1} M_{n-i+1} \leq D. \tag{B.2}$$

However, note that $D/(\prod_{i=1}^{t-1} M_{n-i+1})$ is never an integer due to the co-primality between all M_i. Hence, we can rewrite (B.2) as

$$k \prod_{i=1}^{t-1} M_{n-i+1} < D, \tag{B.3}$$

which completes the proof. □

BIBLIOGRAPHY

Publications as First or Co-Author

[1] J. Milanezi Jr., J. P. C.L. da Costa, F. Römer, R. K. Miranda, M. A.M. Marinho, and G. Del Galdo. M-estimator based chinese remainder theorem with few remainders using a kroenecker product based mapping vector. *Digital Signal Processing, Elsevier, Vol. 87, pp. 60-74, April 2019.*

[2] J. Milanezi Jr., J. P. C.L. da Costa, and G. Del Galdo. A chinese remainder theorem based perfect secret sharing scheme with enhanced secret range values using tensorial operations. *13th International Conference on Signal Processing and Communication Systems, ICSPCS'2019, Australia (submitted).*

[3] L. D. X. Ribeiro, J. Milanezi Jr., J. P. C. L. da Costa, W. F. Giozza, R. K. Miranda, and M. V. Vieira. Pca-kalman based load forecasting of electric power demand. *IEEE Symposium on Signal Processing and Information Technology (ISSPIT), 2016.*

[4] J. Milanezi Jr., J. P. C.L. da Costa, and E. P. de Freitas. Improved radiofrequency energy harvesting based on a rectenna array system and its feasibility evaluation in urban environments. *International Conference on Renewable Energy Research and Application (ICRERA), pp. 561-565, 2014.*

[5] J. Milanezi Jr., J. P. C.L. da Costa, E. P. de Freitas, G. Del Galdo, W. Felber, R. S. Ferreira Junior, and R. K. Miranda. Radiofrequency energy harversting system based on a rectenna array in the urban environment of brasilia, brazil. *International Symposium "New Tendencies of Developing Fundamental and Applied Physics: Problems, Achievements, Prospectives", Nov. 2016, Tashkent, Uzbekistan.*

[6] J. Milanezi Jr., J. P. C.L. da Costa, R. K. Miranda, R. S. Ferreira Jr., G. Del Galdo, E. P. de Freitas, and W. Felber. Radiofrequency energy harvesting system based on a rectenna array in urban environments. *1st International Conference on Signals and Systems 2017 (ICSigSys 2017), Bali.*.

[7] J. Milanezi Jr., J. P. C.L. da Costa, E. P. de Freitas, J. A. A. Gomes, and R. Schmitt. Sustainable electric energy microgeneration system based on electric eels. *International Conference on Renewable Energy Research and Application (ICRERA), pp. 648-652, 2014.*.

[8] G. C. Ornelas, M. V. V. da Silva, J. Milanezi Jr., J. P. C. L. da Costa, F. E. G. de Deus, and G. Del Galdo. First step towards a smart grid communication architecture for the brazilian federal district (df). *4th IFAC Symposium on Telematics Applications, Porto Alegre, Brazil, Nov. 2016.*.

[9] J. Milanezi Jr., J. P. C.L. da Costa, and M. V. Vieira. Iterative inverter module with island-mode detection. *patent under analysis in the Technological Development Support Center (CDT) at UnB.*.

[10] J. Milanezi Jr., J. P. C.L. da Costa, A. Arancibia, L. Weichenberger, R. T. de Sousa Junior, and G. Del Galdo. Data security and trading framework for smart grids in neighborhood area networks. *Transactions on Smart Grids, IEEE, 2019 (submitted).*

References by Other Authors

[11] W. Wang and Z. Lu. Cyber security in the smart grid: Survey and challenges. *Computer Networks, Elsevier, Issue 57, pp. 1344-1371, 2013.*.

[12] V. Tudor, M. Almgren, and M. Papatriantafilou. The influence of dataset characteristics on privacy preserving methods in the advanced metering infrastructure. *Computers & Security, Elsevier, 76, pp. 178-196, 2018.*.

[13] M. Mihaylov, S. Jurado, N. Avellana, K. Van Moffaert, I. M. de Abril, and A. Nowé. Nrgcoin: Virtual currency for trading of renewable energy in smart grids. *11th International Conference on the European Energy Market (EEM14), pp. 1-5, 2014, DOI: 10.1109/EEM.2014.6861213.*

[14] S. Sikdar and K. Rudie. Microgrid level competitive market using dynamic matching. *Proceedings of the Electrical Power & Energy Conference (EPEC), Halifax, NS, Canada, pp. 1-6, 2013.*

[15] REN21. Renewables global status report. *Renewable Energy Policy Network for the 21st Century, Technical Report, 2017.*

[16] R. Li, Q. Wu, and S. S. Oren. Distribution locational marginal pricing for optimal electric vehicle charging management. *IEEE Transactions on Power Systems, Vol. 29, Issue 1, pp. 203-211, 2014.*

[17] K. Wang, Z. Ouyang, R. Krishnan, L. Shu, and L. He. A game theory-based energy management system using price elasticity for smart grids. *IEEE Transactions on Industrial Informatics, Vol. 11, Issue 6, pp. 1607-1616, 2015.*

[18] R. P. McAfee. A dominant strategy double auction. *Journal of Economic Theory, Vol. 56, No. 2, pp. 434-450, 1992.*

[19] The International Energy Agency IEA. Smart grids - tracking clean energy progress. *retrieved from https://www.iea.org/tcep/energyintegration/smartgrids/ on June 22nd 2019.*

[20] X.-G. Xia. On estimation of multiple frequencies in undersampled complex valued waveforms. *IEEE Transactions on Signal Processing, vol. 47, no. 12, December 1999.*

[21] G. Hill. The benefits of undersampling. *Electron. Des., pp. 69-79, 1994.*

[22] H. Xiao and G. Xiao. Notes on crt-based robust frequency estimation. *Signal Processing, Elsevier, Issue, 133, pp. 13–17, 2017.*

[23] P. P. Vaidyanathan and P. Pal. Sparse sensing with co-prime samplers and arrays. *IEEE Transactions on Signal Processing, vol. 59, no. 2, 2011.*

[24] Piya Pal and P. P. Vaidyanathan. Coprime sampling and the music algorithm. *IEEE Digital Signal Processing and Signal Processing Education Meeting (DSP/SPE), pp. 289-294, 2011.*

[25] T. Moon, H. W. Choi, N. Tzou, and A. Chatterjee. Wideband sparse signal acquisition with dual-rate time-interleaved undersampling hardware and multicoset signal reconstruction algorithms. *IEEE Transactions on Signal Processing, Vol.63, Issue 24, pp. 6486-6497, 2015.*

[26] J. Wu, K. Ota, M. Dong, J. Li, and H. Wang. Big data analysis-based security situational awareness for smart grid. *IEEE Transactions on Big Data, Vol. 4, No. 3, pp.408-417, 2018.*

[27] T. Chen and W. Su. Indirect customer-to-customer energy trading with reinforcement learning. *IEEE Transactions on Smart Grid (Early Access), July 2018, DOI: 10.1109/TSG.2018.2857449.*

[28] T. Chen, Q. Alsafasfeh, H. Pourbabak, and W. Su. The next-generation u.s. retail electricity market with customers and prosumers — a bibliographical survey. *Energies 2018, 11, 8; DOI:10.3390/en11010008.*

[29] F. Knirsch, G. Eibl, and D. Engel. Error-resilient masking approaches for privacy preserving data aggregation. *IEEE Transactions on Smart Grid, Vol. 9, No. 4, pp. 3351-3361, 2018.*

[30] H. Chun, K. Ren, and W. Jiang. Privacy-preserving power usage and supply control in smart grid. *Computers & Security, Elsevier, 77 (2018), 709-719.*

[31] R. Mohammadi, H. R. Mashhadi, and M. Shahidehpour. Market-based customer reliability provision in distribution systems based on game theory: A bi-level optimization approach. *IEEE Transactions on Smart Grid, (Early Access), 2018.*

[32] J.-P. Sheu and J.-J. Lin. A multi-radio rendezvous algorithm based on chinese remainder theorem in heterogeneous cognitive radio networks. *IEEE Transactions on Mobile Computing, Issue 99, pp. 1-1, 2018.*

[33] L. Xiao and X.-G. Xia. Robust polynomial reconstruction via chinese remainder theorem in the presence of small degree residue errors. *IEEE Transactions on Circuits and Systems II: Express Briefs, Issue 99, pp. 1-1, 2017.*

[34] G. C. Cardarilli, L. Di Nunzio, R. Fazzolari, L. Gerardi, G. Campolo M. Re, and D. Cascone. A new electric encoder position estimator based on the chinese remainder theorem for the cmg performance improvements. *2017 IEEE International Symposium on Circuits and Systems (ISCAS), pp. 1-4, 2017.*

[35] X. Li, W. Wang, W. Zhang, and Y. Cao. Phase-detection-based range estimation with robust chinese remainder theorem. *IEEE Transactions on Vehicular Technology, Vol, 65, Issue 12, pp. 10132-10137, 2016.*

[36] Z. Y.-Peng, L. Xia, and W. Qiang. Asymmetric cryptography algorithm with chinese remainder theorem. *2011 IEEE 3rd International Conference on Communication Software and Networks, pp. 450-454, 2011.*

[37] X. Shen, Y. Jia, J. Wang, and L. Zhang. New families of balanced quaternary sequences of even period with three-level optimal autocorrelation. *IEEE Communications Letters, Vol. 21, Issue 10, pp. 2146-2149, 2017.*

[38] N. Singh, A. N. Tentu, A. Basit, and V. Ch. Venkaiah. Sequential secret sharing scheme based on chinese remainder theorem. *2016 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), pp. 1-6, 2016.*

[39] L. Harn and M. Fuyou. Multilevel threshold secret sharing based on the chinese remainder theorem. *Information Processing Letters, Elsevier, Issue 114, pp. 504-509, 2014.*

[40] X. Huang, Y. Han, Z. Yan, H. Xian, and W. Lu. Resolution doubled co-prime spectral analyzers for removing spurious peaks. *IEEE Transactions on Signal processing, pp. 2489-2498, vol. 64, No. 10, 2016.*

[41] W. Wang and X.-G. Xia. A closed-form robust chinese remainder theorem and its performance analysis. *IEEE Transactions on Signal Processing, vol. 58, Issue 11, pp. 5655-5666, 2010.*

[42] W. Wang, X. Li, Wei Wang, and X.-G. Xia. Maximum likelihood estimation based robust chinese remainder theorem for real numbers and its fast algorithm. *IEEE Transactions on Signal Processing, Vol. 63, Issue 13, pp. 3317-3331, 2015.*

[43] X.-G. Xia X. Li, W. Wang, and Wei Wang. A robust generalized chinese remainder theorem for two integers. *IEEE Transactions on Information Theory, vol. 62, no. 12, pp. 7491-7504, December 2016.*

[44] S. Bi and W. J. Grossl. The mixed-radix chinese remainder theorem and its applications to residue comparison. *IEEE Transactions on Computers, Vol. 57, Issue 12, pp. 1624-1632, 2008.*

[45] L. Stetson and G. Stark. Peak electrical demands of individuals and groups of rural residential customers. *IEEE Transactions on Industry Applications, vol. 24, no. 5, pp. 772–776, Sept/Oct 1988.*

[46] M. Draper. Modeling weather effects on electric energy sales. *IEEE Proceedings of the Southeastcon, vol. 1, pp. 133–136, 1990.*

[47] J. Milanezi Junior. Spatio-temporal prediction of electric power systems including emergent renewable energy sources. *Master's thesis, University of Brasilia, Brazil, march 2014.*

[48] K. Schaeke. On the kroenecker product. *August 2004, CiteSeer.*

[49] Z. Jiang, J. Wang, Q. Song, and Z. Zhou. A closed-form robust chinese remainder theorem based multibaseline phase unwrapping. *International Conference on Circuits, Devices and Systems (ICCDS), pp. 115-119, 2017.*

[50] L. Xiao, X.-G. Xia, and W. Wang. Multi-stage robust chinese remainder theorem. *IEEE Transactions on Signal Processing, Vol. 62, Issue 18, pp. 4772-4785, 2014.*

[51] A. Koochakzadeh and P. Pal. On the robustness of co-prime sampling. *23rd European Signal Processing Conference (EUSIPCO), pp. 2825-2829, 2015.*

[52] L. Ding, Y. Ye, G. Ye, X. Wang, and Y. Zhu. Bistatic synthetic aperture radar with undersampling for terahertz 2-d near-field imaging. *IEEE Transactions on Terahertz Science and Technology, Vol. 8, Issue 2, pp. 174-182, 2018.*

[53] Wei Wang, X. Li, X.-G. Xia, and W. Wang. The largest dynamic range of a generalized chinese remainder theorem for two integers. *IEEE Signal Processing Letters, vol. 22, no. 2, pp. 254-258, February 2015.*

[54] H. Liao and X.-G. Xia. A sharpened dynamic range of a generalized chinese remainder theorem for multiple integers. *IEEE Transactions on Information Theory, vol. 53, no. 1, pp. 428-433, January 2007.*

[55] K. Kaya and A. A. Selçuk. Sharing dss by the chinese remainder theorem. *Journal of Computational and Applied Mathematics, Elsevier, Issue 259, pp. 495-502, 2014.*

[56] P. W. Adi, Y. P. Astuti, and E. R. Subhiyakto. Imperceptible image watermarking based on chinese remainder theorem over the edges. *IEEE 4th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI), pp. 1-5, September 2017.*

[57] S. K. Singh, V. P. Gopi, and P. Palanisamy. Image security using des and rns with reversible watermarking. *2014 International Conference on Electronics and Communication Systems (ICECS), pp. 1-5.*

[58] S. Iftene. General secret sharing based on the chinese remainder theorem with applications in e-voting. *Electronic Notes in Theoretical Computer Science, Elsevier, Issue 186, pp. 67-84, 2007.*

[59] L. Xiao and X.-G. Xia. Error correction in polynomial remainder codes with non-pairwise coprime moduli and robust chinese remainder theorem for polynomials. *IEEE Transactions on Communications, Vol. 63, No. 3, pp. 605-616, 2015.*

[60] X. Li, H. Liang, and X.-G. Xia. A robust chinese remainder theorem with its applications in frequency estimation from undersampled waveforms. *IEEE Transactions on Signal Processing, Vol. 57, Issue 11, pp. 4314-4322, 2009.*

[61] H. Liang, H. Zhang, and N. Jia. A generalized robust chinese remainder theorem for multiple numbers and its application in multiple frequency estimation with low sampling rates. *IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC), pp. 1-4, 2011.*

[62] H. Liang, X. Li, and X.-G. Xia. Adaptive frequency estimation with low sampling rates based on robust chinese remainder theorem and iir notch filter. *4th IEEE Conference on Industrial Electronics and Applications, pp. 2999-3004, 2009.*

[63] W. Wang, X. Li, and X.-G. Xia. An ml estimation based robust chinese remainder theorem for reals. *2015 IEEE China Summit and International Conference on Signal and Information Processing (ChinaSIP), pp. 363-367, 2015.*

[64] L. Xiao and X.-G. Xia. Frequency determination from truly sub-nyquist samplers based on robust chinese remainder theorem. *Signal Processing, vol. 150, pp. 248-258, 2018.*

[65] G. Xu. On solving a generalized chinese remainder theorem in the presence of remainder errors. *https://arxiv.org/abs/1409.0121, 2014.*

[66] L. Xiao and X.-G. Xia. A new robust chinese remainder theorem with improved performance in frequency estimation from undersampled waveforms. *Signal Processing, Elsevier, vol. 117, pp. 242-246, 2015.*

[67] J. Grossschadl. The chinese remainder theorem and its application in a high-speed rsa crypto chip. *Proceedings 16th Annual Computer Security Applications Conference (ACSAC'00), August 2002.*

[68] A. J. Menezes, P. C. van Oorschot, and S. A. Vanstone. Handbook of applied cryptography. *Crc Press, 1996.*

[69] T. G. Kolda and B. W. Bader. Tensor decompositions and applications. *Sandia National Laboratories, SAND2007-6702, 2007.*

[70] L. Xiao, X.-G. Xia, and H. Huo. Towards robustness in residue number systems. *IEEE Transactions on Signal Processing, Vol. 65, Issue 6, pp. 1497-1510, 2016.*

[71] H. Liang, X. Li, and X.-G. Xia. Adaptive frequency estimation with low sampling rates based on robust chinese remainder theorem and iir notch filter. *4th IEEE Conference on Industrial Electronics and Applications, pp. 2999-3004, 2009.*

[72] E. Lin and L. Monte. Joint frequency and angle of arrival estimation using the chinese remainder theorem. *IEEE Radar Conference (RadarConf), pp. 1547-1551, 2017.*

[73] G. Zhou and X.-G. Xia. Multiple frequency detection in undersampled complex-valued waveforms with close multiple frequencies. *Electronics Letters, Vol. 33, Issue 15, pp. 1294-1295, 1997.*

[74] J. Forrest. The effects of weather on power-system operation. *Journal of the Institution of Electrical Engineers, vol. 93, no. 64, pp. 161–163, April 1946.*

[75] S. Majithia, S. J. Watson, and C. L. Hor. Analyzing the impact of weather variables on monthly electricity demand. *IEEE Transactions on Power Systems, vol. 20, pp. 2078–2085, November 2005.*

[76] R. S. Elias, L. Fang, and M. I. M. Wahab. Electricity load forecasting based on weather variables and seasonalities: a neural network approach. *8th International Conference on Service Systems and Service Management, pp. 1-6, 2011.*

[77] K. Li and N. Tai. Research and application of climatic sensitive short - term load forecasting. *Power & Energy Society General Meeting, pp. 1-5, July 2015.*

[78] J. Asafu-Adjaye. The relationship between energy consumption, energy prices and economic growth: time series evidence from asian developing countries. *Energy Ec, vol. 22, no. 6, pp. 615-625, December 2000.*

[79] G. Altinaya and E. Karagol. Electricity consumption and economic growth: evidence from turkey. *Energy Economics, vol. 27, no. 6, pp. 849–856, November 2005.*

[80] L. Hernandez, C. Baladron, J. M. Aguiar, B. Carro, A. J. Sanchez Esguevillas, J. Lloret, and J. Massana. survey on electric power demand forecasting: Future trends in smart grids, microgrids and smart buildings. *IEEE Communications Surveys and Tutorials, Vol. 16, No. 3, Third Quarter 2014.*

[81] M. Ernoult and R. Mattatia. Short-term load forecasting: new developments at the e. d. f. *Proceedings of the Eighth Power Systems Computation Conference, no. First Edition, August 1984.*

[82] J. H. Park, Y. M. Park, and K. Y. Lee. Composite modeling for adaptive short-term load forecasting. *IEEE Transactions on Power Systems, Vol. 6, No. 2, pp. 450-457, May 1991.*

[83] Y.-Lu and H.-F. Shi. The hourly load forecasting based on linear gaussian state space model. *Proceedings of International Conference on Machine Learning and Cybernetics (ICMLC). IEEE, july 2012, pp. 741 – 747.*

[84] Y. Wang, D. Gu, J. Xu, and J. Li. Back propagation neural network for short-term electricity load forecasting with weather features. *International Conference on Computational Intelligence and Natural Computing, vol. 1, pp. 58-61, 2009.*

[85] R. Faragher. Understanding the basis of the kalman filter via a simple and intuitive derivation. *IEEE Signal Processing Magazine, August 2012, DOI: 10.1109/MSP.2012.2203621.*

[86] B. D. O. Anderson and J. B. Moore. Optimal filtering. *New York: Dover, 2005.*

[87] T.-L. Tien. A new grey prediction model fgm(1, 1). *Mathematical and Computer Modelling, vol. 49, pp. 1416–1426, 2009.*

[88] L. A. D. de Luca, C. M. de Oliveira, and R. S. Wazlawick. Load behavior changes after holidays on thursdays. *Computational Science and Engineering Workshops, pp. 101–106, 2008.*

[89] O. L'eonard S. Borguet. Coupling principal component analysis and kalman filtering algorithms for on-line aircraft engine diagnostics. *Proceedings of the 18th International Society of Air-Breathing Engines (ISABE) Conference, 2007-1275, September 2007.*

[90] X. Dang, Y. Huang, Z. Hao, and Xiong Si. Pca-kalman: device-free indoor humanbehavior detection with commodity wi-fi. *EURASIP Journal on Wireless Communications and Networking (2018) 2018:214, https://doi.org/10.1186/s13638-018-1230-2.*

[91] UK Meteorological Office. Method for calculating heating and cooling degree days. *The Weekly Weather Report, February 1928. [Online]. Available: http://ukclimateprojections.metoffice.gov.uk/22715.*

[92] D. J. Sailor. Relating residential and commercial sector electricity loads to climate - evaluating state level sensitivities and vulnerabilities. *Energy, vol. 26, pp. 645–657, July 2001.*

[93] T. Defraeyea, B. Blockenb, and J. Carmeliet. Convective heat transfer coefficients for exterior building surfaces: Existing correlations and cfd modelling. *Energy Conversion and Management, vol. 52, no. 1, pp. 512-522, January 2011.*

[94] R. E. Kalman. New approach to linear filtering and prediction problems. *Transactions of the ASME - Journal of Basic Engineering, vol. 82 (Series D), pp. 35–45, 1960.*

[95] L. D. X. Ribeiro. Adaptive kalman based forecasting for electric load and distributed generation. *Master Thesis, 659/2017 DM PPGEE, University of Brasilia, April 2017.*

[96] A. Al-Khayari, H. Al-Khayari, S. Al-Nabhani, M. M. Bait-Suwailam, and Z. Nadir. Design of an enhanced rf energy harvesting system for wireless sensors. *2013 IEEE GCC Conference and exhibition, November 17-20, 2013, pp. 479-482.*

[97] K. K. A. Devi, Norashidah Md.Din, C .K. Chakrabarty, and S. Sadasivam. Design of an rf-dc conversion circuit for energy harvesting. *IEEE International Conference on Electronics Design, Systems and Applications (ICEDSA), 2012, pp. 156-161.*

[98] J. Paradiso and T. Starner. Energy scavenging for mobile and wireless electronics. *Pervasive Computing, Vol.4, No.1, pp. 18-27, 2005.*

[99] H. Jabbar, Y. S. Song, and T. T. Jeong. Rf energy harvesting system and circuits for charging of mobile devices. *IEEE Transactions on Consumer Electronics, Vol. 56, No. 1, February 2010, pp. 247-253.*

[100] H. Nishimoto, Y. Kawahara, and T. Asami. Prototype implementation of ambient rf energy harvesting wireless sensor networks. *IEEE Sensors 2010 Conference, pp. 1282-1287.*

[101] C. Mikeka, H. Arai, A. Georgiadis, and A. Collado. Dtv band micropower rf energy-harvesting circuit architecture and performance analysis. *2011 IEEE International Conference on RFID-Technologies and Applications, pp. 561-567, DOI: 10.1109/RFID-TA.2011.6068601.*

[102] M. Ali, L. Albasha, and N. Qaddoumi. Rf energy harvesting for autonomous wireless sensor networks. *8th International Conference on Design and Technology of Integrated Systems in Nanoscale Era (DTIS), IEEE, 2013, pp. 78-81.*

[103] H. J. Visser and R. J. M. Vullers. Rf energy harvesting and transport for wireless sensor network applications: principles and requirements. *Proceedings of the IEEE, Vol. 101, No. 6, June 2013, pp. 1410-1423, DOI: 10.1109/JPROC.2013.2250891.*

[104] H. J. Visser, A. C. F. Reniers, and J. A. C. Theeuwes. Ambient rf energy scavenging: Gsm and wlan power density measurements. *2008 38th European Microwave Conference, pp. 721-724, DOI: 10.1109/EUMC.2008.4751554.*

[105] R. Zhang and C. K. Ho. Mimo broadcasting for simultaneous wireless information and power transfer. *IEEE Transaction on Wireless Communications, Vol. 12, No. 5, May 2013.*

[106] H. Lee, S.-R. Lee, K.-J. Lee, H.-B. Kong, and I. Lee. Optimal beamforming designs for wireless information and power transfer in miso interference channels. *IEEE Transactions on Wireless Communications, vol. 14, no. 9, pp. 4810–4821, September 2015.*

[107] T. Ungan and L. M. Reindl. Concept for harvesting low ambient rf-sources for microsystems. *January 2007.*

[108] H. Yan, J. G. Macias Montero, A. Akhnoukh, L. C. N. de Vreede, and J. N. Burghartz. An integration scheme for rf power harvesting. *Proc. STW Annual Workshop on Semiconductor Advances for Future Electronics and Sensors, Veldhoven, Netherlands, 2005, pp. 64-66.*

[109] J. Wang, Y. Fu, and L. Dong. Modeling of uhf voltage multiplier for radio-triggered wake up circuit. *International Journal of Circuit Theory and Application, 2010, DOI: 10.1002/cta.692.*

[110] N. Barreca, H. M. Saraiva, P. T. Gouveia, J. Tavares, L. M. Borges, F. J. Velez, C. Loss, R. Salvado, P. Pinho, R. Gonçalves, N. B. Carvalho, R. Chavez-Santiago, and I. Balasingham. Antennas and circuits for ambient rf energy harvesting in wireless body area networks. *IEEE 24th International Symposium on Personal, Indoor and Mobile Radio Communications: Fundamentals and PHY Track, 2013, pp.532-537.*

[111] M. Piñuela, P. D. Mitcheson, and S. Lucyszyn. Ambient rf energy harvesting in urban and semi-urban environments. *IEEE Transactions on Microwave Theory and Techniques, Vol. 61, No. 7, July 2013, pp. 2715-2726.*

[112] T. Baleshan. Analysis of distributed beamforming in cooperative communications networks with smart antenna nodes. *PhD dissertation, Queensland University of Technology, October 2014.*

[113] G. Liang, S. R. Weller, F. Luo, J. Zhao, and Z. Y. Dong. Distributed blockchain-based data protection framework for modern power systems against cyber attacks. *IEEE Transactions on Smart Grid, (Early Access), 2018, DOI: 10.1109/TSG.2018.2819663.*

[114] W. Jia, H. Zhu, Z. Cao, X. Dong, and C. Xiao. Human-factor-aware privacy-preserving aggregation in smart grid. *IEEE Systems Journal, Vol. 8, No. 2, pp. 598-607, 2014.*

[115] G. Eibl and D. Engel. Influence of data granularity on smart meter privacy. *IEEE Transactions on Smart Grid, Vol. 6 , Issue 2, pp. 930-939, 2015, DOI: 10.1109/TSG.2014.2376613.*

[116] A. J. Paverd, A. P. Martin, and I. Brown. Security and privacy in smart grid demand response systems. *in: Smart Grid Security — Second International Work- shop, SmartGridSec 2014, Munich, Germany, February 26, 2014, Revised Selected Papers, 2014, pp. 1-15, DOI: 10.1007/978-3-319-10329-7-1.*

[117] S. Kessler, C. M. Flath, and K. Böhm. Allocative and strategic effects of privacy enhancement in smart grids. *Information Systems, 53 (2015) 170-181, Elsevier.*

[118] K. Kvaternik, A. Laszka, M. Walker, D. Schmidt, M. Sturm, M. Lehofer, and A. Dubey. Privacy-preserving platform for transactive energy systems. *arXiv:1709.09597v2 [cs.DC], 30 Jan 2018.*

[119] Z. Guan, G. Si, X. Du, P. Liu, Z. Zhang, and Z. Zhou. Protecting user privacy based on secret sharing with fault tolerance for big data in smart grid. *2017 IEEE International Conference on Communications (ICC), 2017, DOI: 10.1109/ICC.2017.7997371.*

[120] T. Liu, Y. Liu, Y. Mao, Y. Sun, X. Guan, W. Gong, and S. Xiao. A dynamic secret-based encryption scheme for smart grid wireless communication. *IEEE Transactions on Smart Grid Vol. 5, Issue 3, pp. 1175-1182, 2014, DOI: 10.1109/TSG.2013.2264537.*

[121] T. Shiobara, P. Palensky, and H. Nishi. Effective metering data aggregation for smart grid communication infrastructure. *41st Annual Conference of the IEEE Industrial Electronics Society (IECON 2015), pp. 2136-2140, DOI: 10.1109/IECON.2015.7392417.*

[122] S. Salamatian, W. Huleihel, A. Beirami, A. Cohen, and M. Médard. Why botnets work: Distributed brute-force attacks need no synchronization. *IEEE Transactions on Information Forensics and Security, Vol. 14, No. 9, September 2019, DOI: 10.1109/TIFS.2019.2895955.*

[123] M. M. Najafabadi, T. M. Khoshgoftaar, C. Calvert, and C. Kemp. Detection of ssh brute force attacks using aggregated netflow data. *2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA), DOI: 10.1109/ICMLA.2015.20.*

[124] A. M. Ahmed, S. H. Ahmed, and O. H. Ahmed. Enhancing 3d-playfair algorithm to support all the existing characters and increase the resistanceto brute force and frequency analysis attacks. *2017 International Conference on Current Research in Computer Science and Information Technology (ICCIT), DOI: 10.1109/CRCSIT.2017.7965538.*

[125] K. Kaya, A. A. Selçuk, and Z. Tezcan. Threshold cryptography based on asmuth-bloom secret sharing. *Computer and Information Sciences (ISCIS), pp. 935-942, 2006..*

[126] A. Shamir. How to share a secret. *Commun. ACM 1979, 22, pp. 612-613.*

[127] C. Asmuth and J. Bloom. A modular approach to key safeguarding. *IEEE Transactions on Information Theory, 1983, IT-29, pp. 208-210..*

[128] M. Mignotte. How to share a secret. *in Proceedings of Cryptography-Proceedings of the Workshop on Cryptography, LNCS 149, pp. 371–375, Springer, 1983..*

[129] M. Quisquater, B. Preneel, and J. Vandewalle. On the security of the threshold scheme based on the chinese remainder theorem. *Proceedings of the 5th International Workshop on Practice and Theory in Public Key Cryptosystems, PKC 2002, pp.199-210, Paris, France, 2002, DOI: 10.1007/3-540-45664-3_14.*

[130] National Institute of Standards and Technology (NIST). The smart grid interoperability panel – cyber security working group, guidelines for smart grid cyber security. *NISTIR 7628 (2010).*

[131] E. Buchmann, K. Böhm, T. Burghardt, and S. Kessler. Re-identification of smart meter data. *Pers. Ubiquitous Comput. 17(4) (2013) 653-662..*

[132] Y. Gong, Y. Cai, Y. Guo, and Y. Fang. A privacy-preserving scheme for incentive-based demand response in the smart grid. *IEEE Transactions on Smart Grid, Vol. 7, No. 3, pp. 1304-1313, 2016..*

[133] W.-C. Ku and S.-T. Chang. Impersonation attack on a dynamic id-based remote user authentication scheme using smart cards. *IEICE Trans. Commun., Vol. E88–B, No.5, May 2005.*

[134] L. Tamilselvan and V. Sankaranarayanan. Prevention of impersonation attack in wireless mobile ad hoc networks. *IJCSNS International Journal of Computer Science and Network Security, VOL.7 No.3, March 2007.*

[135] A. R. Metke and R. L. Ekl. Security technology for smart grid networks. *IEEE Transactions on Smart Grid, Vol. 1, Issue 1, pp. 99-107, 2010..*

[136] S. Iyer. Cyber security for smart grid, cryptography, and privacy'. *International Journal of Digital Multimedia Broadcasting, Vol. 2011, Article ID 372020, pp. 1-8, 2011..*

[137] A. Barletta, C. Callegari, S. Giordano, M. Pagano, and G. Procissi. Privacy preserving smart grid communications by verifiable secret key sharing. *2015 Intl. Conference on Computing and Network Communications (CoCoNet'15), pp. 199-204, 2015.*

[138] M. S. Rahman, A. Basu, S. Kiyomoto, and M. Z. A Bhuiyan. Privacy-friendly secure bidding for smart grid demand-response. *Information Sciences 379 (2017) 229–240, Elsevier..*

[139] P. Samadi, H. Mohsenian-Rad, R. Schober, V. Wong, and J. Jatskevich. Optimal real-time pricing algorithm based on utility maximization for smart grid. *Proc. of IEEE Conf. on Smart Grid Communications, 2010.*

[140] H. Farhangi. The path of the smart grid. *IEEE Power and Energy Magazine, Vol. 8, No. 1, 2010..*

[141] S. A. Yadav, S. R. Kumar, S. Sharma, and A. Singh. A review of possibilities and solutions of cyber attacks in smart grids. *1st International Conference on Innovation and Challenges in Cyber Security (ICICCS 2016), pp. 60-63..*

[142] C. Neureiter, G. Eibl, A. Veichtlbauer, and D. Engel. Towards a framework for engineering smart-grid-specific privacy requirements. *IECON 2013 - 39th Annual Conference of the IEEE Industrial Electronics Society, DOI: 10.1109/IECON.2013.6699912.*

[143] K. Tazi, F. Abdi, and M. F. Abbou. Review on cyber-physical security of the smart grid: Attacks and defense mechanisms. *2015 3rd International Renewable and Sustainable Energy Conference (IRSEC), April 2016, DOI: 10.1109/IRSEC.2015.7455127.*

[144] S. Shapsough, F. Qatan, R. Aburukba, F. Aloul, and A. R. Al Ali. Smart grid cyber security: Challenges and solutions. *International Conference on Smart Grid and Clean Energy Technologies (ICSGCE), pp. 170-175, 2015.*

[145] C. Richardson, N. Race, and P. Smith. A privacy preserving approach to energy theft detection in smart grids. *2016 IEEE International Smart Cities Conference (ISC2), pp. 1-4, 2016.*

[146] D. He, N. Kumar, S. Zeadally, A. Vinel, and L. T. Yang. Efficient and privacy-preserving data aggregation scheme for smart grid against internal adversaries. *IEEE Transactions on Smart Grid, Vol. 8, Issue 5, pp. 2411-2419, 2017.*

[147] K. Iwamura and K. Tokita. Fast secure computation based on a secret sharing scheme for $n < 2k - 1$. *4th International Conference on Mobile and Secure Services (MobiSecServ), pp. 1-5, 2018.*

[148] O. G. Abood, M. A. Elsadd, and S. K. Guirguis. Investigation of cryptography algorithms used for security and privacy protection in smart grid. *2017 Nineteenth International Middle East Power Systems Conference (MEPCON), pp. 644-649, 2017.*

[149] W. Wei, F. Liu, and S. Mei. Energy pricing and dispatch for smart grid retailers under demand response and market price uncertainty. *IEEE Transactions on Smart Grid, Vol. 6, No. 3, pp. 1364-1374, 2015.*

[150] K. Ma, G. Hu, and C. J. Spanos. Distributed energy consumption control via real-time pricing feedback in smart grid. *IEEE Transactions on Control Systems Technology, Vol. 22, Issue 5, pp. 1907-1914, 2014.*

[151] P. Gope and B. Sikdar. An efficient privacy-preserving dynamic pricing-based billing scheme for smart grids. *2018 IEEE Conference on Communications and Network Security (CNS), pp. 1-2, 2018.*

[152] W. Tushar, C. Yuen, D. B. Smith, and H. Vincent Poor. Price discrimination for energy trading in smart grid: A game theoretic approach. *IEEE Transactions on Smart Grid, Vol. 8, No. 4, pp. 1790-1801, 2017.*

[153] N. Liu, X. Yu, C. Wang, C. Li, L. Ma, and J. Lei. Energy-sharing model with price-based demand response for microgrids of peer-to-peer prosumers. *IEEE Transactions on Power Systems, Vol. 32, Issue 5, pp. 3569-3583, 2017.*

[154] P. Jacquot, O. Beaude, S. Gaubert, and N. Oudjane. Analysis and implementation of an hourly billing mechanism for demand response management. *IEEE Transactions on Smart Grid (Early Access), 2018.*

[155] J. Lee, J. Guo, J. K. Choi, and M. Zukerman. Distributed energy trading in microgrids: A game-theoretic model and its equilibrium analysis. *IEEE Transactions on Industrial Electronics, Vol. 62, No. 6, pp. 3524-3533, 2015.*

[156] C. P. Mediwaththe, E. R. Stephens, D. B. Smith, and A. Mahanti. A dynamic game for electricity load management in neighborhood area networks. *IEEE Transactions on Smart Grid, Vol. 7, No. 3, pp. 1329-1336, 2016.*

[157] L. P. Qian, Y. Wu, Y. J. Zhang, and J. Huang. Demand response management via real-time electricity price control in smart grids. *Smart Grid — Networking, Data Management, and Business Models, CRC Press, Taylor & Francis Group, 2016.*

[158] J. Zhu, M. Z. Q. Chen, Z. Zuo, and B. Du. A new pricing scheme for controlling energy storage devices in smart grid. *2014 American Control Conference, pp. 2912-2917, 2014.*

[159] T. Voice, P. Vytelingum, S. Ramchurn, A. Rogers, and N. Jennings. Decentralised control of micro-storage in the smart grid. *AAAI Conference on Artificial Intelligence, 2011.*

[160] R. Verschae, T. Kato, and T. Matsuyama. Energy management in prosumer communities: A coordinated approach. *Energies, Vol. 9, No. 7, 2016.*

[161] F. Guo, C. Wen, and Z. Li. Distributed optimal energy scheduling based on a novel pd pricing feedback strategy in smart grid. *2015 IEEE 10th Conference on Industrial Electronics and Applications (ICIEA), pp. 208-213, 2015, DOI: 10.1109/I-CIEA.2015.7334112.*

[162] M. Ye, G. Hu, and C. J. Spanos. Decentralized optimal load scheduling using extremum seeking-based optimization. *UC Berkeley: Center for Research in Energy Systems Transformation (CREST), 2014, retrieved from https://escholarship.org/uc/item/85d412c7.*

[163] Y. Wu, X. Tan, L. Qian, D. H. K. Tsang, W.-Z. Song, and L. Yu. Optimal pricing and energy scheduling for hybrid energy trading market in future smart grid. *IEEE Transactions on Industrial Informatics, Vol. 11, No. 6, pp. 1585-1596, 2015.*

[164] Y. S. Foo. Eddy, H. B. Gooi, and S. X. Chen. Multi-agent system for distributed management of microgrids. *IEEE Transactions on Power Systems, Vol. 30, Issue 1, pp. 24-34, 2015.*

[165] A.-H. Mohsenian-Rad, V. W. S. Wong, J. Jatskevich, R. Schober, and A. Leon-Garcia. Autonomous demand-side management based on game-theoretic energy consumption scheduling for the future smart grid. *IEEE Transactions on Smart Grids, Vol. 1, No. 3, pp. 320-331, 2010.*

[166] P. Chakraborty, E. Baeyens, P. P. Khargonekar, K. Poolla, and P. Varaiya. Analysis of solar energy aggregation under various billing mechanisms. *IEEE Transactions on Smart Grid (Early Access), 2018.*

[167] L. Gkatzikis, I. Koutsopoulos, and T. Salonidis. The role of aggregators in smart grid demand response markets. *IEEE Journal on Selected Areas in Communications, Vol. 31, No. 7, pp. 1247-1257, Jul. 2013.*

[168] Y. Wang, W. Saad, Z. Han, H. V. Poor, and T. Başar. A game-theoretic approach to energy trading in the smart grid. *arXiv:1310.1814 [cs.GT], October 2013.*

[169] P. Vytelingum, T. D. Voice, S. D. Ramchurn, A. Rogers, and N. R. Jennings. Agent-based micro-storage management for the smart grid. *Proc. of the Ninth International Conference on Autonomous Agents and Multiagent Systems, pp. 39-46, 2010.*

[170] M. Kim, S. Parkt, J. K. Choi, and J. Lee. Energy independence of energy trading system in microgrid. *2017 IEEE Innovative Smart Grid Technologies — Asia (ISGT-Asia), 2017, DOI: 10.1109/ISGT-Asia.2017.8378441.*

[171] D. Ilic, P. G. da Silva, S. Karnouskos, and M. Griesemer. An energy market for trading electricity in smart grid neighbourhoods. *2012 6th IEEE International Conference on Digital Ecosystems and Technologies (DEST), July 2012, DOI: 10.1109/DEST.2012.6227918.*

[172] I. S. Bayram, M. Z. Shakir, M. Abdallah, and K. Qaraqe. A survey on energy trading in smart grid. *2014 IEEE Global Conference on Signal and Information Processing (GlobalSIP), February 2015, DOI: 10.1109/GlobalSIP.2014.7032118.*

[173] N. Yaagoubi and H. T. Mouftah. A distributed game theoretic approach to energy trading in the smart grid. *2015 IEEE Electrical Power and Energy Conference (EPEC), pp. 203-208, 2015, DOI: 10.1109/EPEC.2015.7379950.*

[174] R. Ghorani, M. Fotuhi-Firuzabad, and M. Moeini-Aghtaie. Optimal bidding strategy of transactive agents in local energy markets. *IEEE Transactions on Smart Grid (Early Access), 2018.*

[175] O. Zinaman, A. Aznar, C. Linvill, N. Darghouth, T. Dubbeling, and E. Bianco. Grid-connected distributed generation: Compensation mechanism basics. *National Renewable Energy Laboratory, Oct. 2017.*

[176] E. McKenna and M. Thomson. Photovoltaic metering configurations, feed-in tariffs and the variable effective electricity prices that result. *IET Renew. Power Gener., Vol. 7, No. 3, pp. 235-245, May 2013.*

[177] C. Hu, X. Liao, and X. Cheng. Verifiable multi-secret sharing based on lfsr sequences. *Theoretical Computer Science, Elsevier, Issue 445, pp. 52-62, 2012.*

[178] M. George and P. Alfke. Linear feedback shift registers in virtex devices. *Xilinx XAPP210 (v1.3), pp. 1-5, April 30, 2007.*

[179] J. P. C. L. da Costa and L. Weichenberger. Reliable intra-system communication for wireless battery management systems. *Deutsches Patent- und Markenamt (DPMA).*

[180] A. T. Magis and N. D. Price. The top-scoring 'n' algorithm: a generalized relative expression classification method from small numbers of biomolecules. *BMC Bioinformatics · September 2012, DOI: 10.1186/1471-2105-13-227.*

[181] J. Hao, R. J. Piechocki, D. Kaleshi, W. H. Chin, and Z. Fan. Optimal malicious attack construction and robust detection in smart grid cyber security analysis. *2014 IEEE International Conference on Smart Grid Communications (SmartGridComm), 2015, DOI: 10.1109/SmartGridComm.2014.7007752.*

[182] J. Duan and M.-Y. Chow. A resilient consensus-based distributed energy management algorithm against data integrity attacks. *IEEE Transactions on Smart Grid (Early Access), 2018, DOI: 10.1109/TSG.2018.2867106.*

[183] Q. Tang, K. Yang, D. Zhou, Y. Luo, and F. Yu. A real-time dynamic pricing algorithm for smart grid with unstable energy providers and malicious users. *IEEE Internet of Things Journal, Vol. 3, No. 4, pp. 554-562, 2016.*

[184] Y.-C. Hung and G. Michailidis. Modeling and optimization of time-of-use electricity pricing systems. *IEEE Transactions on Smart Grid (Early Access), June 2018, DOI: 10.1109/TSG.2018.2850326.*

ERKLÄRUNG

Ich versichere, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Die aus anderen Quellen direkt oder indirekt übernommenen Daten und Konzepte sind unter Angabe der Quelle gekennzeichnet.

Weitere Personen waren an der inhaltlich-materiellen Erstellung der vorliegenden Arbeit nicht beteiligt. Insbesondere habe ich hierfür nicht die entgeltliche Hilfe von Vermittlungs bzw. Beratungsdiensten (Promotionsberater oder anderer Personen) in Anspruch genommen. Niemand hat von mir unmittelbar oder mittelbar geldwerte Leistungen für Arbeiten erhalten, die im Zusammenhang mit dem Inhalt der vorgelegten Dissertation stehen.

Die Arbeit wurde bisher weder im In- noch im Ausland in gleicher oder ähnlicher Form einer Prüfungsbehörde vorgelegt. Ich bin darauf hingewiesen worden, dass die Unrichtigkeit der vorstehenden Erklärung als Täuschungsversuch bewertet wird und gemäß § 7 Abs. 10 der Promotionsordnungden Abbruch des Promotionsverfahrens zur Folge hat.

Brasília, der 16. 08. 2019

Jayme Milanezi Junior

About the Author

Jayme Milanezi Junior

He graduated in Electrical Engineering from the Military Institute of Engineering (IME-2004), holds a Master's degree (2014) and a PhD (2019) in Electrical Engineering from the University of Brasília. He completed part of his PhD as a Sandwich PhD student at the Technical University of Ilmenau, Germany, and has published papers as the first author. His research spans diverse fields such as digital signal processing, energy recycling, applications in cryptography, and privacy protection of users in Smart Grids. He currently works at the Secretary of Innovation and Energy Transition of the National Electric Energy Agency (STE/ANEEL) and also as a Professor at the University of the Federal District (UnDF) teaching Computer Networks II.

AYA EDITORA
2024