

# Técnicas de Processamento de Sinais e Telecomunicações

**Henrique Ajuz Holzmann  
(Organizador)**

**Henrique Ajuz Holzmann**

(Organizador)

# Técnicas de Processamento de Sinais e Telecomunicações

Atena Editora  
2019

2019 by Atena Editora  
Copyright © Atena Editora  
Copyright do Texto © 2019 Os Autores  
Copyright da Edição © 2019 Atena Editora  
Editora Executiva: Prof<sup>a</sup> Dr<sup>a</sup> Antonella Carvalho de Oliveira  
Diagramação: Karine de Lima  
Edição de Arte: Lorena Prestes  
Revisão: Os Autores

O conteúdo dos artigos e seus dados em sua forma, correção e confiabilidade são de responsabilidade exclusiva dos autores. Permitido o download da obra e o compartilhamento desde que sejam atribuídos créditos aos autores, mas sem a possibilidade de alterá-la de nenhuma forma ou utilizá-la para fins comerciais.

### **Conselho Editorial**

#### **Ciências Humanas e Sociais Aplicadas**

Prof. Dr. Álvaro Augusto de Borba Barreto – Universidade Federal de Pelotas  
Prof. Dr. Antonio Carlos Frasson – Universidade Tecnológica Federal do Paraná  
Prof. Dr. Antonio Isidro-Filho – Universidade de Brasília  
Prof. Dr. Constantino Ribeiro de Oliveira Junior – Universidade Estadual de Ponta Grossa  
Prof<sup>a</sup> Dr<sup>a</sup> Cristina Gaio – Universidade de Lisboa  
Prof. Dr. Deyvison de Lima Oliveira – Universidade Federal de Rondônia  
Prof. Dr. Gilmei Fleck – Universidade Estadual do Oeste do Paraná  
Prof<sup>a</sup> Dr<sup>a</sup> Ivone Goulart Lopes – Istituto Internazionele delle Figlie de Maria Ausiliatrice  
Prof. Dr. Julio Candido de Meirelles Junior – Universidade Federal Fluminense  
Prof<sup>a</sup> Dr<sup>a</sup> Lina Maria Gonçalves – Universidade Federal do Tocantins  
Prof<sup>a</sup> Dr<sup>a</sup> Natiéli Piovesan – Instituto Federal do Rio Grande do Norte  
Prof<sup>a</sup> Dr<sup>a</sup> Paola Andressa Scortegagna – Universidade Estadual de Ponta Grossa  
Prof. Dr. Urandi João Rodrigues Junior – Universidade Federal do Oeste do Pará  
Prof<sup>a</sup> Dr<sup>a</sup> Vanessa Bordin Viera – Universidade Federal de Campina Grande  
Prof. Dr. Willian Douglas Guilherme – Universidade Federal do Tocantins

#### **Ciências Agrárias e Multidisciplinar**

Prof. Dr. Alan Mario Zuffo – Universidade Federal de Mato Grosso do Sul  
Prof. Dr. Alexandre Igor Azevedo Pereira – Instituto Federal Goiano  
Prof<sup>a</sup> Dr<sup>a</sup> Daiane Garabeli Trojan – Universidade Norte do Paraná  
Prof. Dr. Darllan Collins da Cunha e Silva – Universidade Estadual Paulista  
Prof. Dr. Fábio Steiner – Universidade Estadual de Mato Grosso do Sul  
Prof<sup>a</sup> Dr<sup>a</sup> Girlene Santos de Souza – Universidade Federal do Recôncavo da Bahia  
Prof. Dr. Jorge González Aguilera – Universidade Federal de Mato Grosso do Sul  
Prof. Dr. Ronilson Freitas de Souza – Universidade do Estado do Pará  
Prof. Dr. Valdemar Antonio Paffaro Junior – Universidade Federal de Alfenas

#### **Ciências Biológicas e da Saúde**

Prof. Dr. Benedito Rodrigues da Silva Neto – Universidade Federal de Goiás  
Prof.<sup>a</sup> Dr.<sup>a</sup> Elane Schwinden Prudêncio – Universidade Federal de Santa Catarina  
Prof. Dr. Gianfábio Pimentel Franco – Universidade Federal de Santa Maria  
Prof. Dr. José Max Barbosa de Oliveira Junior – Universidade Federal do Oeste do Pará

Profª Drª Natiéli Piovesan – Instituto Federal do Rio Grande do Norte  
Profª Drª Raissa Rachel Salustriano da Silva Matos – Universidade Federal do Maranhão  
Profª Drª Vanessa Lima Gonçalves – Universidade Estadual de Ponta Grossa  
Profª Drª Vanessa Bordin Viera – Universidade Federal de Campina Grande

### **Ciências Exatas e da Terra e Engenharias**

Prof. Dr. Adélio Alcino Sampaio Castro Machado – Universidade do Porto  
Prof. Dr. Eloi Rufato Junior – Universidade Tecnológica Federal do Paraná  
Prof. Dr. Fabrício Menezes Ramos – Instituto Federal do Pará  
Profª Drª Natiéli Piovesan – Instituto Federal do Rio Grande do Norte  
Prof. Dr. Takeshy Tachizawa – Faculdade de Campo Limpo Paulista

### **Conselho Técnico Científico**

Prof. Msc. Abrãao Carvalho Nogueira – Universidade Federal do Espírito Santo  
Prof. Dr. Adaylson Wagner Sousa de Vasconcelos – Ordem dos Advogados do Brasil/Seccional Paraíba  
Prof. Msc. André Flávio Gonçalves Silva – Universidade Federal do Maranhão  
Prof.ª Drª Andreza Lopes – Instituto de Pesquisa e Desenvolvimento Acadêmico  
Prof. Msc. Carlos Antônio dos Santos – Universidade Federal Rural do Rio de Janeiro  
Prof. Msc. Daniel da Silva Miranda – Universidade Federal do Pará  
Prof. Msc. Eliel Constantino da Silva – Universidade Estadual Paulista  
Prof.ª Msc. Jaqueline Oliveira Rezende – Universidade Federal de Uberlândia  
Prof. Msc. Leonardo Tullio – Universidade Estadual de Ponta Grossa  
Prof.ª Msc. Renata Luciane Polsaque Young Blood – UniSecal  
Prof. Dr. Welleson Feitosa Gazel – Universidade Paulista

<b>Dados Internacionais de Catalogação na Publicação (CIP) (eDOC BRASIL, Belo Horizonte/MG)</b>	
T255	Técnicas de processamento de sinais e telecomunicações [recurso eletrônico] / Organizador Henrique Ajuz Holzmann. – Ponta Grossa, PR: Atena Editora, 2019.  Formato: PDF Requisitos de sistema: Adobe Acrobat Reader Modo de acesso: World Wide Web Inclui bibliografia ISBN 978-85-7247-449-8 DOI 10.22533/at.ed.498190807  1. Tecnologia da informação. 2. Telecomunicações. I. Holzmann, Henrique Ajuz.  CDD 338.47
<b>Elaborado por Maurício Amormino Júnior – CRB6/2422</b>	

Atena Editora  
Ponta Grossa – Paraná - Brasil  
[www.atenaeditora.com.br](http://www.atenaeditora.com.br)  
contato@atenaeditora.com.br



## APRESENTAÇÃO

A obra Técnicas de Processamento de Sinais e Telecomunicações está organizada de maneira a atender a temas atuais sobre a área de telecom e processamento de sinais de maneira sucinta e otimizada, sendo dividido em 17 capítulos sequenciais.

A transmissão de dados juntamente com suas vertentes representa um dos principais pilares para o progresso econômico de uma nação e para o atendimento de inúmeras necessidades da humanidade, estando presente nos mais diversos setores. Desenvolve-la de maneira eficiente é uma busca constante de grandes empresas e pesquisadores, buscando otimizar e agilizar o processo de troca de informações.

Produzir conhecimento nestas áreas é de extrema importância, a fim de gerar desenvolvimento e ampliar possibilidades nos mais diversos campos. Desta forma um compendio de temas e abordagens que facilitam as relações entre temas referentes a comunicação e processamento de sinais em diferentes níveis de profundidade em pesquisas, envolvendo aspectos técnicos, científicos e humanos é trazido nesta obra.

Boa leitura!

Henrique Ajuz Holzmann

## SUMÁRIO

<b>CAPÍTULO 1</b> .....	<b>1</b>
ANTENA DE MICROFITA RETANGULAR PARA APLICAÇÃO EM 2,5 GHZ UTILIZANDO SUBSTRATO METAMATERIAL	
Almir Souza e Silva Neto Bruno Pontes Alves da Silva Matheus Mesquita Correa Humberto César Chaves Fernandes Ronilson Mendes Fonseca	
<b>DOI 10.22533/at.ed.4981908071</b>	
<b>CAPÍTULO 2</b> .....	<b>7</b>
BANDWIDTH ENHANCEMENT OF AN ULTRA WIDE BAND PLANAR INVERTED F-ANTENNA	
Pedro Paulo Ferreira do Nascimento Glauco Fontgalland Raymundo de Amorim Júnior Tagleorge Marques Silveira Rodrigo César Fonseca da Silva	
<b>DOI 10.22533/at.ed.4981908072</b>	
<b>CAPÍTULO 3</b> .....	<b>14</b>
COMPORTAMENTO DE MODELOS DE DIFRAÇÃO SOBRE MÚLTIPLOS GUMES DE FACA EM VHF E UHF	
Lorenço Santos Vasconcelos Gilberto Arantes Carrijo	
<b>DOI 10.22533/at.ed.4981908073</b>	
<b>CAPÍTULO 4</b> .....	<b>27</b>
ON-CHIP KOCH FRACTAL ANTENNA ARRAY FOR 60 GHZ ISM BAND APPLICATION	
Paulo Fernandes da Silva Júnior Ewaldo Eder Carvalho Santana Mauro Sérgio Pinto Filho Almir Souza e Silva Neto Elder Eldervitch Carneiro de Oliveira Paulo Henrique da Fonseca Silva Alexandre Jean René Serres Raimundo Carlos Silvério Freire	
<b>DOI 10.22533/at.ed.4981908074</b>	
<b>CAPÍTULO 5</b> .....	<b>36</b>
PROJETO E ANÁLISE DE UM ARRANJO LINEAR DE ANTENAS UTILIZANDO A CURVA FRACTAL DE KOCH	
Elder Eldervitch Carneiro de Oliveira Pedro Carlos de Assis Júnior Marcelo da Silva Vieira Rodrigo César Fonseca da Silva	
<b>DOI 10.22533/at.ed.4981908075</b>	

<b>CAPÍTULO 6</b> .....	<b>48</b>
FINDING REPEATER PLACEMENT FOR P2P WIRELESS LINKS WITH NLOS IN EXTREMELY MOUNTAINOUS REGIONS	
Alvaro Javier Ortega	
<b>DOI 10.22533/at.ed.4981908076</b>	
<b>CAPÍTULO 7</b> .....	<b>60</b>
NOVA ARQUITETURA DE DEMODULADOR $\pi/3$ -BPSK PARA OS SATÉLITES DO SISTEMA BRASILEIRO DE COLETA DE DADOS	
Flavia Vasconcelos Maia	
Antonio Macilio Pereira de Lucena	
Francisco de Assis Tavares Ferreira da Silva	
<b>DOI 10.22533/at.ed.4981908077</b>	
<b>CAPÍTULO 8</b> .....	<b>73</b>
PROPOSTA DE UM NOVO ALGORITMO QOS-AWARE PARA O ESCALONAMENTO <i>DOWNLINK</i> LTE-A EM CENÁRIOS DE TRÁFEGO MISTO: UMA COMPARAÇÃO DE DESEMPENHO	
Júnio Moreira	
Éderson Rosa da Silva	
Paulo Roberto Guardieiro	
<b>DOI 10.22533/at.ed.4981908078</b>	
<b>CAPÍTULO 9</b> .....	<b>85</b>
SERVIÇO DE L2VPN EM REDES DE <i>BACKBONE</i> IP: ESTUDO DE CASO DA REDECOMEP-RIO	
Pedro Henrique Diniz da Silva	
Natália Castro Fernandes	
Nilton Alves Jr.	
Márcio Portes de Albuquerque	
<b>DOI 10.22533/at.ed.4981908079</b>	
<b>CAPÍTULO 10</b> .....	<b>101</b>
SISTEMA DISTRIBUÍDO PARA DETECÇÃO DE AMEAÇAS EM REDES UTILIZANDO <i>DEEP LEARNING</i>	
Fábio César Schuartz	
Mauro Sérgio Pereira Fonseca	
Anelise Munaretto	
<b>DOI 10.22533/at.ed.49819080710</b>	
<b>CAPÍTULO 11</b> .....	<b>113</b>
UM MÓDULO DE DEFESA PARA ATAQUES DDOS NA CAMADA DE APLICAÇÃO USANDO ESTRATÉGIAS SELETIVAS	
Túlio Albuquerque Pascoal	
João Henrique Gonçalves Corrêa	
Vivek Nigam	
Iguatemi Eduardo da Fonseca	
<b>DOI 10.22533/at.ed.49819080711</b>	

<b>CAPÍTULO 12</b> .....	<b>125</b>
AN EMPIRICAL RATE BALANCED ALIEN XTALK MITIGATION METHOD FOR G.FAST SYSTEMS	
Diego de Azevedo Gomes	
Cláudio de Castro Coutinho Filho	
João Victor Costa Carmona	
Evaldo Gonçalves Pelaes	
<b>DOI 10.22533/at.ed.49819080712</b>	
<b>CAPÍTULO 13</b> .....	<b>135</b>
REPRESENTAÇÃO ESPARSA UTILIZANDO WAVELETS E VARIAÇÃO TOTAL APLICADOS AO PROCESSAMENTO DE SINAIS DE DESCARGAS PARCIAIS	
Paulo Vitor do Carmo Batista	
Hilton de Oliveira Mota	
<b>DOI 10.22533/at.ed.49819080713</b>	
<b>CAPÍTULO 14</b> .....	<b>152</b>
REDUÇÃO DE DIMENSÕES USANDO TRANSFORMADA DE KARHUNEN-LOÈVE EM SISTEMAS MIMO MASSIVO DISTRIBUÍDO COM <i>FRONTHAUL</i> LIMITADO	
Ricardo de Souza Cerqueira	
André Noll Barreto	
<b>DOI 10.22533/at.ed.49819080714</b>	
<b>CAPÍTULO 15</b> .....	<b>167</b>
WSN COVERAGE IMPROVEMENT WITH ROF IN BUS TOPOLOGY FOR SMART CITIES	
Raphael Montali da Assumpção	
Indayara Bertoldi Martins	
Frank Herman Behrens	
Omar Carvalho Branquinho	
Fabiano Fruett	
<b>DOI 10.22533/at.ed.49819080715</b>	
<b>CAPÍTULO 16</b> .....	<b>179</b>
MODELO ELETROMAGNÉTICO DE UM ARRANJO PLANAR DE NANODIPOLOS SOBRE PLANO DE OURO ATRAVÉS DA FUNÇÃO DE GREEN 3D	
André Felipe Souza da Cruz	
Nadson Welkson Pereira de Souza	
Karlo Queiroz da Costa	
<b>DOI 10.22533/at.ed.49819080716</b>	
<b>CAPÍTULO 17</b> .....	<b>194</b>
AVALIAÇÃO DE FADIGA MUSCULAR LOCALIZADA EM SINAIS ELETROMIOGRÁFICOS UTILIZANDO TAXA DE AMOSTRAGEM VARIÁVEL NO TEMPO	
Jean Kevyn Correia Pessoa	
Pedro Henrique Melgaço de Oliveira Martins	
Thiago Raposo Milhomem de Carvalho	
<b>DOI 10.22533/at.ed.49819080717</b>	
<b>SOBRE O ORGANIZADOR</b> .....	<b>207</b>

## ANTENA DE MICROFITA RETANGULAR PARA APLICAÇÃO EM 2,5 GHz UTILIZANDO SUBSTRATO METAMATERIAL

### **Almir Souza e Silva Neto**

Instituto Federal de Educação, Ciência e Tecnologia do Maranhão (IFMA), Departamento de Engenharia Elétrica.

São Luís – MA

### **Bruno Pontes Alves da Silva**

Instituto Federal de Educação, Ciência e Tecnologia do Maranhão (IFMA), Departamento de Engenharia Elétrica.

São Luís – MA

### **Matheus Mesquita Correa**

Instituto Federal de Educação, Ciência e Tecnologia do Maranhão (IFMA), Departamento de Engenharia Elétrica.

São Luís – MA

### **Humberto César Chaves Fernandes**

Universidade Federal do Rio Grande do Norte (UFRN),

Departamento de Engenharia Elétrica.

Natal - RN

### **Ronilson Mendes Fonseca**

Instituto Federal de Educação, Ciência e Tecnologia do Maranhão (IFMA), Departamento de Engenharia Elétrica.

São Luís – MA

**RESUMO:** Este artigo apresenta uma análise de antenas de microfita, utilizando substrato metamaterial, para aplicação na frequência de 2,5 GHz. Foram simuladas quatro antenas

na frequência de 2,5 GHz, sendo três com substrato metamaterial e uma antena padrão. A antena 1 utilizou 14 espiras quadradas, a antena 2 utilizou 27 espiras quadradas, a antena 3, 27 anéis partidos. Através dos resultados simulados observou-se que a antena padrão apresentou uma largura de banda de 20 MHz e perda de retorno de -12,49 dB, enquanto a antena 1 apresentou uma largura de banda de 40 MHz e uma perda de retorno de -16,89 dB. As antenas 2 e 3 apresentaram um deslocamento da frequência desejada, sendo que a antena 3 apresentou uma largura de banda de 60 MHz. Embora a antena com metamaterial seja mais espessa, foi possível obter uma redução de 94 % da área do patch e redução no volume de 79%, quando comparadas a antena padrão para 2,5 GHz.

**PALAVRAS-CHAVE:** Antena, microfita, metamaterial.

**ABSTRACT:** This article presents an analysis of microstrip antennas, using metamaterial substrate, for application in the 2.5 GHz frequency. Four antennas were simulated at the 2.5 GHz frequency, three of them with metamaterial substrate and one standard antenna. The antenna 1 used 14 square turns, the antenna 2 used 27 square turns, the antenna 3, 27 split rings. Through the simulated results it was observed that the standard antenna had a



bandwidth of 20 MHz and a return loss of -12.49 dB, while antenna 1 had a bandwidth of 40 MHz and a return loss of -16, 89 dB. Antennas 2 and 3 showed a desired frequency shift, with antenna 3 having a bandwidth of 60 MHz. Although the antenna with metamaterial was thicker, it was possible to obtain a reduction of 94% of the area of the patch and reduction in the 79% compared to the standard 2.5 GHz antenna.

**KEYWORDS:** Antenna, microstrip, metamaterial.

## 1 | INTRODUÇÃO

A permissividade elétrica ( $\epsilon$ ) e a permeabilidade magnética ( $\mu$ ) são dois parâmetros que podem determinar as propriedades elétricas e magnéticas dos materiais. Estes, em conjunto, determinam o comportamento da onda eletromagnética quando se propaga através do material. Normalmente, a permissividade elétrica ( $\epsilon$ ) e a permeabilidade magnética ( $\mu$ ) são positivas em meios convencionais. A permissividade elétrica ( $\epsilon$ ) pode ser negativa em alguns materiais, porém não é conhecido nenhum material em estado natural com permeabilidade magnética ( $\mu$ ) negativa. Porém, alguns materiais, conhecidos como LHM (Left-Handed Materials) apresentam, simultaneamente, permissividade efetiva ( $\epsilon_{\text{eff}}$ ) e permeabilidade efetiva ( $\mu_{\text{eff}}$ ) com valores negativos, além de possuírem índice de refração ( $n$ ) negativo, no qual resulta na inversão dos fenômenos eletromagnéticos (VESELAGO et al., 1968; SOUSA NETO et al., 2011).

Estes materiais são chamados de metamateriais, onde o prefixo meta refere-se à natureza distinta de suas características eletromagnéticas (SCIENCE,2006; ITOH, 2006). Por causa dessas características excêntricas, os metamateriais estão sendo aplicados em sistemas eletromagnéticos e ópticos. As propriedades desses meios artificiais vão além de sua funcionalidade, uma vez que é permitida a combinação de sinais entre os parâmetros de permissividade e permeabilidade. Esses parâmetros são relacionados ao índice de refração  $n$  dado por (VESELAGO,1968; ITOH, 2006) :

$$n = \pm \sqrt{\mu_r \epsilon_r} \quad (1)$$

Em que  $\mu_r$  é a permeabilidade relativa e  $\epsilon_r$  é a permissividade relativas relacionadas à permeabilidade no espaço livre dadas por:  $\mu_0 = \mu/\mu_r = 4\pi \times 10^{-7} \text{ N/A}^2$  e a permissividade no espaço livre dadas por:  $\epsilon_0 = \epsilon/\epsilon_r = 8,854 \times 10^{-12} \text{ C}^2\text{N}^{-1}\text{m}^{-2}$ . As combinações de sinais para  $\epsilon$  e  $\mu$  são quatro (+,+), (+,-), (-,+), e (-,-) (VESELAGO et al., 1968; SMITH et al., 2000; SOUSA NETO et al., 2011).

## 2 | ANTENA DESIGN

Uma antena de microfita foi simulada com substrato metamaterial para a frequência de 2,5 GHz sobre um substrato RO3006 da Rogers Corporation, com

constante dielétrica de 6,15 e espessura  $p = 6$  mm.

A Figura 1 mostra a geometria do patch, o anel partido e a espira quadrada, com as seguintes dimensões:  $r = 2,7$  mm,  $s = 0,4$  mm,  $g = 0,5$ ,  $l = 5$  mm e  $t = 1$  mm. Os modelos tratam-se de ressoadores RLC com frequência ressonante  $\omega_0 = 1/\sqrt{LC}$ .

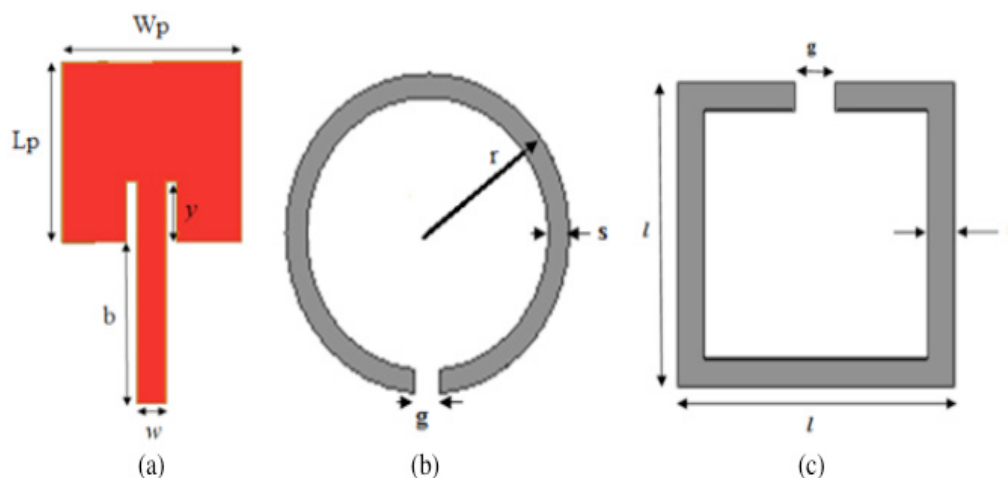


Figura 1. Geometria. a) Patch, b) SRR único e c) Espira quadrada.

Fonte: Elaborado pelo autor

Para efeito de comparação foi simulada uma antena retangular padrão para a frequência de 2,5 GHz utilizando o mesmo substrato RO3006.

A Tabela 1 mostra a comparação entre as dimensões da antena padrão e a antena com substrato metamaterial.

Dimensão (mm)	Antena Padrão	Antena com Metamaterial
W	56,7	12,56
L	75,05	18
p	1,52	6
Wp	31,73	9
Lp	23,95	9
b	23,76	8
w	2,23	1,5
y	8,24	3

Tabela 1 - Comparativo entre as dimensões da antena padrão e a antena com metamaterial.

Fonte: Elaborado pelo autor

Percebe-se que a antena com metamaterial possui dimensões reduzidas quando comparadas com a antena padrão. Apesar da antena com metamaterial ser mais espessa houve uma redução de 94 % da área do patch e redução no volume de 79%, quando comparadas a antena padrão para 2,5 GHz, utilizando um substrato RO3006 com constante dielétrica de 6,15 e espessura 1,52 mm. São analisadas quatro

antenas, sendo três antenas com metamaterial e uma sem metamaterial (padrão) para a frequência de 2,5 GHz. A primeira, a antena 1, possui 14 espiras quadradas e a segunda, a antena 2, apresenta 27 espiras quadradas e a terceira, antena 3, com 27 anéis partidos, e a quarta é uma antena padrão para a frequência de 2,5 GHz, conforme mostra a Figura 2.

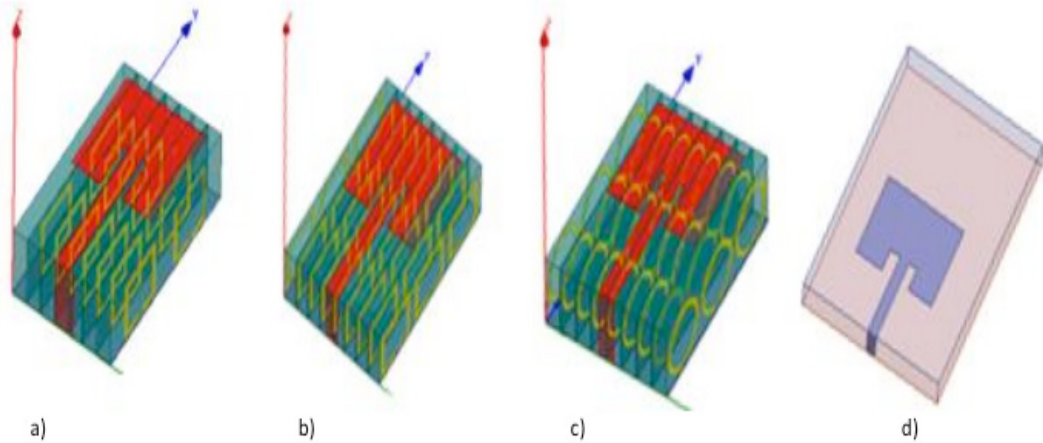


Figura 2. Antenas de microfita com substrato metamaterial. a) Antena 1 e b) Antena 2, c) Antena 3 e d) Antena padrão.

Fonte: Elaborado pelo autor

A figura 3 mostra o diagrama de radiação 2D das antenas 1 e a Padrão em 2,5 GHz.

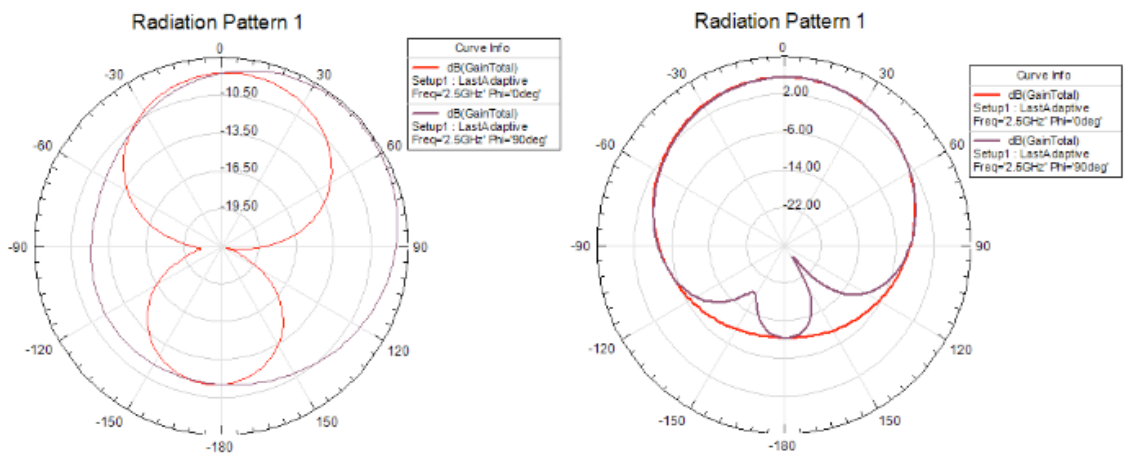


Figura 3. Diagrama de radiação (2D) da Antena 1 e da Antena Padrão em 2,5 GHz.

Fonte: Elaborado pelo autor

A comparação da perda de retorno em função da frequência para as quatro antenas é mostrada na Figura 4.

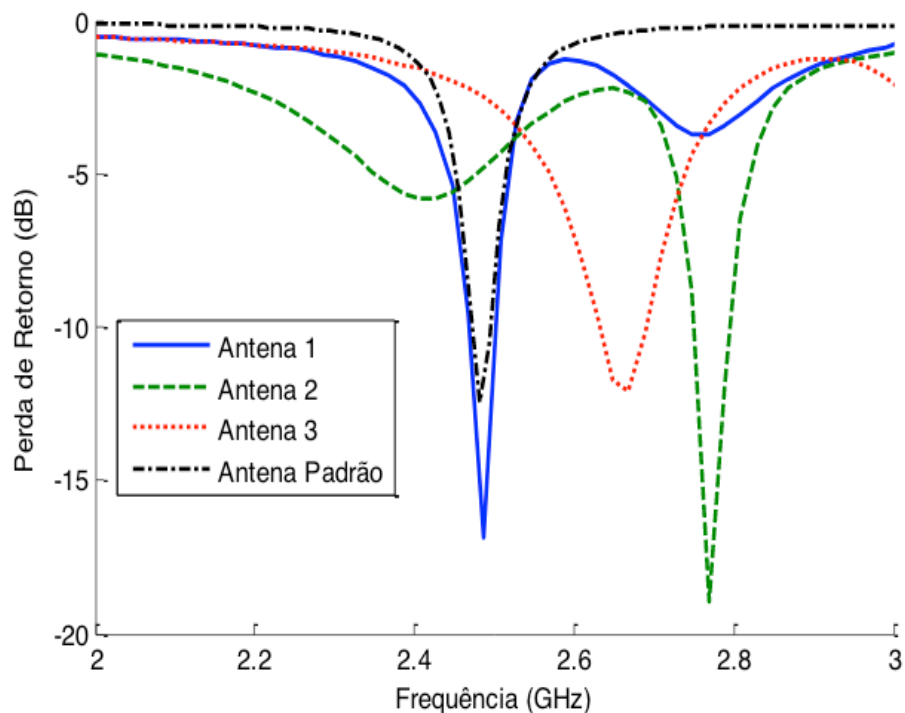


Figura 4. Comparação da perda de retorno entre a Antena 1, Antena 2, Antena 3 e Antena padrão.

Fonte: Elaborado pelo autor

A tabela 2 mostra um comparativo entre as antenas.

Antena	Frequência de Ressonância (GHz)	Perda de Retorno (dB)	Largura de Banda ( $S_{11} < -10$ dB) (MHz)
Antena 1	2,487	-16,89	40 (2,46 a 2,50 GHz)
Antena 2	2,769	-18,98	40 (2,75 a 2,79 GHz)
Antena 3	2,668	-12,11	60 (2,62 a 2,68 GHz)
Antena Padrão	2,48	-12,49	20 (2,47 a 2,49 GHz)

Tabela 2 - Comparativo de frequência de ressonância, perda de retorno e largura de banda entre as antenas metamateriais.

Fonte: Elaborado pelo autor

De acordo com a Figura 4, a antena 1 apresenta uma frequência de ressonância muito próxima da frequência de interesse (2,5GHz) . Para antena 2, com o aumento da quantidade de espiras quadradas há um deslocamento da frequência de ressonância, porém um melhor casamento de impedância. Com a antena 3, observa-se um deslocamento de frequência e um aumento de largura de banda. A antena padrão apresentou uma largura de banda menor que a antena 1 e uma perda de retorno de -12,49 dB, enquanto que a antena 1 apresentou uma perda de retorno de -16,89 dB.

### 3 | CONCLUSÕES

Foram feitas simulações com três antenas utilizando substrato metamaterial para a frequência de 2,5 GHz, sendo que a primeira possui 14 espiras quadradas no substrato, a segunda 27 espiras quadradas e a terceira 27 anéis partidos. A partir da curva de perda de retorno percebe-se um deslocamento da frequência a partir do momento que aumenta a quantidade espiras quadrada e ocorre um aumento de largura de banda para as estruturas em anéis. Embora a antena com metamaterial seja mais espessa, houve uma redução de 94 % da área do patch e redução no volume de 79%, quando comparadas a antena padrão para 2,5 GHz, utilizando um substrato RO3006 com constante dielétrica de 6,15 e espessura 1,52 mm.

### REFERÊNCIAS

ITO, T.; CALOZ, C. **Electromagnetic metamaterials: transmission line theory and microwave applications**. New Jersey, USA, John Wiley & Sons, Inc., 2006.

SCIENCE **Breakthrough the year: The runners-up**. Science, v. 314, no. 5807, pp. 1850-1855, 2006.

SMITH, D. R.; PADILHA, W. J.; VIER, D. C.; NEMAT-NASSER, S. C.; SCHULTZ, S. **Composite medium with simultaneously negative permeability and permittivity**. Phys. Rev. Lett., vol. 84, no. 18, pp. 4184-4187, 1 Maio, 2000.

SOUSA NETO, M. P.; FERNANDES, H. C. C. **Full-Wave Analysis of Rectangular Microstrip Antenna on Metamaterial**. In: INTERNATIONAL MICROWAVE AND OPTOELECTRONICS CONFERENCE Proceedings... IMOC 2011, Natal, Rio Grande do Norte, Brasil, 2011.

VESELAGO, V. G. **The electrodynamics of substances with simultaneously negative values of “ $\epsilon$ ” and “ $\mu$ ”**, Soviet Physucs Uspekhi. v. 10, no. 4, p. 509-514, 1968.



## BANDWIDTH ENHANCEMENT OF AN ULTRA WIDE BAND PLANAR INVERTED F-ANTENNA

**Pedro Paulo Ferreira do Nascimento**

pedro.nascimento@ee.ufcg.edu.br

Universidade Federal de Campina Grande, UFCG  
Campina Grande – PB

**Glauco Fontgalland**

Universidade Federal de Campina Grande, UFCG  
Campina Grande – PB

**Raymundo de Amorim Júnior**

Universidade Federal de Campina Grande, UFCG  
Campina Grande – PB

**Tagleorge Marques Silveira**

Universidade Federal de Campina Grande, UFCG  
Campina Grande – PB

**Rodrigo César Fonseca da Silva**

Universidade Estadual da Paraíba, UEPB  
Patos – PB

Simulation, Enhancement, Bandwidth

**RESUMO:** Este trabalho apresenta o acréscimo de largura de banda de uma Antena-F Planar Invertida (PIFA) de até 1,55 GHz, por adição de slots no elemento irradiador da estrutura. Os parâmetros S, ganho e diagrama de irradiação, após a análise, foram comparados com um modelo de PIFA já publicado. As simulações foram feitas utilizando-se do software comercial CST Design Studio. Um acréscimo de largura de banda de 53% foi alcançado.

**PALAVRAS-CHAVE:** Antena-F Planar Invertida, PIFA, Simulação, Acréscimo, Largura de Banda

**ABSTRACT:** This paper presents the bandwidth enhancement of a Planar Inverted F-Antenna (PIFA) up to 1.55 GHz as result of adding slots in the radiator element of the structure. The S-Parameters, gain, and radiation pattern after analyzed were used for comparison purpose with a published PIFA model. The simulations were made using the commercial software CST Design Studio. As result, a bandwidth enhancement of approximately 53% was achieved.

**KEYWORDS:** Planar Inverted F-Antenna, PIFA,

### 1 | INTRODUCTION

The PIFA (Planar Inverted-F Antenna) is largely utilized in various applications, especially in the mobile phones, as it has low cost, reduced size, and easy fabrication and implementation. Typically, the PIFA consists of: (i) ground plane; (ii) planar element; (iii) short-circuit plate or pin; (iv) feeding mechanism.

The first publication of PIFA appeared in the IEEE literature in 1987 (TAGA, 1987). Historically, its name comes from the evolution of Inverted-F Antenna. With similar design, the PIFA has its developing more closely related to

the Patch Antenna (RAY e CHAUDHURI, 2011). On the other hand, the Inverted-F Antenna is based on monopole antenna of a quarter of wavelength. The PIFA is an evolution of the Inverted-F Antenna with the substitution of the wire for a plate in order to enhance the bandwidth (CHATTHA, HUANG, *et al.*, 2011). Despite that, the bandwidth is still somewhat small and many techniques and parametric analysis have been done (CHATTHA, HUANG, *et al.*, 2011), (CHATTHA, HUANG, *et al.*, 2012). These studies include the insertion of slots in the radiator (SAHU e JYOTI, 2014) or in the ground plane (PICHER, ANGUERA, *et al.*, 2012), in order to enhance the bandwidth.

Motivated by the large application of PIFA in portable devices and by the limited results obtained by previous works to increase the broadband characteristics of the antenna, in this paper it is proposed the insertion of new slots in the design to enhance the bandwidth of an already published PIFA model.

## 2 | GEOMETRY DESIGN

Using the published model proposed in CHATTHA, HUANG, *et al.*, (2011), the slot insertion technique was used with the objective of improving the bandwidth of the model. In order to do that, firstly simulations were made in the commercial software CST Design Studio to recover the initial design and model parameters, as it can be observed in the Figure 1. The FR4 was used as substrate on the ground plane. The parameters and dimensions used were the same as in CHATTHA, HUANG, *et al.*, (2011) with the exception of  $W_s = 0.008\lambda$  and  $dr_x = 0.055\lambda$ , which were defined after optimization and have presented the best overall results in the 1.25 GHz – 3 GHz band. In CHATTHA, HUANG, *et al.*, (2011) the wavelength was defined as  $\lambda = 100$  mm (central frequency  $f_c = 3$  GHz), but in this study it was adopted  $\lambda = 162.05$  mm and, therefore, a central frequency of  $f_c = 1.85$  GHz. The reason for this choice was the enhancement of the lower frequencies response. It must be observed that one of the parasitic element proposed in CHATTHA, HUANG, *et al.*, (2011) was not add to the proposed structure, since it was not observed a significant contribution to the antenna.

Once the initial model was validated, the antenna's design by insertion of slots in the planar element and the ground plane was initiated. A careful parametric study and optimizations were made and it was possible to obtain a reasonable model, which can be seen in the Figure 2, whose enhancements in the bandwidth are considerable.

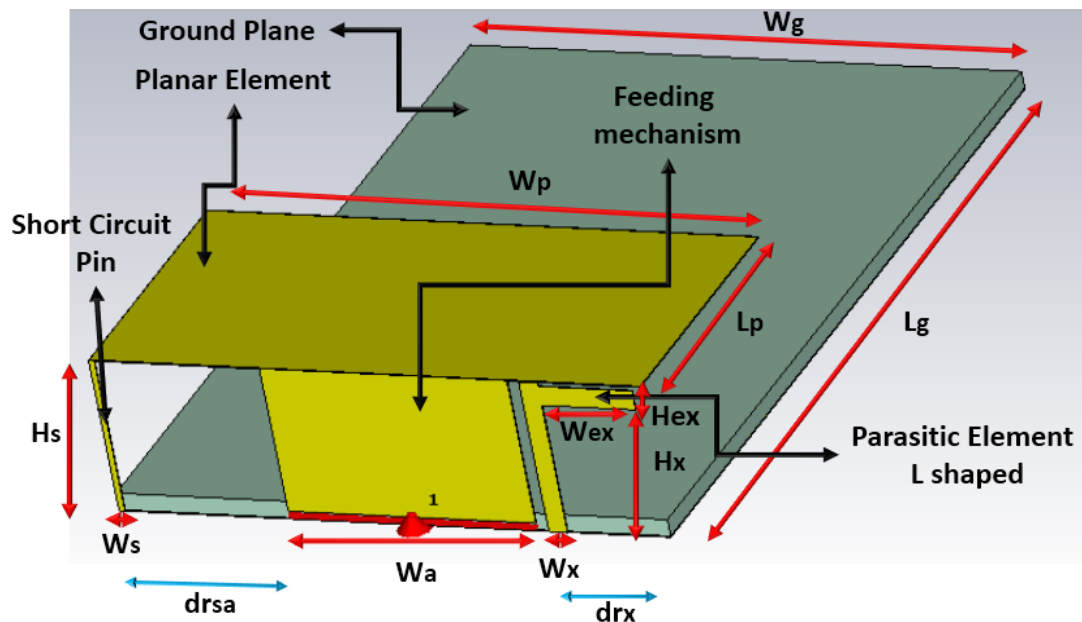


Figure 1 – Original Model as proposed by CHATTHA, HUANG, et al., (2011).

The modifications that presented the best results are: (i) Insertion of slots in the planar element and (ii) Removal of the FR4 Substrate. The inserted slots, A, B, and C, can be observed in the Figure 2. Lastly, the FR4 substrate was removed and the simulations has showed a substantial improvement. The final dimensions were organized in Table 1.

### 3 | RESULTS

For the original model, with  $\lambda = 162.05$  mm, the operation frequency band observed was from 1.28 GHz to 2.29GHz. In the new model, the enhancement in the upper frequency to 2.83GHz was obtained, which represents an improvement of about 540 MHz in the bandwidth, i.e. a bandwidth enhancement of 53%, considering VSWR  $< 2$ .

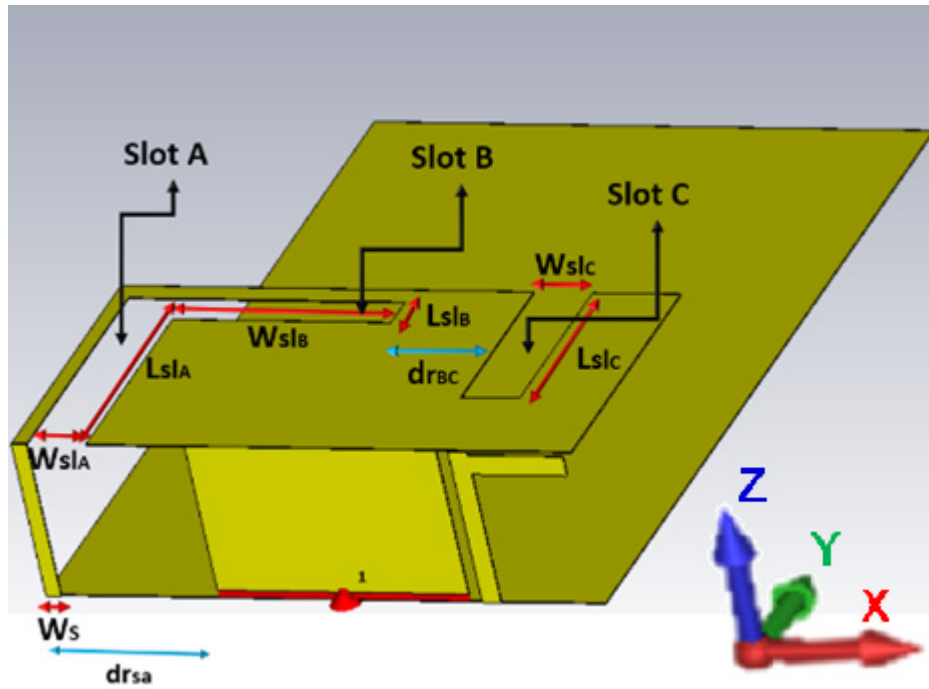


Figure 2 – Final model with the details of the slots and the FR4 substrate removal

Dimension	Value (mm)	Dimension	Value (mm)
$W_{sl_A}$	$0.031 \lambda$	$L_{sl_A}$	$0.136 \lambda$
$W_{sl_B}$	$0.112 \lambda$	$L_{sl_B}$	$0.019 \lambda$
$W_{sl_C}$	$0.031 \lambda$	$L_{sl_C}$	$0.098 \lambda$
$W_s$	$0.008 \lambda$	$dr_{SA}$	$0.055 \lambda$
$dr_{BC}$	$0.038 \lambda$		

Table 1 – Dimensions for the Final Model

The first slot inserted, Slot A, has the effect of reducing the reflection coefficient and slightly enhancing the bandwidth at higher frequencies. The addition of the second slot, Slot B, creates a resonance frequency at various frequencies according to the values chosen, and it greatly increases the bandwidth. The addition of the third slot, Slot C, shifts the resonance frequency and further expand the operation intervals to higher frequencies. The removal of the FR4 substrate and the insertion of the slots greatly enhances the reflection coefficient. All the curves for the reflection coefficient's results of the original model and the proposed model with insertion of slots, and FR4 substrate removal were presented in the Figure 3.

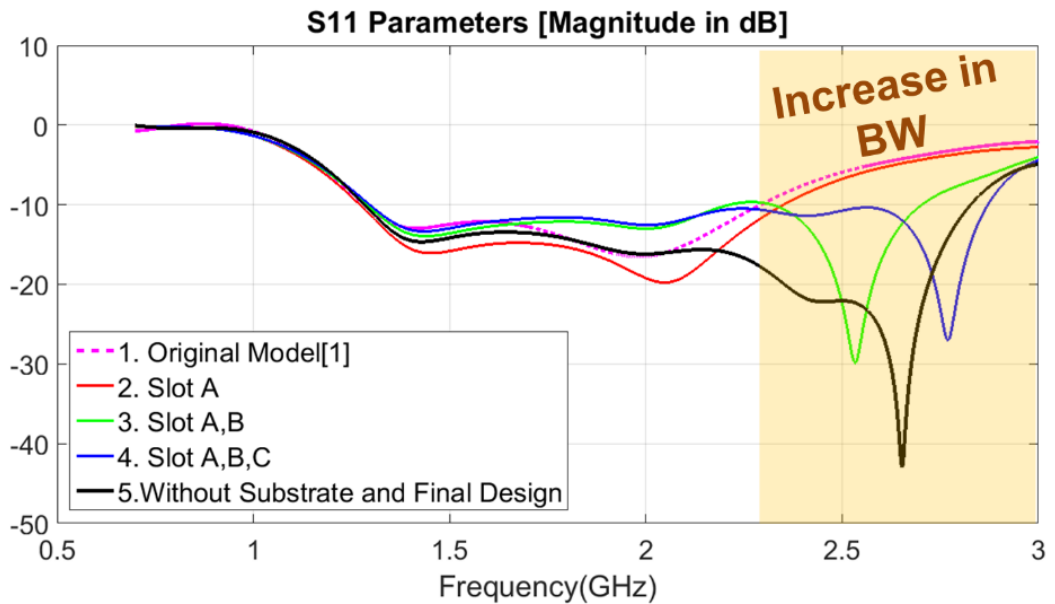


Figure 3 – Curves for simulated S11 Parameters

The maximum simulated gain over frequency band was analyzed in the models and it is presented in Figure 4. The new model has a great gain improvement in the frequencies up to about 1.8 GHz and a slight gain reduction above 2.2 GHz. Therefore, an equalized gain response was achieved. Looking at the  $S_{11}$  parameters these changes show that the slight gain reduction is accompanied with a greater reduction of the reflection coefficient value in the affected frequencies, i.e., a much better matching characteristic. These results are due to the many slots inserted, and consequentially, the reduction of conductive material, i.e., the radiation parts.

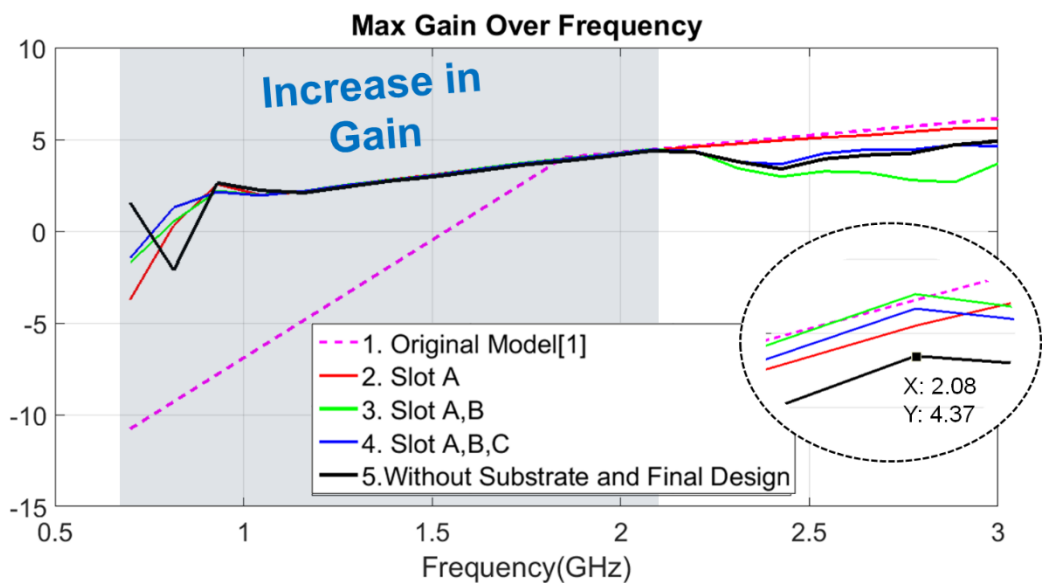


Figure 4 – Maximum simulated gain over frequency; for the analyzed models

The radiation pattern simulated is presented in Figure 5, a directional pattern was obtained in the azimuth plane whose Front to Back Ratio was 6.3 dB. Which is the expected result for the PIFA and therefore the changes did not impact on the radiation



pattern of the antenna.

## 4 | CONCLUSIONS

A good enhancement of the bandwidth, approximately 53% in comparison with the original model, and a great gain improvement in the frequencies up to 1.8 GHz and a slight reduction in the frequencies above 2.08 GHz was observed.

One possible application of this antenna is in GPS systems, as the bands L1, L3 and L4 are all within the range obtained for the PIFA with the reduction of its weight by substrate removal.

With these results, the objective is to keep expanding the bandwidth in order to obtain smaller operation frequencies and to keep the miniaturization of the antenna. The next step is to build a prototype made with copper plates sustained with an electromagnetic invisible material, such as Styrofoam, for structural support and to measure the antenna's parameters.

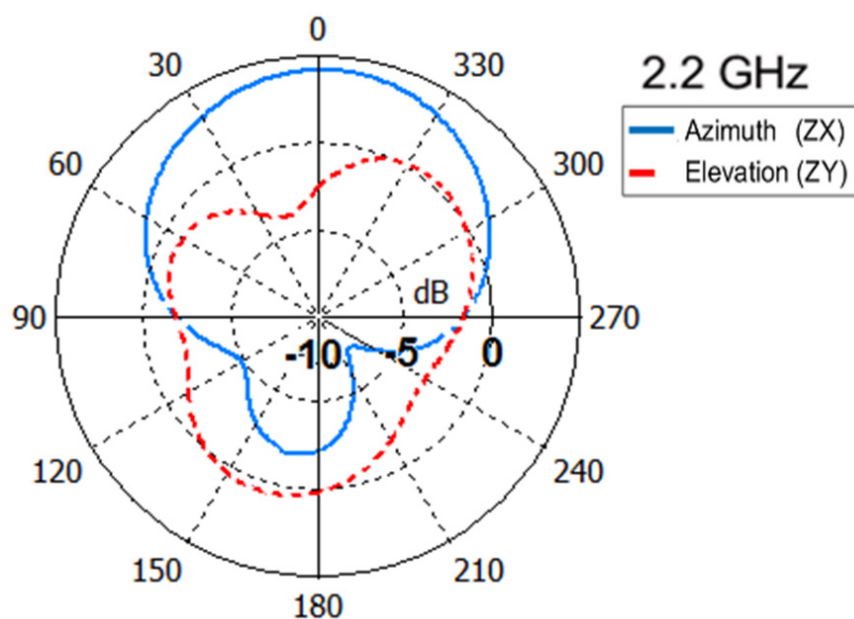


Figure 5 – Radiation pattern for the 2.2 GHz frequency for both the azimuth and the elevation plane

## REFERENCES

CHATTHA, H. et al. **A Comprehensive Parametric Study of Planar Inverted-F Antenna.** *Wireless Engineering and Technology*, 3, n. 1, 2012. 1-11.

CHATTHA, H. T. et al. **Bandwidth enhancement techniques for planar inverted-F antenna.** *IET Microwaves, Antennas & Propagation*, 5, n. 15, 9 Dezembro 2011. 1872-1879.

PICHER, C. et al. **Analysis of the Human Head Interaction in Handset Antennas with Slotted Ground Planes.** *IEEE Antennas and Propagation Magazine*, v. 54, n. 2, p. 36-59, Abril 2012.

RAY, J. A.; CHAUDHURI, S. R. B. **A review of PIFA technology.** Indian Antenna Week (IAW). Kolkata: [s.n.]. 2011. p. 1-4.

SAHU, L. K.; JYOTI, R. **Design and development of triple band U Slot PIFA antenna**. 2014 IEEE International Microwave and RF Conference (IMaRC). Bangalore: [s.n.]. 2014. p. 144-147.

TAGA, T. **Performance analysis of a built-in planar inverted F antenna**. *IEEE Journal on Selected Areas in Communications*, 5, n. 5, Junho 1987. 921-929.

## COMPORTAMENTO DE MODELOS DE DIFRAÇÃO SOBRE MÚLTIPLOS GUMES DE FACA EM VHF E UHF

### Lorenço Santos Vasconcelos

Universidade Federal de Uberlândia, Faculdade de Engenharia Elétrica  
Uberlândia – Minas Gerais

### Gilberto Arantes Carrijo

Universidade Federal de Uberlândia, Faculdade de Engenharia Elétrica  
Uberlândia – Minas Gerais

**RESUMO:** Neste trabalho, são comparados cinco diferentes métodos de cálculo de difração sobre múltiplos gumes de faca em áreas rurais. Os modelos de Bullington, Epstein-Peterson, Japonês, Deygout e Giovaneli foram implementados no MATLAB® e os seus resultados, analisados. Constatou-se que o comportamento relativo dos modelos tem estrita relação apenas com o número de gumes de faca do terreno e com a frequência de operação. Esses resultados fornecem uma base importante na escolha de métodos para se fazer análises de difração em enlaces de rádio em VHF e UHF.

**PALAVRAS-CHAVE:** difração, Giovaneli, gume de faca, terreno irregular

**ABSTRACT:** In this paper, we compared five different methods of multiple knife-edge diffraction calculation in rural areas. The Bullington, Epstein-Peterson, Japanese, Deygout and Giovaneli models were implemented in MATLAB® and the results analyzed. It is found that the relative behavior of the tested models is closely related to only the number of knife-edges and operating frequency. These results provide an important guidance in the choice of methods to do diffraction analysis in radio links in the VHF and UHF bands.

Deygout and Giovaneli models were implemented in MATLAB® and the results analyzed. It is found that the relative behavior of the tested models is closely related to only the number of knife-edges and operating frequency. These results provide an important guidance in the choice of methods to do diffraction analysis in radio links in the VHF and UHF bands.

**KEYWORDS:** diffraction, Giovaneli, irregular terrain, knife-edge

### 1 | INTRODUÇÃO

A propagação de ondas de rádio sobre terrenos irregulares envolve a atenuação causada pela difração nos vários obstáculos do terreno. Embora esses obstáculos sejam de formas e tamanhos muito variados, em frequências como VHF (*Very High Frequency*) e UHF (*Ultra High Frequency*), eles podem ser representados, com alguma perda de fidelidade, por gumes de faca.

A perda por difração em um gume de faca é calculada analiticamente pelas equações de Fresnel-Kirchhoff ou pelas equações da UTD (Teoria Uniforme da Difração) e da GTD (Teoria Geométrica da Difração). No entanto, quando se deseja calcular a atenuação devido a uma série de gumes de faca, o cálculo torna-se muito mais complicado e de difícil solução.

Millington *et al.* [1] desenvolveram soluções analíticas para o caso de dois gumes de faca e Vogler [12], baseado no trabalho de Furutsu [6], desenvolveu uma complexa expressão para o caso de múltiplos gumes de faca.

Devido à alta complexidade dos cálculos analíticos, foram criados métodos simples, mas que não têm grandes comprovações teóricas do seu funcionamento. Esses modelos são utilizados em análises de enlaces de rádio com o intuito de se fazer previsões da atenuação do sinal e da área de cobertura. Os modelos mais difundidos são: Bullington [3], Epstein-Peterson [5], Japonês [1], Deygout [4] e Giovaneli [7].

Tendo em vista a existência de vários métodos diferentes, é interessante realizar estudos comparativos para compreender o seu comportamento e os fatores que influenciam os resultados. Tais estudos são ponto de partida para guiar projetistas de sistemas na escolha de quais métodos utilizar. Com isso, o objetivo deste trabalho é comparar os cinco métodos citados, aplicando-os em vários percursos reais de propagação em terreno irregular e utilizando diferentes frequências nas faixas de VHF e UHF.

Portanto, este artigo foi estruturado da seguinte maneira: a seção 2 trata de aspectos teóricos da difração em múltiplos gumes de faca e traz alguns comentários sobre os métodos estudados. A seção 3 apresenta a metodologia empregada nos testes e as seções 4 e 5 trazem os resultados obtidos e as conclusões, respectivamente.

## 2 | DIFRAÇÃO POR MÚLTIPLOS GUMES DE FACA

A UTD [9], com base na GTD [8], postula que o campo elétrico difratado por um gume de faca é dado por uma expressão do tipo da Equação :

$$E^d = E_0 \bar{D}(\cdot) A(\cdot) e^{-j(\cdot)}, \quad (1)$$

em que  $\bar{D}(\cdot)$  é um coeficiente de difração que é função da geometria do gume e dos ângulos de difração;  $A(\cdot)$  representa a variação em amplitude do campo difratado em função da geometria da difração e da onda incidente e  $e^{-j(\cdot)}$  representa a variação em fase do campo difratado. Dessa maneira, pode-se dizer que a atenuação total é o produto de um fator de difração,  $\bar{D}$ , e um fator de propagação,  $A$ .

Pogorzelski [11] mostrou que a abordagem de Epstein-Peterson considera corretamente o fator de difração, mas é otimista em relação ao fator de propagação. Esse autor também mostrou que a solução de Deygout considera o fator de propagação correto, mas é pessimista com relação ao fator de difração. Já o método Japonês, de maneira semelhante ao de Epstein-Peterson, considera corretamente o fator de difração, porém corrige, parcialmente, o fator de propagação. Com isso, é natural que os resultados mais exatos estejam em algum ponto intermediário entre as soluções do método Japonês e do método de Deygout.

Por outro lado, a alternativa sugerida por Giovaneli corrige o fator de difração

incorreto do modelo de Deygout. Assim, essa abordagem considera corretamente os dois fatores e isso faz com que ela produza resultados mais consistentes com a UTD, GTD e com os resultados obtidos por Vogler.

Giovaneli comprovou a eficácia de sua abordagem ao encontrar erros menores que 0,9 dB entre as suas previsões e valores medidos experimentalmente. Bibb *et al.* [2] também mostraram que o modelo de Giovaneli se mantém mais próximo dos resultados de Vogler em várias situações.

Nesse sentido, esse trabalho considerou o modelo de Giovaneli como referência para os outros modelos. Assim, os resultados dos quatro primeiros modelos foram comparados com os resultados do modelo de Giovaneli.

### 3 I METODOLOGIA EMPREGADA

Para fazer a comparação entre os modelos, foi implementado um programa, em MATLAB®, capaz de realizar as seguintes tarefas: obter o perfil vertical do terreno; converter o perfil obtido em uma representação de gumes de faca; e realizar o cálculo da atenuação por difração utilizando os cinco métodos escolhidos.

A obtenção do perfil do terreno é feita, automaticamente, pelo programa utilizando-se o serviço Google Elevation API. O *software* faz uma requisição HTTP (Hyperext Transfer Protocol) ao servidor cuja resposta é um objeto JSON (JavaScript Object Notation) que contém as informações de altitude entre dois pontos escolhidos. Após a obtenção do perfil, o programa computa os gumes de faca, como é ilustrado na Figura 1.

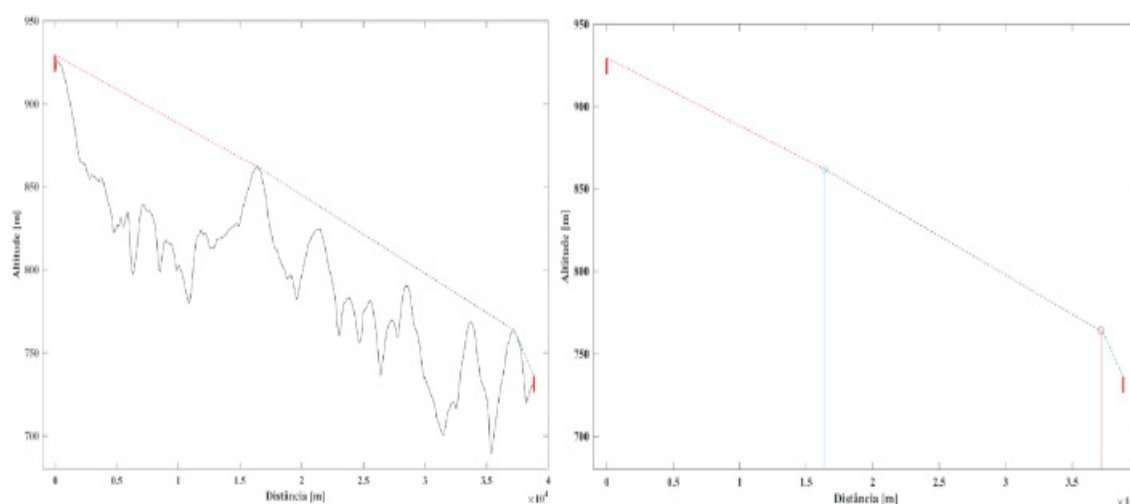


Figura 1: Conversão do perfil real em gumes de faca.

As análises feitas nesse trabalho seguiram, basicamente, dois procedimentos. O primeiro consistiu em avaliar os modelos apenas em função do terreno de propagação. Para isso, os cálculos foram realizados em diferentes regiões.

Definiu-se 6 áreas de teste distintas, compostas por 2401 pontos de recepção



diferentes. Para cada uma das 6 áreas, os cálculos foram feitos com 4 alturas diferentes de antena transmissora (10, 20, 50 e 100 m) e mesma frequência. A Tabela 1 resume as características das áreas de teste.

Área	Região/Cidade	Frequência [MHz]
1	Sul de Uberlândia	575,1428 (canal 31)
2	Noroeste de Uberlândia	
3	Nordeste de Frutal	
4	Sudoeste de Frutal	
5	Oeste de Patos de Minas	
6	Norte de Patos de Minas	

Tabela 1: Áreas de teste deste trabalho.

A Figura 2 mostra uma das regiões de teste analisadas. Nela, o ponto vermelho é o transmissor e os pontos azuis são os pontos de recepção.

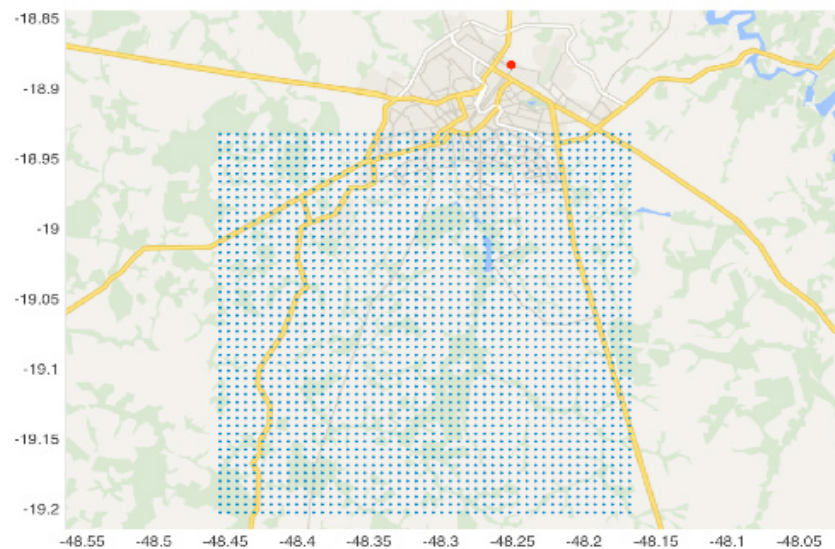


Figura 2: Área de teste 1 da Tabela 1.

O segundo procedimento consistiu em analisar o comportamento dos modelos selecionados em relação à frequência de operação. Para isso, os mesmos cálculos do primeiro procedimento foram realizados, mas para todos os 68 canais VHF e UHF (54 – 800 MHz).

## 4 | RESULTADOS

A primeira análise verificou o desvio de cada método em relação ao modelo de Giovaneli em função do terreno de propagação. Os Gráficos 1 a 4, a seguir, mostram os desvios para cada caso testado na área de teste 1 (a mesma da Figura 1).

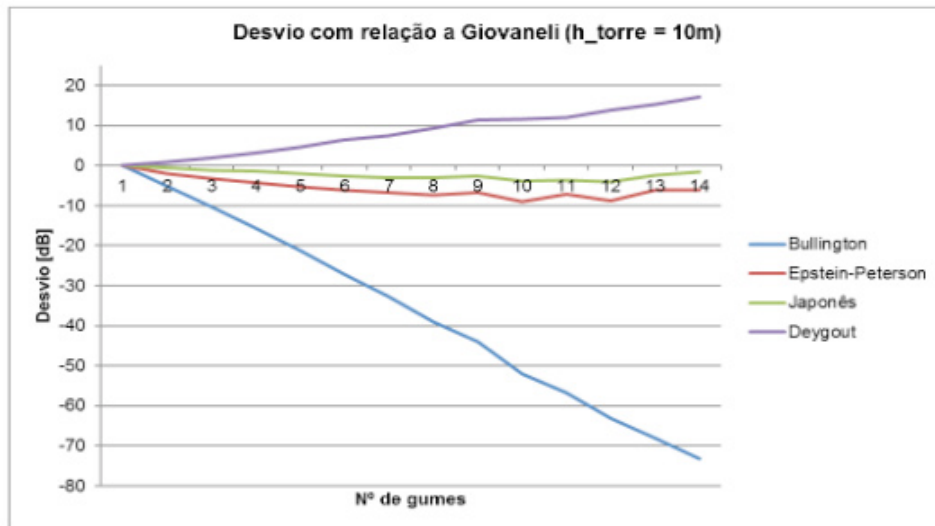


Gráfico 1: Desvio na área de teste 1 e para altura da antena transmissora = 10 m.



Gráfico 2: Desvio na área de teste 1 e para altura da antena transmissora = 20 m.

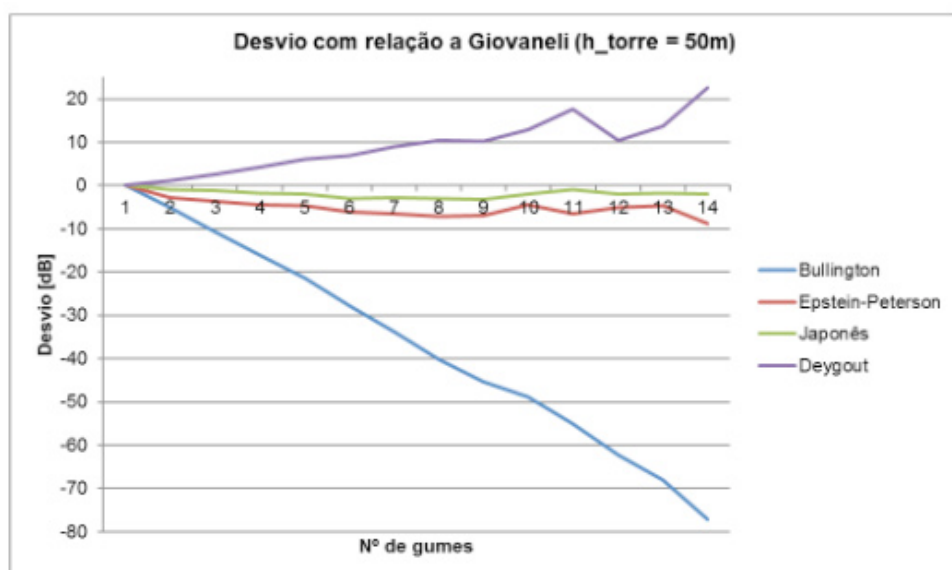


Gráfico 3: Desvio na área de teste 1 e para altura da antena transmissora = 50 m.

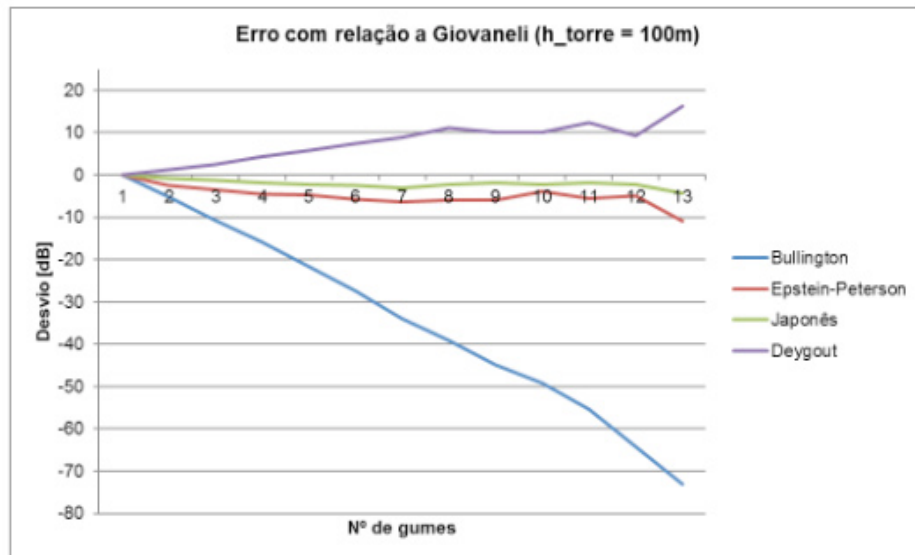


Gráfico 4: Desvio na área de teste 1 e para altura da antena transmissora = 100 m.

A partir desses resultados, verificou-se que a diferença entre os modelos segue um padrão, independentemente da altura da antena transmissora. Também foi possível perceber que, em todos os quatro casos, o modelo de Giovaneli (abscissa) ficou intermediário em relação aos modelos de Deygout (limite superior) e o Japonês (limite inferior), comprovando o que foi observado na Seção 2 deste trabalho.

Ao realizar as análises para as outras 5 áreas de teste restantes (Uberlândia, Patos de Minas e Frutal), verificou-se que o mesmo padrão se repetiu, embora os terrenos sejam completamente diferentes. Então, partiu-se para uma análise conjunta de todos os pontos, apenas em função do número de gumes de faca, independentemente da região analisada ou das alturas da antena transmissora. Isso é mostrado no Gráfico 5.



Gráfico 5: Desvio geral com relação a Giovaneli.

Observando o Gráfico 5, verifica-se que o mesmo padrão dos outros gráficos se repete. Então, isso sugere que a diferença entre os métodos de Bullington, Epstein-

Peterson, Japonês e Deygout, com relação ao método de Giovaneli, depende apenas da quantidade de gumes de faca no terreno.

O método de Bullington é o mais otimista e provê bons resultados apenas para um, ou, no máximo, dois gumes de faca. Quanto mais gumes há no terreno, mais ele se torna otimista, enquanto o modelo de Deygout torna-se pessimista, desviando-se bastante do modelo de Giovaneli.

Já o comportamento dos modelos de Epstein-Peterson e o Japonês é muito parecido. Eles se mantêm praticamente uniformes com relação ao número de gumes, mas o modelo Japonês é mais próximo do modelo de Giovaneli. Isso também pode ser observado na Figura 3, que traz as perdas por difração, em [dB], de todos os métodos.

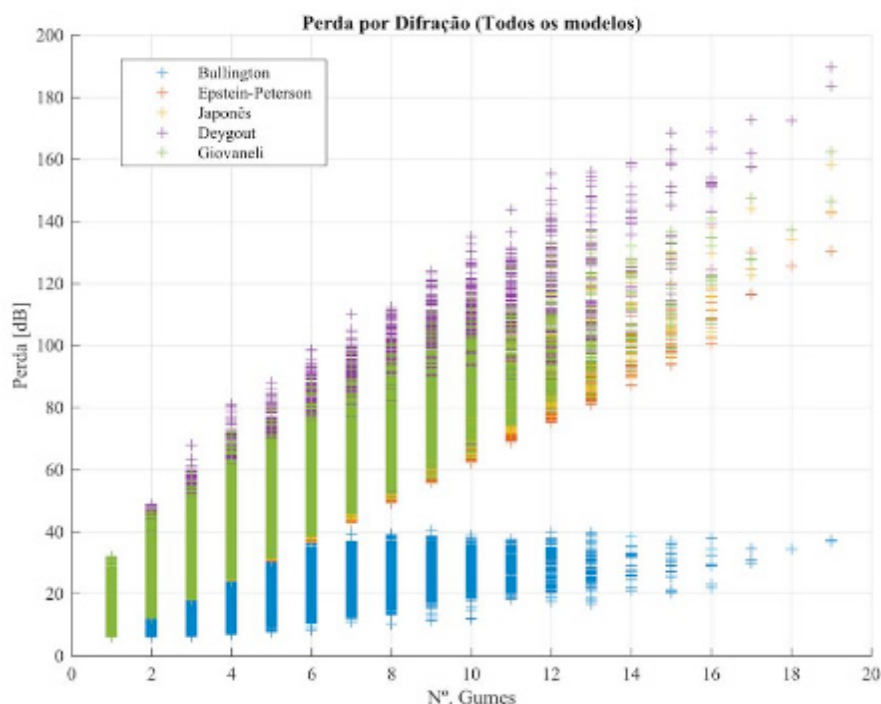


Figura 3: Perdas por difração de todos os métodos

O programa desenvolvido também foi feito para plotar, geograficamente, os dados calculados e isso é mostrado em seguida. A Figura 4 mostra o número de gumes de faca entre o transmissor e cada ponto da área de teste 1 para altura de antena transmissora de 100 m.

A Figura 5 ilustra o desvio padrão entre os modelos de Epstein-Peterson, Japonês e Deygout com relação ao modelo de Giovaneli nas mesmas condições da Figura 4. O modelo de Bullington foi omitido dessa análise para se ter uma boa escala de cores nas figuras, pois ele é muito divergente dos outros métodos.

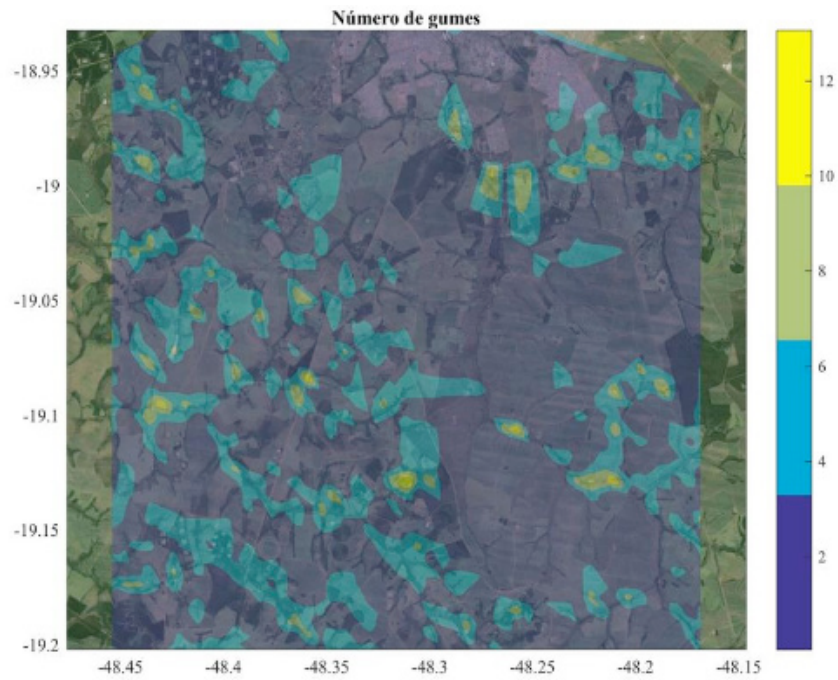


Figura 4: Número de gumes de faca entre o transmissor e cada ponto.

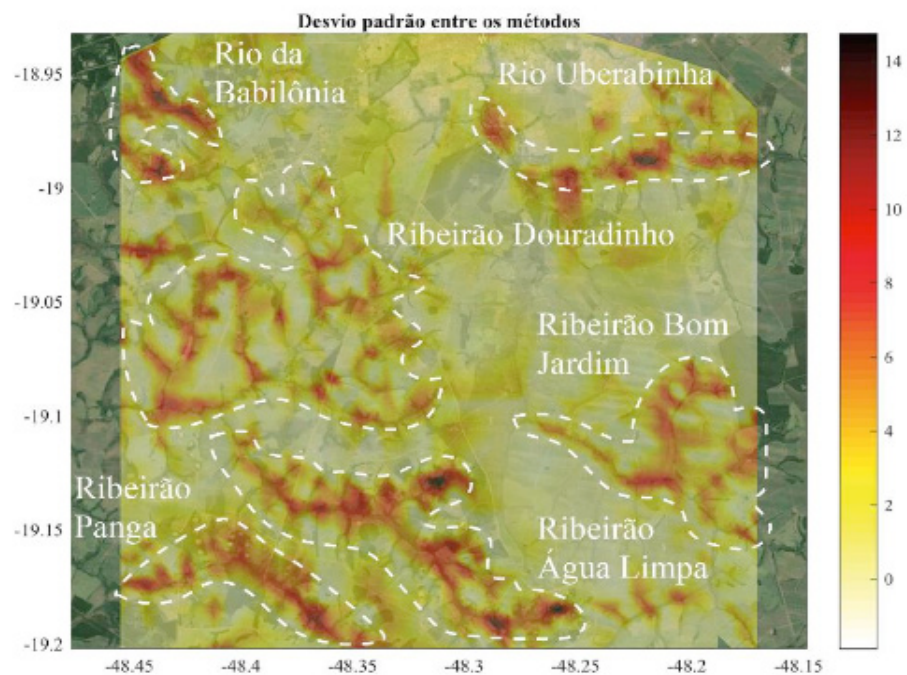


Figura 5: Desvio padrão entre os métodos com relação à Giovaneli.

Analisando as Figuras 4 e 5, é possível confirmar o que está mostrado nos Gráficos 1 a 5 e na Figura 3 (que a diferença entre os métodos e Giovaneli aumenta com o número de gumes). Nas mesmas regiões em que, na Figura 4, o percurso contém mais gumes, o desvio padrão entre os métodos é grande na Figura 5.

Curiosamente, se for feita uma análise hidrográfica, percebe-se que a maioria desses pontos são margens de pequenos rios e córregos, como foi identificado na Figura 5. Isso ocorre, pois, em geral, as margens desses componentes hidrográficos são formadas por vales e depressões. Assim, essas regiões de margens são sombreadas



por vários gumes de faca.

A segunda abordagem foi analisar o comportamento dos modelos de propagação em função da frequência de operação. Para isso, os cálculos foram realizados para todos os 68 canais VHF e UHF.

A Figura 6 mostra o comportamento dos modelos de Epstein-Peterson, Japonês e Deygout em relação ao modelo de Giovaneli em função da frequência e do número de gumes. A Figura 7 mostra o desvio padrão entre esses métodos e o de Giovaneli.

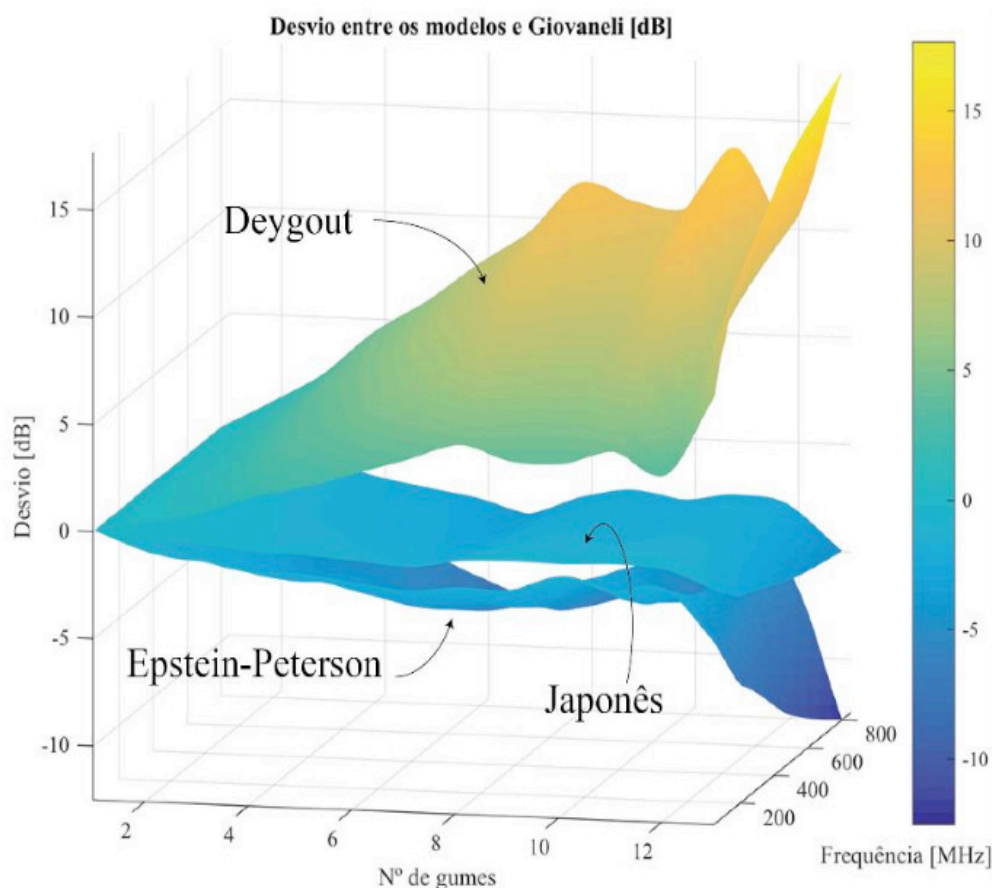


Figura 6: Comportamento dos modelos com relação a Giovaneli.

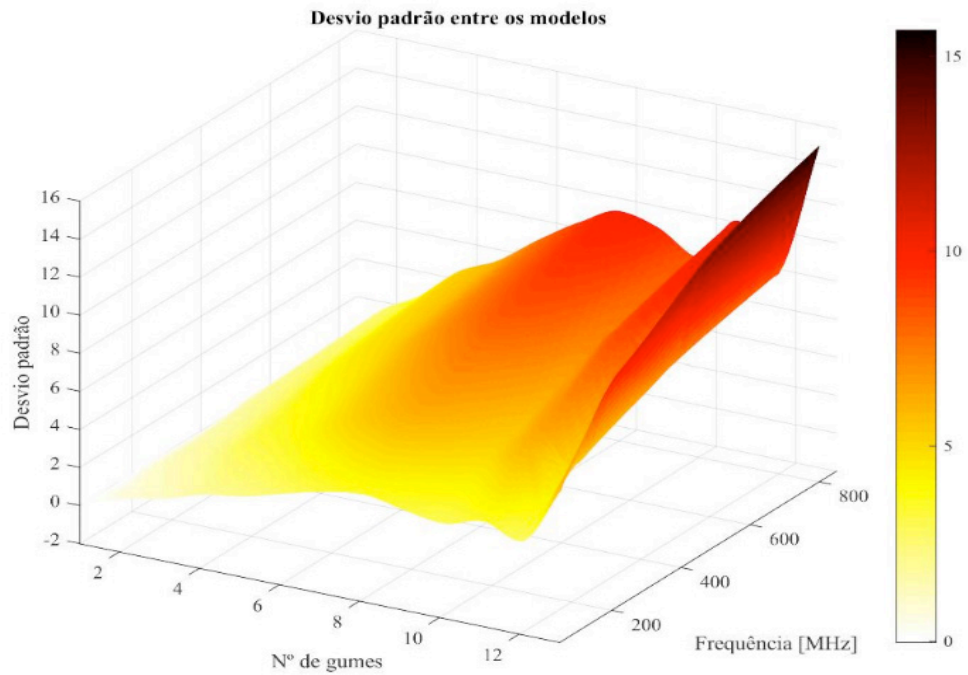


Figura 7: Desvio padrão entre os modelos e o método de Giovaneli.

Em concordância com os resultados anteriores, as Figuras 6 e 7 mostram que os métodos passam a divergir mais com o aumento do número de gumes de faca. Mas, ao introduzir o eixo das frequências, elas também mostram que as diferenças entre os métodos são acentuadas com o aumento da frequência, com exceção do modelo Japonês, que se manteve praticamente constante em relação ao modelo de Giovaneli.

Finalmente, foi avaliado o tempo de processamento dos cinco modelos estudados, como mostra o Gráfico 6.

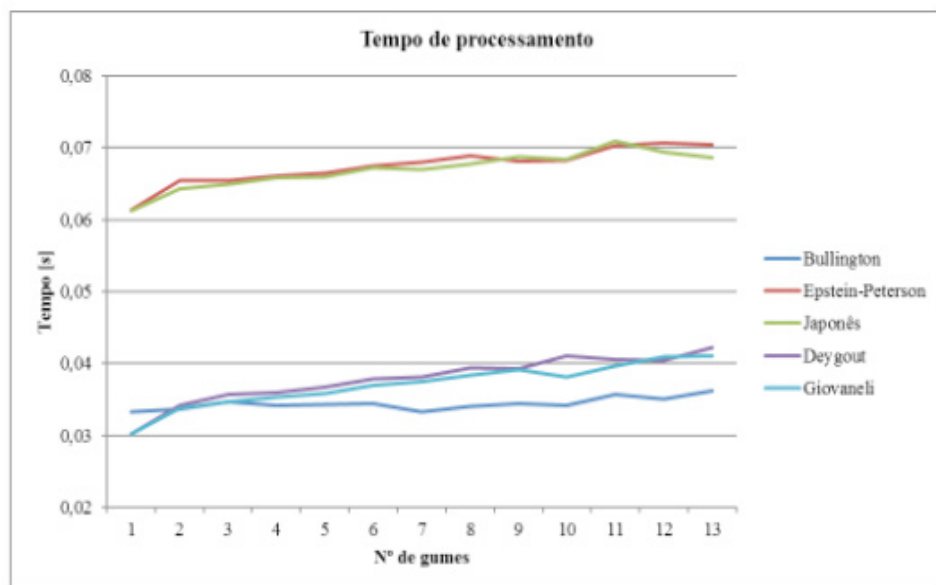


Gráfico 6: Tempo de processamento dos modelos.

Percebe-se que o modelo mais rápido é o de Bullington, o que era esperado devido à sua simplicidade. Observa-se, também que o seu tempo de processamento é



independente do número de gumes, pois esse modelo sempre utiliza, no máximo, dois gumes para realizar os cálculos.

Em seguida, aparecem o modelo de Giovaneli e o modelo de Deygout. Também era esperado que os tempos de processamento desses modelos fossem próximos, pois sua metodologia é bem parecida e isso realmente ocorreu.

Os modelos mais demorados foram os de Epstein-Peterson e o Japonês, o que também era esperado, pois eles realizam operações intensivas em todos os gumes de faca do terreno. Em geral, esses dois modelos são 1,8 vezes mais demorados do que o de Giovaneli.

## 5 | CONCLUSÃO

Este trabalho procurou realizar um estudo comparativo entre cinco diferentes métodos de cálculo de difração sobre múltiplos gumes de faca. Uma vez que o método de Giovaneli é considerado, por vários autores, como o mais exato na maioria das vezes (Seção 2 desse trabalho), então esse método foi considerado como referência para os outros quatro métodos analisados.

Os resultados apresentados nos Gráficos 1 a 5 e nas Figuras 3, 4 e 5 mostraram que a diferença entre os modelos de Bullington, Epstein-Peterson, Japonês e Deygout, com relação ao modelo de Giovaneli, são independentes da região analisada e da altura da antena transmissora, mas dependem apenas do número de gumes de faca. Ou seja, em geral, características como: distância entre gumes, altura média dos gumes, altura efetiva da antena transmissora, distância do enlace, entre outras, não interferem no comportamento dos modelos em relação a Giovaneli, conseqüentemente com relação à UTD, GTD e Vogler.

Ficou claro que o método de Bullington é o menos preciso e o que mais se desvia do modelo de Giovaneli. O método de Deygout fornece bons resultados quando são considerados dois ou três gumes, mas, a partir daí, desvia-se bastante de Giovaneli e passa a ser muito pessimista. O método de Epstein-Peterson forneceu resultados razoáveis e se manteve próximo do modelo de Giovaneli para toda a faixa de gumes testados. Isso mostra que, em geral, é um método confiável para ser utilizado em diversas situações, mas otimista. Já o modelo Japonês foi o que forneceu os melhores resultados. Ele se manteve muito próximo do modelo de Giovaneli em todos os casos e é ligeiramente otimista.

Além disso, as Figuras 4 e 5 mostraram que regiões de margens de pequenos rios e córregos devem ser cuidadosamente analisadas, pois os vales que podem estar presentes aumentam, consideravelmente, o número de gumes, elevando a divergência entre os métodos e o desvio com relação à realidade. Por isso, nessas áreas, é mais indicado utilizar os modelos de Giovaneli, Japonês e Epstein-Peterson.

A segunda análise levou em consideração a frequência de operação. As Figuras 6 e 7 mostraram que o comportamento dos modelos (em relação a Giovaneli) também

depende da frequência de operação. Em todos os casos, o aumento da frequência de operação acentuou o desvio dos modelos com relação ao modelo de Giovaneli, com exceção do modelo Japonês. O comportamento desse método se manteve aproximadamente constante para todas as quantidades de gumes de faca e para toda a faixa VHF e UHF.

Foi possível concluir, então, que, dentre os modelos analisados, o melhor é o Japonês. Uma vez que ele é mais próximo de Giovaneli, então ele também é mais próximo dos resultados de Vogler, da UTD e da GTD.

Como resultado geral deste trabalho, pode-se dizer que:

- para perfis de terreno com até três gumes de faca e frequências no começo da faixa de VHF, praticamente todos os cinco modelos podem ser utilizados;
- à medida que o número de gumes de faca aumenta, os modelos de Giovaneli, Epstein-Peterson e Japonês ainda podem ser utilizados, mas, a partir de, aproximadamente, 400 MHz, o modelo de Epstein-Peterson passa a prover piores resultados; e
- basicamente, os modelos de Giovaneli e Japonês podem ser utilizados para uma grande gama de gumes de faca e toda a banda de VHF e UHF.

## REFERÊNCIAS

**Atlas of radio wave propagation curves for frequencies between 30 and 10,000 Mc/s.** Radio Research Lab, Ministry of Postal Services, Tokyo, Japan, pp. 172–9, 1957.

BIBB, D. A; DANG, J; YUN, Z; ISKANDER, M. F. **Computational accuracy and speed of some knife-edge diffraction models.** Antennas and Propagation Society International Symposium (APSURSI), IEEE, pp. 705-706, 2014.

BULLINGTON, K. **Radio propagation at frequencies above 30 Mc.** Proc IRE, 35(10), pp. 1122–36, 1947.

DEYGOUT, J. **Multiple knife-edge diffraction of microwaves.** IEEE Trans., AP14(4), pp. 480–9, 1966.

EPSTEIN, J; PETERSON, D. W. **An experimental study of wave propagation at 850 MC.** Proc. IRE, 41(5), pp. 595–611, 1953.

FURUTSU, K. **On the theory of radiowave propagation over inhomogeneous earth.** J. Res. NBS, 67D, pp. 39–62, 1963.

GIOVANELI, C. L. **An analysis of simplified solutions for multiple knife-edge diffraction.** IEEE Trans., AP32(3), pp. 297–301, 1984.

KELLER, J. B. **Geometrical theory of diffraction.** J. Opt. Soc. Am., 52, pp. 116–30, 1962.

KOUYOUJIAN, R. G; PATHAK, P. H. **A uniform geometrical theory of diffraction for an edge in a perfectly conducting surface.** Proc. IEEE, 62(11), pp. 1448–61, 1974.

MILLINGTON, G; HEWITT, R; IMMIRZI, F. S. **Double knife-edge diffraction in field-strength**

**prediction.** IEE Monograph 507E, pp. 419–29, 1963.

POGORZELSKI, R. J. **A note on some common diffraction link loss models.** Radio Science, 17, pp. 1536–40, 1983.

VOGLER, L. E. **The attenuation of electromagnetic waves by multiple knife-edge diffraction.** NTIA Report 81–86. Available as PB82-139239, National Technical Information Service, Springfield VA, 1981.

## ON-CHIP KOCH FRACTAL ANTENNA ARRAY FOR 60 GHZ ISM BAND APPLICATION

### **Paulo Fernandes da Silva Júnior**

Post-graduation Program in Computer Engineering and Systems (PECS), State University of Maranhão. São Luís - Maranhão, Brazil, E-mail: pfs1224@gmail.com

### **Ewaldo Eder Carvalho Santana**

Post-graduation Program in Computer Engineering and Systems (PECS), State University of Maranhão. São Luís - Maranhão, Brazil.

### **Mauro Sérgio Pinto Filho**

Post-graduation Program in Computer Engineering and Systems (PECS), State University of Maranhão - Maranhão. São Luís, Brazil.

### **Almir Souza e Silva Neto**

Federal Institute of Educations, Science and Technology of Maranhão (IFMA), São Luís - Maranhão, Brazil.

### **Elder Eldervitch Carneiro de Oliveira**

State University of Paraíba (UEPB), João Pessoa - Paraíba, Brazil.

### **Paulo Henrique da Fonseca Silva**

Pos-Graduation Programa in Electrical Engineering, Federal Institute of Educations, Science and Technology of Paraíba (IFPB), João Pessoa - Paraíba, Brazil.

### **Alexandre Jean René Serres**

Center of Electriccal Engineering and Informatics, Federal University of Campina Grande (UFCG), Campina Grande - Paraíba, Brazil.

### **Raimundo Carlos Silvério Freire**

Center of Electriccal Engineering and Informatics, Federal University of Campina Grande (UFCG), Campina Grande - Paraíba, Brazil.

**ABSTRACT:** The Industrial Scientific and Medical band in 60 GHz (57 GHz to 64 GHz) is being proposed to high data rate wireless transmission requiring antennas with efficient design. In this paper we present an on-chip array aperture couple antenna, using Koch fractal geometry. The proposed antenna is composed by a transmission feed line, an aperture and patch element built in aluminum with 2 micrometers of thickness lying on two layers of silicon with height of 200 micrometers. Those dimensions were calculated according to the effective wavelength for resonance frequency at 60 GHz in order to match at 50 Ohms. The proposed antennas have been simulated using the ANSYS software, and the results show an omnidirectional radiation pattern, with half power beamwidth greater than 124 degrees, gain of 16.7 dBi, and polarization close to the circular polarization with axial ratio less of 2.22.

**KEYWORDS:** *Antenna; Design ISM Simulation.*

## 1 | INTRODUCTION

The increasing use of the electromagnetic spectrum for various technologies has implied the necessity of more bandwidth resulting in to research the utilization of higher frequency bands, especially in the millimeter wave range (30 - 300 GHz). Industrial scientific and medical

(ISM) band in 60 GHz range (57–64 GHz), normalized by Federal Communication Commission (FCC) have bandwidth of 7 GHz for high data rate wireless communication (NIKNEJAD e HASHEMI). Services in the ISM 60 GHz include technologies like Wireless Fidelity (Wi-Fi), wireless personal local area network (WPAN), and communication of fifth generation (5G) (ZHAO e REYNAERT).

Allied to these new systems, there is the necessity of device's portability, and the development of broadband antennas in Silicon at 60 GHz has the advantage to permit their utilization inside integrated circuit. Considering this scenario, the mutual interference and electromagnetic coupling, and the standards indicated for each technology, aperture-coupled antennas appear to be a good solution since it promotes gain with compact structures (NIKNEJAD e HASHEMI), (ZHAO e REYNAERT), (ZHANG, LI e LIU), (KHAM et al).

This paper presents the development of an on-chip array Koch fractal aperture-coupled antenna, operating in ISM application at 60 GHz band, built in silicon and aluminum. The rest of the paper is divided as follows. Section II shows the basic structure of aperture-coupled antennas. Section III shows the antennas proposed and the materials and methods used. Section IV presents the different results achieved, and finally conclusions are drawn in section V.

## 2 | APPERTURE-COUPLE ANTENNAS

Aperture-coupled antennas are compact broadband antennas composed by two or three layers used in many applications and technologies [5-8]. Generally, these layers are the transmission feed line, aperture, patch elements, and in some cases, one parasite patch element, separated by dielectrics with different permittivity. According to (STUZMAN e THIELE, 2013), layers with different permittivity can be used to filter spurious radiations. The association of more one layer with a radiating element upper, but smaller than the lower layer patch, promotes the union of the two resonant frequencies close it is yield a broadband effect.

In the model of the aperture-coupled antennas, the transmission line is feed generating an electromagnetic field through the substrate to the slot layer, concentrating part of the power through the aperture, illuminating the patch in the higher layer (STUZMAN e THIELE, 2013), (KIM), (PACHI, KAMARKAR e LAW). Fig. 1 shows structure of aperture-coupled antenna with three layers.

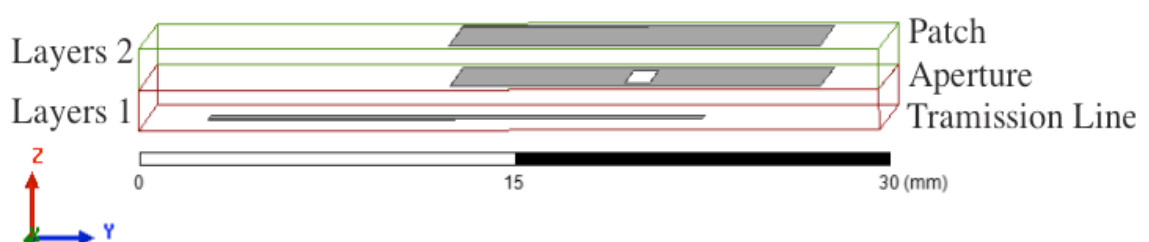


Fig. 1. Structure of aperture-coupled antenna.

In (BERNARDO et al) was compared aperture-coupled patch antennas operating in industrial scientific and medical band at 60 GHz (57 – 64 GHz), with square and circular shapes, the square patch presented a more compact structure with dimension less than 1.55 mm. The conception of patch antennas, coupled with the use of low cost dielectric and radiating element, such as silicon and aluminum, enables a versatile antenna design. From results of (BERNARDO et al), this paper presents a Koch fractal aperture-coupled antenna array for ISM band at 60 GHz.

### 3 | MATERIALS AND METHODS

From a mathematical point of view, a fractal is a set in Euclidean space with specific properties, such as self-similarity or self-affinity, simple and recursive definition, irregular shape, and natural appearance (FALCONER, 2003). Fractal geometry is the study of sets with these properties, which are too irregular to be described by calculus or traditional Euclidian geometry language (SULLIVAN, 2000), (MANDELBROT, 1982). The Koch fractal is used, generally, in patch antenna, to increase the perimeter of the radiating element and consequently decrease the resonance frequency in a smaller area that implies a reduction of gain and greater current density.

Fractals are resort to conventional classes, such as geometrical fractals, algebraic fractals and stochastic fractals (FALCONER, 2003). Two common methods used to generate mathematical fractals are Iterated Function Systems (IFS) and Lindenmayer Systems (FALCONER, 2003), (MANDELBROT, 1982).

Antenna array is used to alter or improve parameters such as bandwidth and gain [5]. In this work, antennas with identical dimensions but different shapes are used to increasing antenna gain in the operating frequencies.

The first dimensions of the designed aperture-couple are related to the effective wavelength,  $\lambda_{eff}$  at 60.5 GHz resonant frequency. From simulations in the ANSYS software it was possible to optimize the dimensions taking the  $\lambda_{eff}$  given by

$$\lambda_{eff} = \frac{c}{f_0 \sqrt{\frac{\epsilon_r + 1}{2}}}, \quad (1)$$

where  $c$  is the light speed in vacuum,  $f_0$  is the resonance frequency in Hz, and  $\epsilon_r$  is the relative permittivity of the material.

The aperture-coupled patch antennas were designed using silicon as dielectric material, with thickness of 200  $\mu\text{m}$ , and dielectric permittivity,  $\epsilon = 11.9$ , loss tangent of 0.001, and radiating elements of aluminum, with thickness of 2  $\mu\text{m}$ .

Fig. 2 shows the proposed on-chip Koch fractal aperture-coupled array antenna,

with Koch fractal level 0 (square) Fig. 1(a), Koch fractal level 1 Fig. 1(b), array Koch fractal level 1 Fig. 1(c), and structure antenna Fig. 1(d).

Fig. 3 shows the dimensions in function to  $\lambda_{eff}$  of Koch level 0 (square shape) Fig. 3(a), Koch level 1 Fig. 3(b), and array Koch level 1, Fig. 3(c), and antennas in the chip of area of 1 cm<sup>2</sup>. Koch fractal level 1 has been built with insertion of square of 1/3 of total dimension of central element, in borders of the patch antenna. This promotes an 1/3 increase in total structure of patch element. In fractal geometry applied to antenna project there is an increase of the electrical perimeters and reduction of resonance frequency, thus, the Koch level 0 (square) was projected with central frequency of 70 GHz,  $\lambda_{eff} = 1.6875$ . In the case of Koch fractal level 1, it was used a central frequency at 60.5 GHz with  $\lambda_{eff} = 1.9611$  mm, increasing the perimeter of 33.33%.

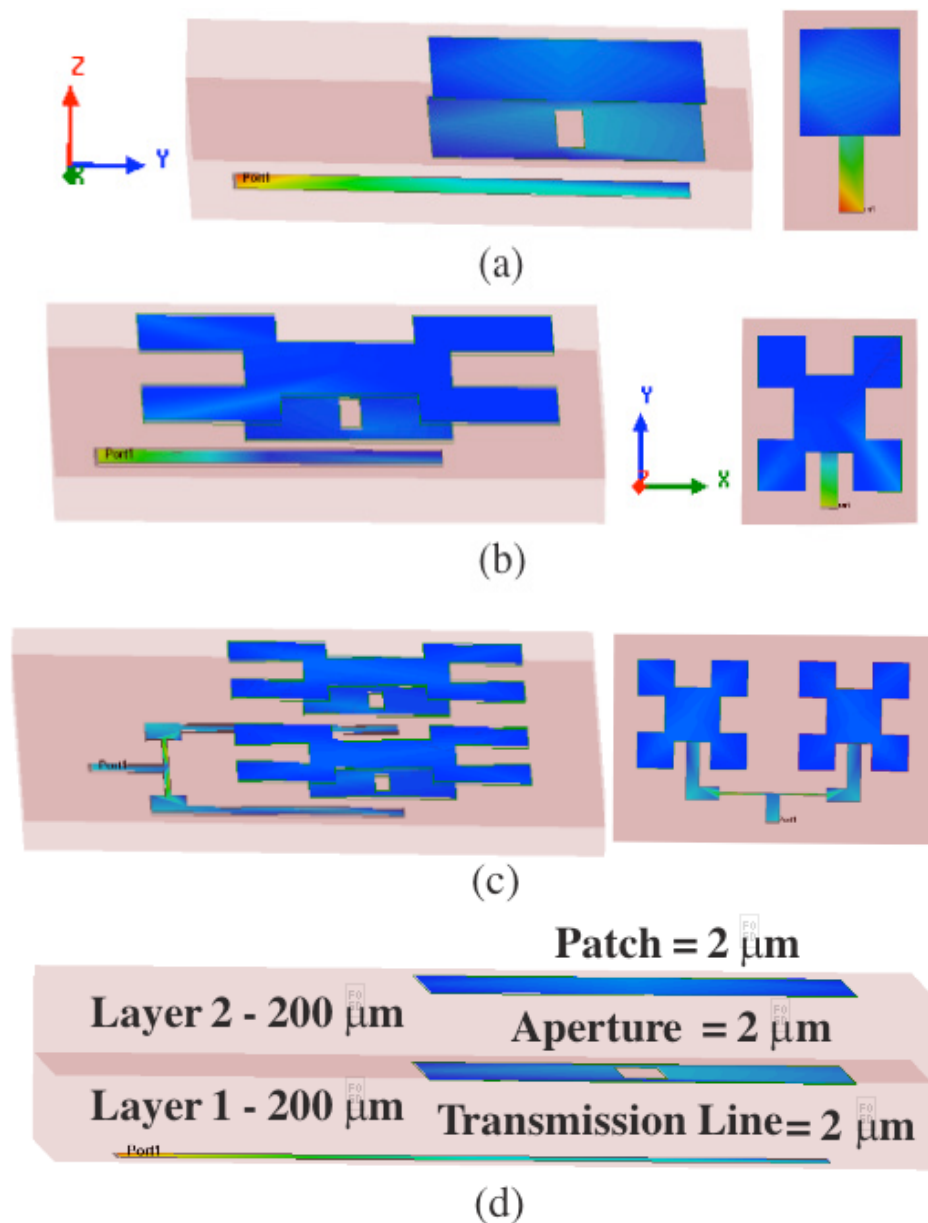


Fig.2. Structure of on-chip Koch fractal aperture-coupled antennas array: a) Koch fractal level 0 (square); b) Koch fractal level 1; c) Koch level 1 array.



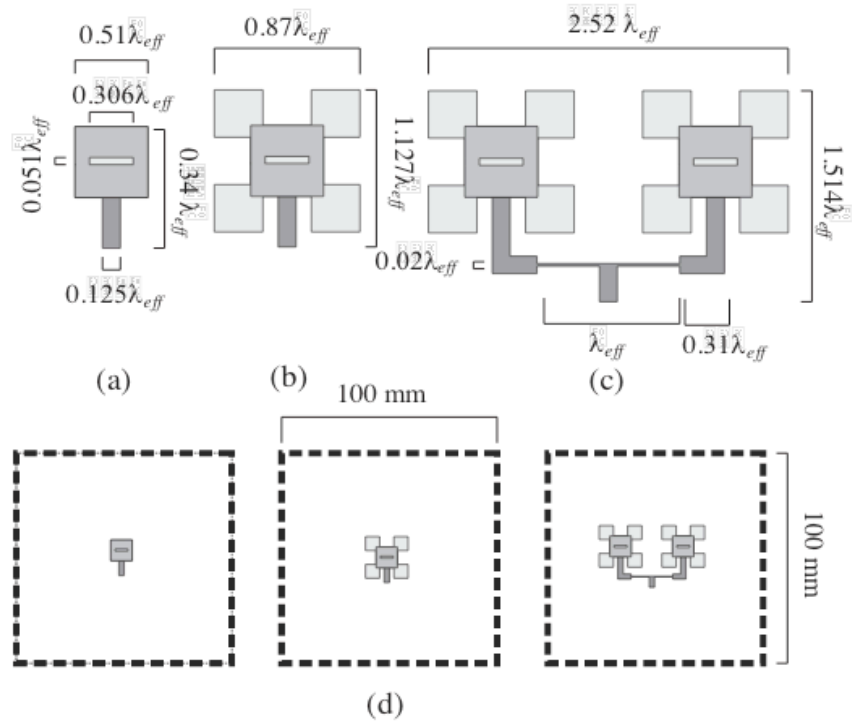


Fig.3. Structure of on-chip Koch fractal aperture-coupled antennas: a) Koch fractal level 0 (square); b) Koch fractal level 1; c) array Koch level 1.

#### 4 | ON-CHIP ARRAY ANTENNA WITH KOCH FRACTAL GEOMETRY

Fig. 4 shows  $IS_{11}$  values for Koch fractal aperture-coupled antennas from 50 GHz to 90 GHz. Table I presents values of first resonance frequency,  $f_1$ , second resonance frequency,  $f_2$ , central resonance frequency,  $f_0$ , bandwidth, BW, and return loss of simulated antennas.

It can be observed that the on-chip Koch fractal antennas level 1 obtained variation in resonance frequency of 14.28%, and bandwidth of 34.95%, covering part of ISM 60 GHz (57 – 64 GHz). The greater bandwidth has been observed in Koch fractal level 1, with small difference in resonance frequency between Koch fractal level 1 and array Koch fractal level 1 of 1.11%.

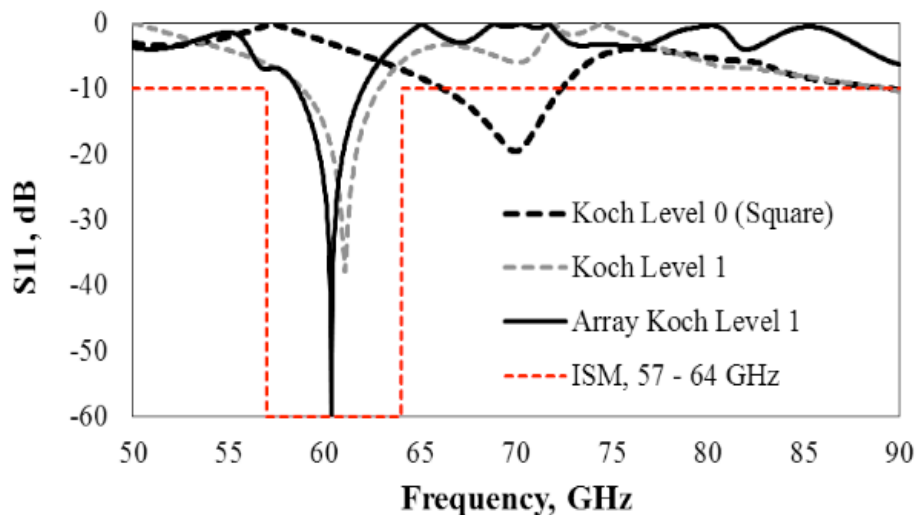
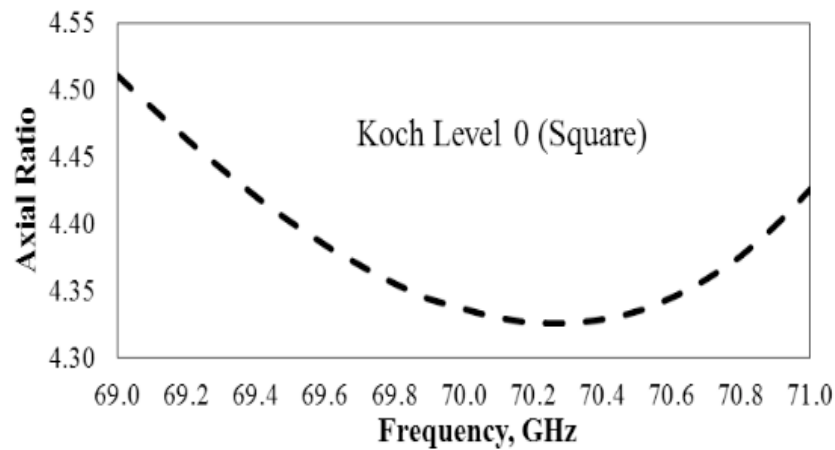


Fig.4. Comparison of  $S_{11}$  parameter of simulated on-chip Koch fractal aperture-coupled antennas.

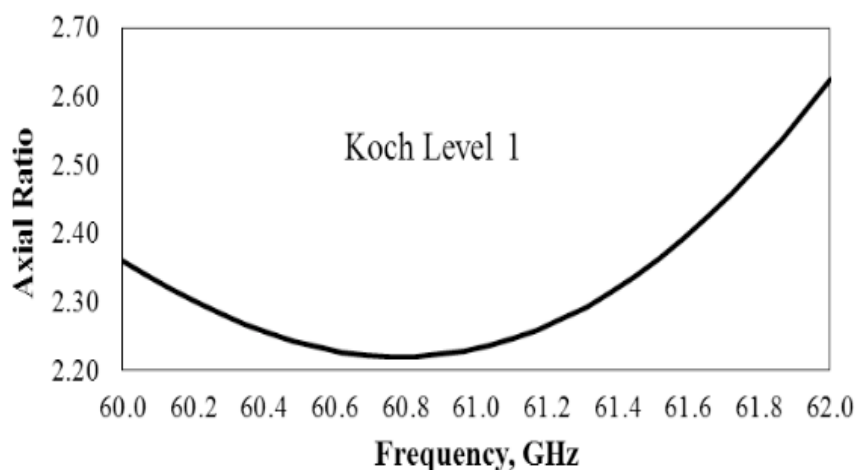
Antenna	$f_0$ (GHz)	$f_1$ (GHz)	$f_2$ (GHz)	Bandwidth (GHz)	Return Loss (dB)
Koch 0	70	66.20	72.35	6.15	-19.50
Koch 1	61.06	58.85	62.85	4.00	-36.15
Array Koch 1	60.38	58.60	62.00	3.40	-59.67

Table 1 Results of On-Chip Koch Fractal Aberture-Coupled Atennas

From axial ratio shown in Fig. 5, it can be noticed that the Koch fractal level 1, in resonance frequencies have polarization close to circular polarization Fig. 4(b), i.e., results of axial ration close to 1, indicating that the antenna can receive/transmit signals with low polarization losses, regardless of its position. This result can be associated to the relation of antenna dimension and aperture.



(a)



(b)

Fig.5. Axial ratio of aperture-coupled antenna: a) Koch level 0; b) Koch level 1.

Fig. 6 shows Gain (dBi) curves in resonance frequency of on-chip Koch fractal

antennas. According to [11], there is a relationship between the quantity of metal in the radiating element and the gain in the patch antennas. Thus, the array Koch fractal level 1 achieves gain of 16.7 dBi, presenting an increase of 10.66 dBi (142.9%) and 8.63 dBi (106.9%) compared to the Koch fractal level 0 and the Koch fractal level 1, respectively.

Radiation pattern, with HPBW indications in resonance frequencies of Koch fractal aperture-coupled antennas are shown in Fig. 7. Antennas presents omnidirectional radiation pattern, with greater variations for electric fields ( $\theta = 0^\circ$ ), magnetic fields ( $\theta = 90^\circ$ ), and half power beamwidth (HPBW) for array Koch fractal antennas of 46.15%, Fig.7(c), compared to Koch fractal 1, demonstrating that array Koch fractal level 1 presented more illumination area than Koch fractal level 0 (square) and Koch fractal level 1.

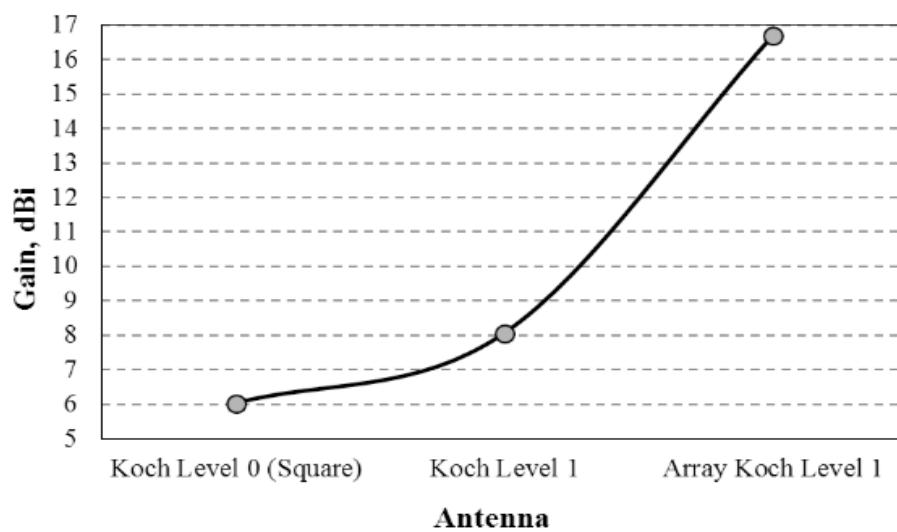


Fig.6. Comparison of gain of on-chip Koch fractal aperture-coupled antennas and array.

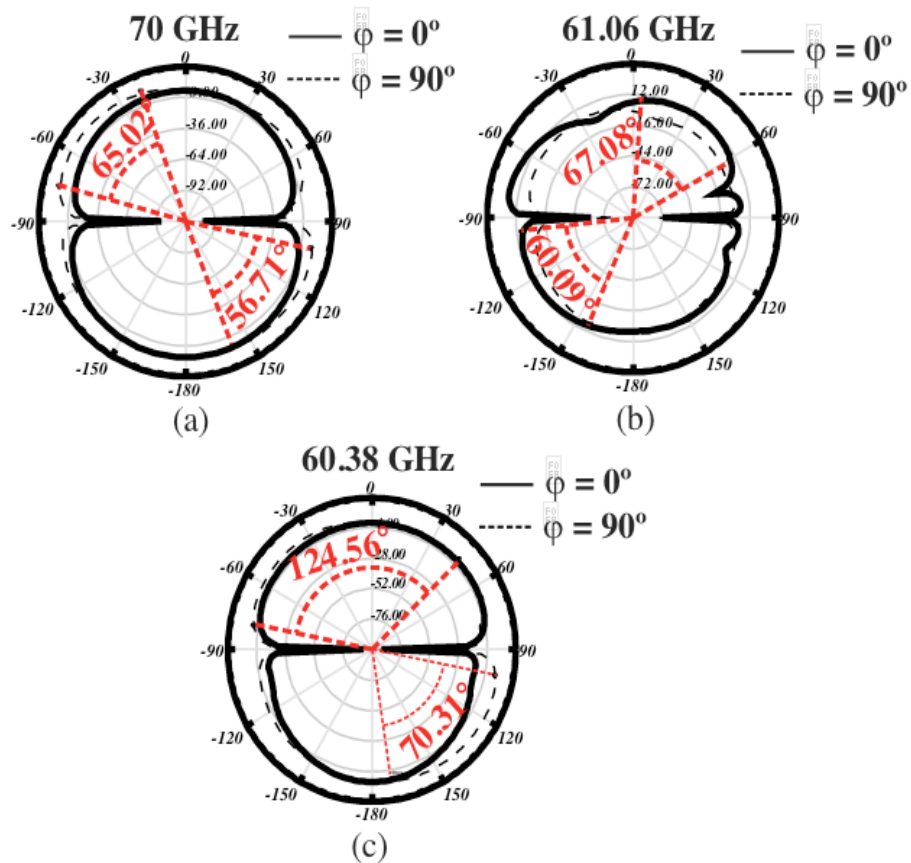


Fig.7. Radiation patterns of on-chip aperture-coupled antennas: a) Koch fractal level 0 (square); b) Koch fractal level 1; c) array Koch level 1.

## 5 | CONCLUSIONS

In this paper, it was presented the simulation of an on-chip array Koch fractal aperture-coupled antennas operating in industrial scientific and medical band at 60 GHz (57 GHz to 64 GHz). The proposed aperture-coupled antennas were designed with patch element built in aluminum with 2 micrometers, on two layers of silicon with 200 micrometers. All dimensions were calculated according to the effective wavelength for resonance frequency at 60 GHz to match to 50 Ohms. The use of Koch fractal and array Koch fractal allowed the development of aperture-coupled antennas of greater perimeter with more metal in patch elements promoting increase in the gain. Gain of array Koch fractal level 1 observed was of 16.7 dBi, increasing of 142% compared to the Koch fractal level 0 (square), with bandwidth operating in ISM Band, polarization closeto the circular polarization with axial ratio less of 2.22, and half power beamwidth greater than 124 degrees in resonance frequencies.

## 6 | ACKNOWLEDGMENT

We greatly appreciate the COPELE/UFCG, CAPES/COPEX, and PECS/UEMA by support and funding these institutions, without which this work would not be possible.

## REFERENCES

- NIKNEJAD, A. M.; HASHEMI, H. **mm-Waves Silicon Technology 60 GHz and Beyond**. New York: Springer, 2008.
- ZHAO, D.; REYNAERT, P. **CMOS 60-GHz and E-band Power Amplifiers and Transmitters**. New York: Springer, 2015.
- ZHANG, Y.; LI, B.; LIU, Y. **A procedure to design aperture-coupled stacked-microstrip antenna based on experiment**. *Microwave and Opt. Technol. Letters*, vol. 35, no. 3, 2002, pp. 244-247.
- KAM, D. G.; LUI, D.; NATARAJAN, A.; REYNOLDS, S.; CHEN, H.; FLOYD, B. A. **LTCC packages with embedded phased-array antennas for 60 GHz communications**. *IEEE Microwave and Wireless Components Letters*, vol. 21, no. 3, 2011, pp. 142-144
- STUTZMAN, W. L.; THIELE, G. A. **Antenna Theory and Design**. 3rd ed. Hoboken: Wiley, 2013.
- KIM, J. P. **Optimum design of an aperture-coupled microstrip patch antenna**. *Microwave and Opt. Technol. Letters*, vol. 39, no. 1, 2003, pp. 75-78.
- PACHI, S. K.; KAMARKAR, N. C.; LAW, C. L. **An EM-coupled dual-polarized microstrip patch antenna for RFID applications**. *Microwave and Opt. Technol. Letters*, vol. 39, no. 5, 2003, pp. 355-360.
- BERNARDO, M. B.; FREIRE, R. C. S.; FARIAS, A. K. Q.; ARAUJO, J. I. L.; MORAIS, S. M. A.; SERRES, A. J. R.; SILVA JUNIOR, P. F. **Aperture-coupled on-chip patch antennas for ISM band 60 GHz**. In: 2017 2<sup>nd</sup> International Symposium on Instrumentation Systems, Circuits and Transducers (INSCIT), 2017, Fortaleza, 2017. p. 1-4.
- FALCONER, K. **Fractal geometry: mathematical foundations and application**. 2. ed. Londres: Wiley, 2003. 337p.
- SULLIVAN, D. M. **Electromagnetic Simulation Using the FDTD Method**. IEEE Press, 2000.
- MANDELBROT, B. B. **The fractal geometry of nature**. 3. ed. Nova York: W. H. Freeman and Co., 1982. 468p.

## PROJETO E ANÁLISE DE UM ARRANJO LINEAR DE ANTENAS UTILIZANDO A CURVA FRACTAL DE KOCH

### **Elder Eldervitch Carneiro de Oliveira**

Universidade Estadual da Paraíba  
Departamento de Computação  
João Pessoa - PB

### **Pedro Carlos de Assis Júnior**

Universidade Estadual da Paraíba  
Departamento de Física  
Patos - PB

### **Marcelo da Silva Vieira**

Universidade Estadual da Paraíba  
Departamento de Física  
Patos - PB

### **Rodrigo César Fonseca da Silva**

Universidade Estadual da Paraíba  
Departamento de Física  
Patos - PB

**RESUMO:** A busca por dispositivos leve, compactos, de baixo custo e sem perda de desempenho vem impulsionando as pesquisas na área de antenas planares, principalmente devido as inúmeras aplicações presentes nos dias de hoje, tais como, internet das coisas e aplicações sem fio de uma maneira geral. A maioria dessas aplicações utilizam altas frequências, principalmente na faixa de 2,44 GHz e 5,80 GHz. Nesse contexto, uma antena é um dispositivo capaz de captar e transmitir ondas eletromagnéticas pelo espaço livre.

A geometria fractal aplicado em projetos de antenas nos permite obter dispositivos com dimensões reduzidas e comportamento multibanda. O projeto do arranjo de antenas proposto foi auxiliado pelo *Ansoft Designer™* que implementa o método dos momentos em uma série de análises numéricas. Posteriormente a fase inicial de projeto e simulação, o arranjo de antenas fractal foi construído e alimentado pela técnica de linha transmissão em microfita (TLM). Em um segundo momento, o arranjo proposto teve seus parâmetros de espalhamento (perda de retorno e coeficiente de onda estacionária) medidos com um analisador de redes vetorial modelo N5230A. Uma discussão a respeito dos resultados simulados e medidos foram apresentados e discutidos, apresentando uma boa concordância entre eles visando aplicações na faixa de frequência das micro-ondas, mais precisamente na banda de frequência ISM (*Industrial Scientific and Medical*).

**PALAVRAS-CHAVE:** Arranjo de Antenas, Fractal de Koch, Miniaturização, Micro-ondas.

**ABSTRACT:** The search for devices lightweight, compact, low-cost, performance-lossless has been boosting research in the area of planar antennas, mainly due to the numerous applications present today, such as internet of things and wireless applications in a way general. Most of these applications use high

frequencies, mainly in the range of 2.44 GHz and 5.80 GHz. In this context, an antenna is a device capable of capturing and transmitting electromagnetic waves through the free space. The fractal geometry applied in antenna designs allows us to obtain devices with reduced dimensions and multiband behavior. The proposed antenna array design was aided by Ansoft Designer™ that implements the moments method in a series of numerical analyzes. After the initial phase of design and simulation, the array of fractal antennas was constructed and fed by the technique of transmission line in microstrip (TLM). In a second moment, the proposed array had its scattering parameters (return loss and voltage standing wave ratio) measured with a vector network analyzer model N5230A. A discussion about the simulated and measured results was presented and discussed, showing good agreement among them for applications in the microwave frequency range, more precisely in the ISM (Industrial Scientific and Medical) frequency band.

**KEYWORDS:** Antenna array, Koch fractal, Miniaturization, Microwave.

## 1 | INTRODUÇÃO

Uma antena sempre foi um requisito primordial em todos os tipos de sistema de comunicação, seja na transmissão ou recepção dos sinais elétricos (BALANIS, 1997). Com o advento das tecnologias nos diversos serviços e setores de telecomunicações e seu crescimento considerável nas últimas duas décadas, principalmente no uso das tecnologias de comunicação sem fio, exemplo das tecnologias de terceira (3G), quarta (4G) e mais recentemente quinta geração, o chamado 5G (ainda em fase experimental na maior parte do mundo), passou-se a investir cada vez mais em dispositivos compactos, leves, com um custo acessível e sem perda de desempenho. Nesse contexto, as antenas planares surge como uma solução interessante para aplicações em altas frequências, em virtude de suas características atrativas, requisitos fundamentais aos dispositivos modernos. A utilização da geometria fractal aliado a projetos de antenas vem suprir parte dos requisitos primordiais nos sistemas de comunicação sem fio moderno, em virtude de sua geometria proporcionar um comportamento multibanda para a antena, além de ser possível reduzir as dimensões da antena em virtude do aumento de seu comprimento elétrico proporcionado pela utilização de um fractal em seu projeto (MANDELROT, 1983). Normalmente, antenas planares como é o caso de uma antena de microfita, apresenta baixo valor de diretividade para uma comunicação a longas distâncias. Diante desse fato, a junção de dois ou mais elementos irradiantes (patches), formando uma grande rede de alimentação denominada arranjos de antenas, surge como uma solução plausível e eficiente em projeto de antenas, com o intuito de maximizar as características de radiação do disposto proposto (ALAM, et al., 2009).

A seção II descreve o processo de geração da curva fractal de Koch de nível 2. Na seção III é descrito o projeto do arranjo proposto. Os resultados simulados e experimentais são apresentados na Seção IV. A seção V apresenta as conclusões desse artigo.



## 2 | GEOMETRIA FRACTAL – GERAÇÃO DA CURVA DE KOCH

A geometria fractal remete a uma figura, estrutura ou mesmo uma forma irregular, capaz de ser subdividida indefinidamente, na qual uma parte dessa subdivisão é uma réplica idêntica ao todo (MANDELBROT, 1983). O termo e por consequência a geometria fractal foi inicialmente proposta pelo matemático e pesquisador Benoit Mandelbrot na década de 1970, onde sua principal contribuição foi o estudo e descrição de um vasto grupo de formas complexas que possuem auto similaridade em sua estrutura geométrica, e cujas dimensões são fracionárias (MANDELBROT, 1983). Os percussores no estudo e implementação da geometria fractal em projeto de antenas foram Puente (1998) e Cohen (1997). A principal motivação para a implementação dessa geometria em projeto de dispositivos de micro-ondas se dá por duas características intrínsecas dessa geometria, são elas: i) a propriedade de auto similaridade, o qual é responsável pelo comportamento multibanda da antena, e ii) a propriedade de preenchimento do espaço, o qual é responsável pelo aumento do comprimento elétrico da antena e por consequência, consegue-se uma redução considerável nas dimensões do dispositivo.

Aliado a isso, diferentes curvas fractais podem ser obtidas por diferentes técnicas, dentre as quais se destaca o método IFS (*Iterative Function System*) para geração de curvas fractais. Nesse artigo em especial optamos por utiliza essa técnica por motivos de simplicidade na geração das curvas. A transformação usada,  $q_i: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ , para a obtenção da curva de Koch de nível 2 por meio do método IFS é dada pela Equação (1), (MANDELBROT, 1991). A Figura 1 ilustra os passos envolvidos no processo de construção da curva fractal de Koch de nível 2 a partir dos lados de um quadrado (forma inicial da curva).

$$q_i(x, y) = \frac{1}{3} \begin{pmatrix} \cos\theta_i & -\text{sen}\theta_i \\ \text{sen}\theta_i & \cos\theta_i \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} \begin{pmatrix} e_i \\ f_i \end{pmatrix} \quad (1)$$

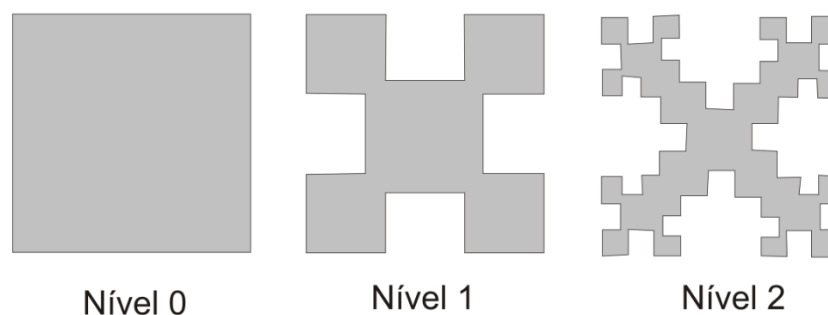


Figura 1 - Diferentes níveis da curva de Koch

### 3 | ESTRUTURA DO ARRANJO DE ANTENAS DE NÍVEL 2 PROPOSTO

As dimensões do arranjo de antenas proposto foram calculadas por meio de modelos matemáticos aproximados por meio das Equações (2) – (5). Em seguida, estas dimensões iniciais foram otimizadas pelo simulador *Ansoft Designer™* para análise precisa do comportamento eletromagnético das antenas (ANSYS Inc.). A estrutura projetada denominada de AKR2 foi baseada na curva fractal de Koch de nível 2 e foi excitada por linha de microfita, conforme mostra a Figura 2, para uma frequência de ressonância ( $F_r$ ) em 2,44 GHz, com aplicação direta na banda ISM (*Industrial Scientific and Medical*).

$$W = \frac{c}{2F_r} \sqrt{\frac{2}{\epsilon_r + 1}} \quad (2)$$

$$\epsilon_{reff} = \frac{\epsilon_r + 1}{2} + \frac{\epsilon_r - 1}{2} \left[ 1 + \frac{12h}{W} \right]^{-1/2} \quad (3)$$

$$\frac{\Delta L}{h} = 0,412 \frac{(\epsilon_{reff} + 0,3) \left( \frac{W}{h} + 0,264 \right)}{(\epsilon_{reff} - 0,258) \left( \frac{W}{h} + 0,813 \right)} \quad (4)$$

$$L = \frac{c}{2F_r \sqrt{\epsilon_r}} - 2\Delta L \quad (5)$$

Em que,  $c$  é a velocidade da luz,  $W$  e  $L$  são as dimensões do patch irradiante e  $\Delta L$  é um ajuste para compensar o efeito de borda (ALAM, et al., 2009). O cálculo do *inset-fed* ( $y_0$ ) foi calculado por meio da Equação (6) (RAMESH, et al. 2003).

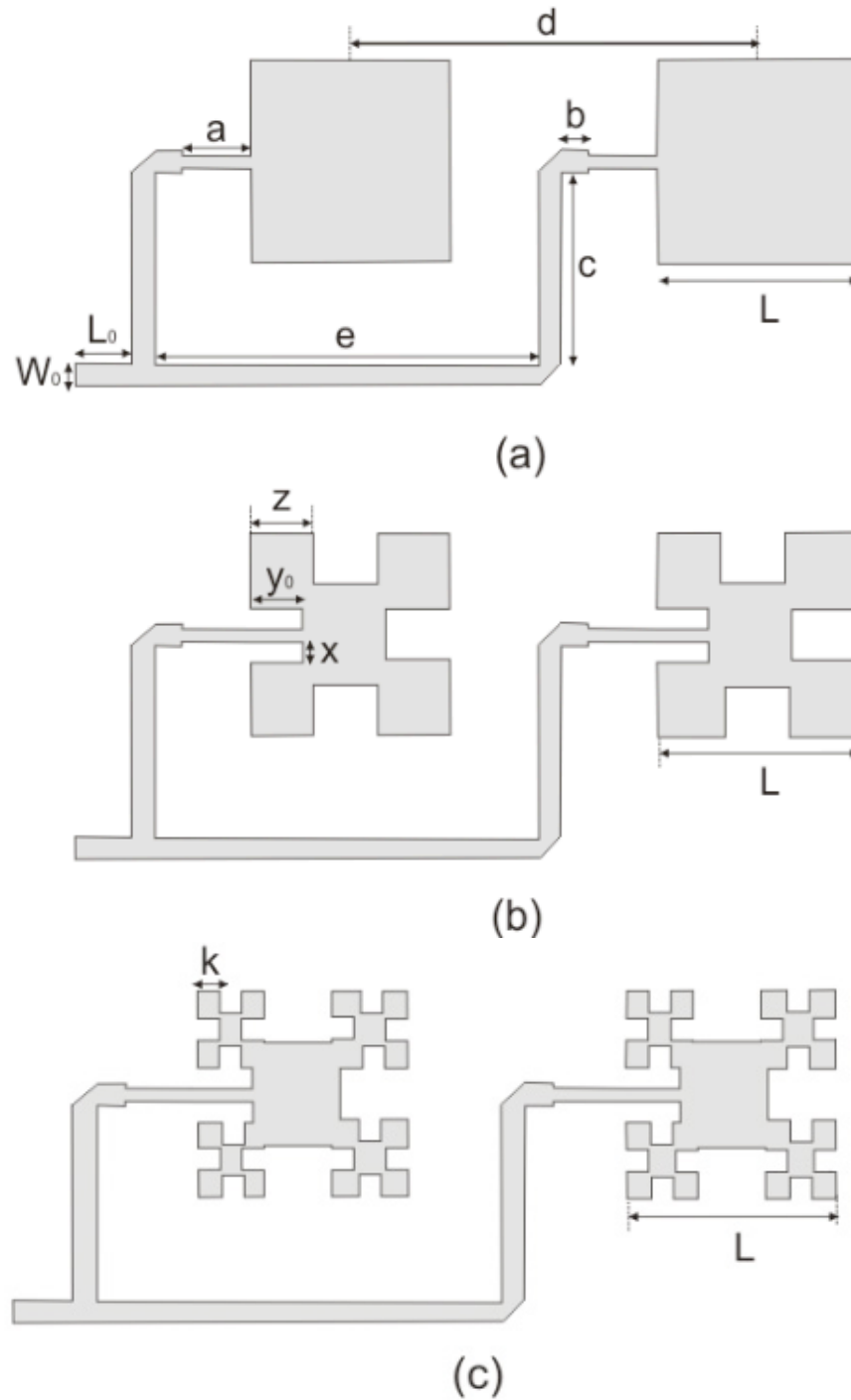


Figura 2 - Arranjos de antenas fractal: a) AKQ0 [8], b) AKQ1 [8] e c) AKR2

$$\frac{y_0}{10^{-4}} = \frac{L}{2} \left( \begin{array}{l} 0,001699\varepsilon_r^7 + 0,13761\varepsilon_r^6 - 6,1783\varepsilon_r^5 + 93,187\varepsilon_r^4 - \\ -682,69\varepsilon_r^3 + 2561,9\varepsilon_r^2 - 4043\varepsilon_r + 6697 \end{array} \right) \quad (6)$$

Uma excitação da antena foi realizada pela técnica de linha de transmissão em microfita, usando um conector SMA de  $50 \Omega$  [1]. O comprimento da linha de alimentação ( $L_0$ ) foi considerado  $\frac{1}{4}$  do comprimento de onda na frequência de operação de 2,44 GHz. Um protótipo de arranjo foi construído, simulado e medido (ver Figura 3), onde foi utilizada uma placa de circuito impresso de fibra de vidro (FR4), com as seguintes características: espessura ( $h = 1,5 \text{ mm}$ ), permissividade elétrica ( $\varepsilon_r = 4,4$ ) e tangente

de perda ( $\tan\delta = 0,02$ ).

Para a construção do arranjo de nível 2, partiu-se inicialmente de dois projetos executados anteriormente, onde na ocasião foram projetados dois arranjos de antenas, um de nível 0 e o outro de nível I, denominados respectivamente de AKQ0 e AKQ1, conforme mostrado em (RODRIGUES, et al., 2016). A distância ( $d$ ) entre os centros dos dois elementos radiantes foi de 59,0 mm o que corresponde a aproximadamente 0,5 do comprimento de onda guiado ( $\lambda_g$ ), essa distância mínima entre eles tem por finalidade evitar o acoplamento mútuo (RODRIGUES, et al., 2016). As dimensões do arranjo da figura 3 estão mostradas na Tabela 1.

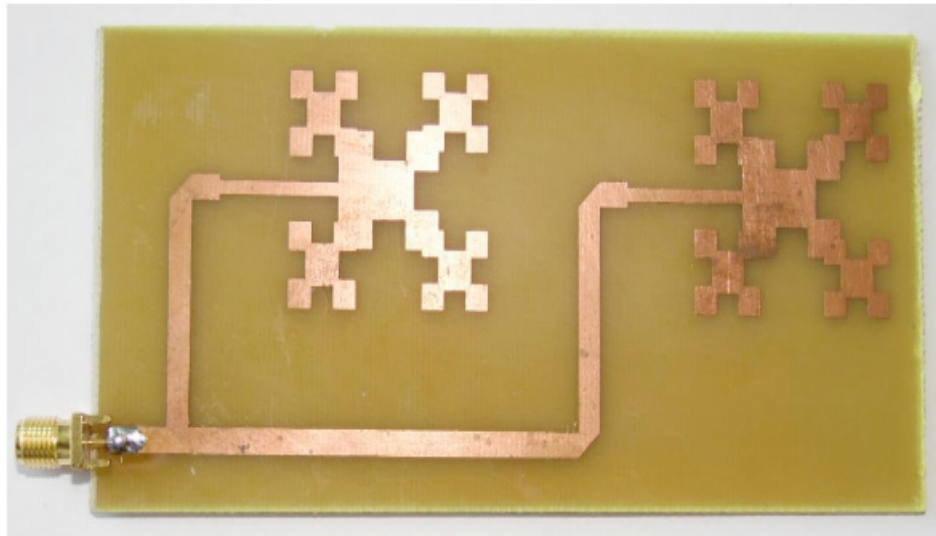


Figura 3 - Arranjos de antenas fractal nível 2 (AKR2) construído

Baseado no projeto inicial do arranjo proposto e por meio do cálculo do fator de compressão (CF), dado na Equação (7), as dimensões do arranjo AKR2 (ver Tabela 1) foram redimensionada com a finalidade de projetar um novo arranjo de antenas patch fractal de nível 2 com dimensões reduzidas, denominado AKR2 miniaturizado (AKR2 mini) para 2,44 GHz, pois com o aumento do comprimento elétrico da estrutura, conseguimos uma redução nas dimensões do arranjo.

Parâmetro	Dimensão (AKR2)	Dimensão (AKR2 mini)
L	28,7 mm	21,1 mm
d	59,0 mm	43,3 mm
$L_0$	8,37 mm	6,15 mm
$W_0$	3,12 mm	2,29 mm
e	55,9 mm	41,0 mm
b	4,0 mm	2,94 mm
c	27,1 mm	19,9 mm
a	17,4 mm	12,79 mm
k	3,19 mm	2,34 mm

Tabela 1 - Dimensões do arranjo AKR2 e AKR2 mini propostos

As dimensões desse novo arranjo projetado (AKR2 mini) com dimensões

reduzidas também são mostradas na Tabela 1. Baseado na Equação (7) para o cálculo do fator de compressão, nós conseguimos uma redução do tamanho do arranjo AKR2 mini de 26,5% quando comparado ao arranjo AKRQ0 (RODRIGUES, et al., 2016).

$$CF = \left( \frac{F_r(\text{GHz})}{2,45(\text{GHz})} \right) \times 100\% \quad (7)$$

#### 4 | RESULTADOS SIMULADOS E MEDIDOS PARA O ARRANJO PROPOSTO

Um protótipo de arranjo de antenas de microfita utilizando a geometria fractal foi construído, conforme mostrado anteriormente na Figura 3. Os primeiros resultados obtidos foram simulados no software *Ansoft Designer™* que implementa o método dos momentos (MoM) na análise da estrutura projetada.

Com o intuito de validar o modelo do arranjo proposto, a estrutura foi analisada experimentalmente e teve seus mais diversos parâmetros de espalhamento investigado, dentre eles: a Frequência de ressonância ( $F_r$ ), perda de retorno (RL), Largura de banda (BW), largura de banda percentual (BW%) e coeficiente de onda estacionária (VSWR), obtidos por meio de ensaios de laboratório no Instituto Federal de Educação, Ciência e Tecnologia da Paraíba (IFPB) por um analisador de redes vetorial modelo N5230A.

A perda de retorno (RL) no terminal de entrada da antena é definida como sendo a relação, em dB, entre as potências incidente e refletida nesse termina, onde a mesma pode ser expressa em termos no coeficiente de reflexão (S11), conforme Equação (8).

$$RL = -20\log |S11| \quad (8)$$

A Figura 4 ilustra os resultados obtidos para a perda de retorno simulado e medido para o arranjo AKR2. É possível observar uma boa concordância entre os resultados com baixas perdas de retorno ( $RL < -13$  dB), o que caracteriza um bom casamento de impedância para essa estrutura. A estrutura apresentou ainda uma frequência de ressonância medida igual a 1,65 GHz e uma estreita largura de banda igual a 25 MHz.

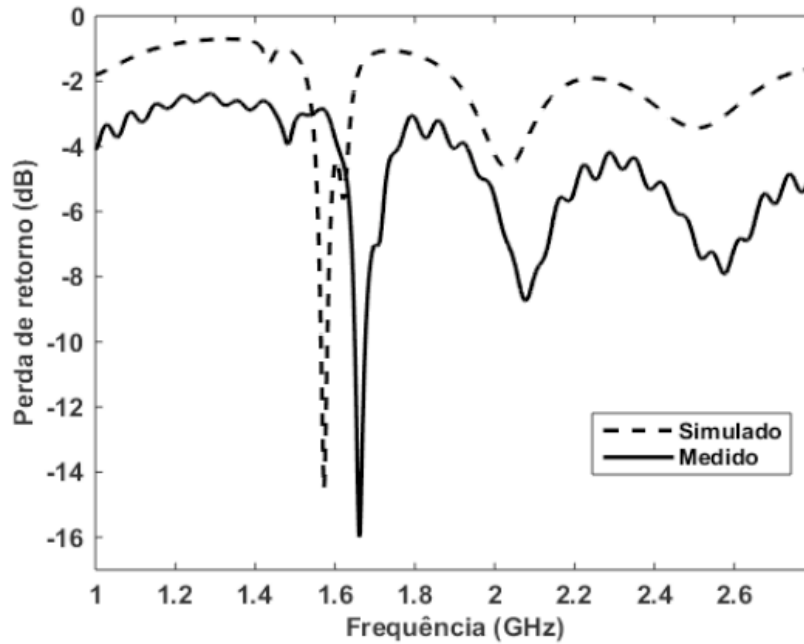


Figura 4 - Resultados de perda de retorno para o arranjo AKR2 proposto

Por sua vez, o coeficiente de onda estacionária é um parâmetro que mede o quanto de energia está voltando na linha de transmissão sob a forma de onda estacionária, e a mesma pode ser dada como sendo a relação entre os coeficientes de reflexão das ondas incidentes e refletidas sobre um terminal, conforme mostra a Equação (9).

$$VSWR = \frac{1+|S_{11}|}{1-|S_{11}|} \quad (9)$$

A Figura 5 ilustra o resultado medido do coeficiente de onda estacionária (VSWR) para o arranjo de antenas AKR2, onde podemos constatar um valor medido de 1,4. Valor este abaixo da condição limite de funcionamento, que corresponde a um VSWR igual a 2.

Uma característica intrínseca da geometria fractal é que a mesma quando aplicada a uma estrutura dessa natureza faz com que ocorra um aumento do comprimento elétrico da estrutura, ocasionado assim uma diminuição na sua frequência inicial de projeto. Dessa forma, se faz necessário reprojeter o arranjo para operar na frequência inicialmente desejada, que é a de 2,45 GHz, para que isso ocorra, foi aplicado um valor percentual de redução, neste caso foi 26,5% em suas dimensões, obtido pelo cálculo do fator de compressão, conforme Equação (7).

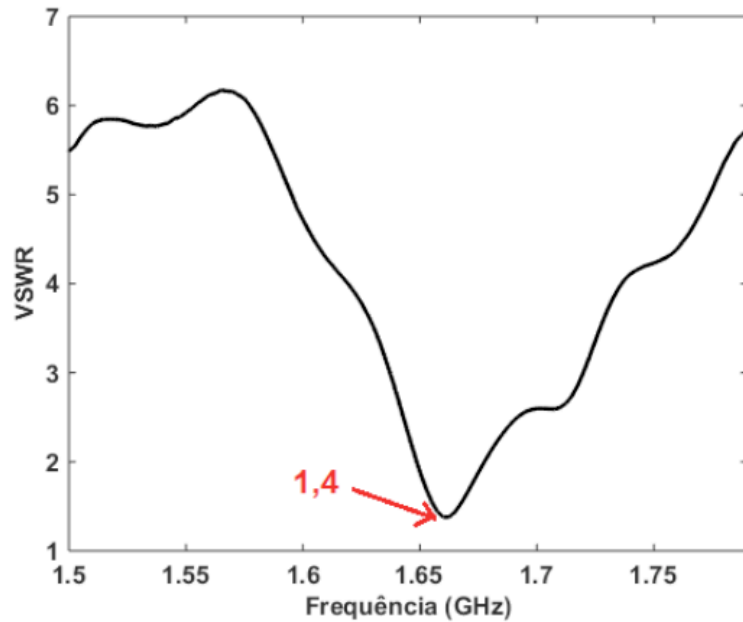


Figura 5 - Resultado experimental do coeficiente de onda estacionária (VSWR) para o arranjo de antenas AKR2

A Figura 6 seguinte ilustra o resultado de perda de retorno para o arranjo reprojeto para a frequência de projeto que é de 2,44 GHz. Uma boa concordância entre os resultados simulado e medido foi alcançada, onde o novo arranjo reprojeto apresentou uma frequência de ressonância medida de 2,45 GHz. Esse arranjo ainda apresentou uma baixa perda de retorno ( $RL < -25$  dB) caracterizando um bom casamento de impedância para estrutura, além de apresentar uma largura de banda (BW) de 80 MHz, superior ao primeiro arranjo com 100% de suas dimensões físicas.

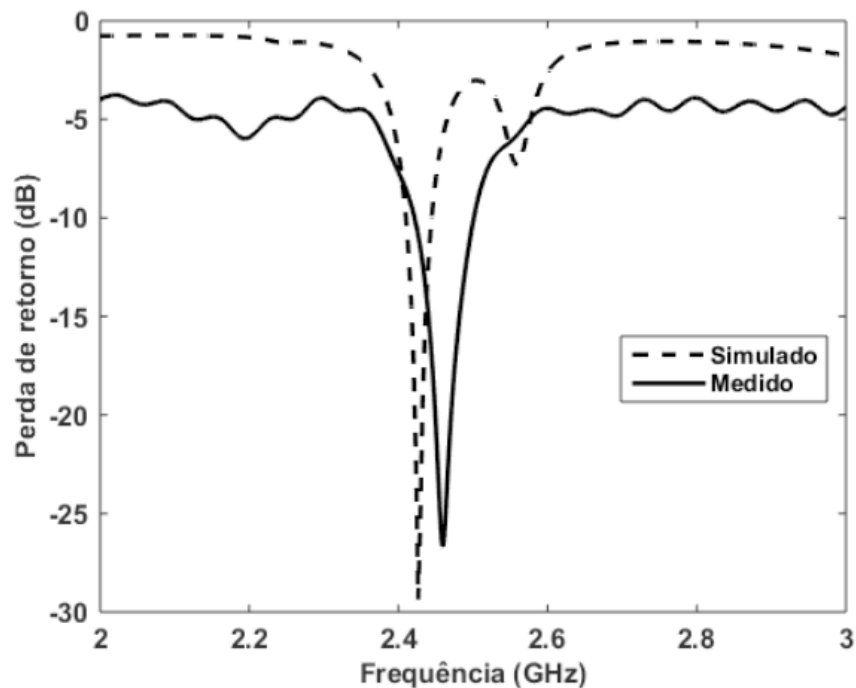


Figura 6 - Resultado de perda de retorno para o arranjo AKR2 mini



O resultado medido da impedância de entrada para o arranjo AKR2 mini é mostrado na Figura 7, onde o valor obtido foi de  $43,58 \Omega$ . Por sua vez, o resultado medido do coeficiente de onda estacionária para o arranjo miniaturizado AKR2 mini foi 1,1 e pode ser visto na Figura 8.

A Figura 9 ilustra um comparativo dos diagramas de radiação 2D, no Plano H, para os arranjos de antenas AKR2 e AKR2 mini, propostos neste artigo quando comparado ao arranjo AKQ0 convencional (RODRIGUES, et al. 2016).

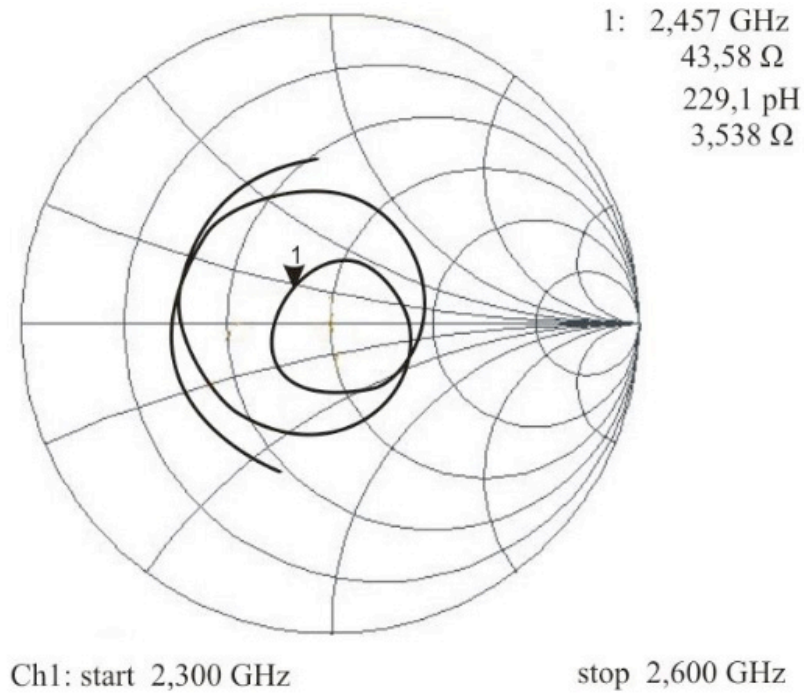


Figura 7 - Impedância de entrada medida na carta de Smith para o arranjo AKR2 mini, na frequência de 2,45 GHz

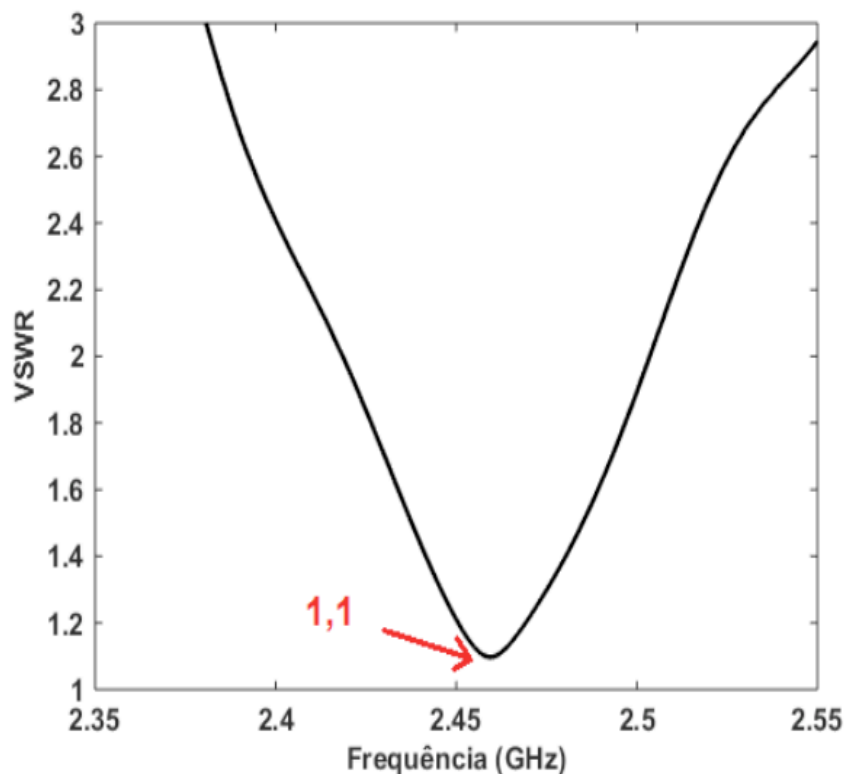


Figura 8 - Resultado de VSWR para o arranjo miniaturizado

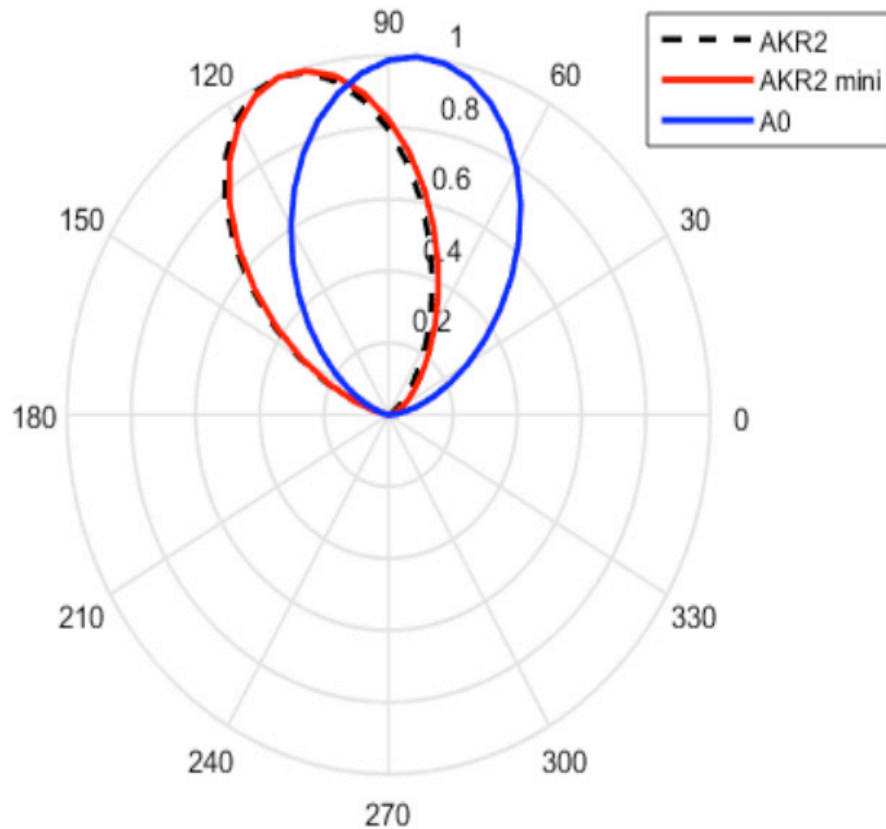


Figura 9 - Resultado simulado dos diagramas de radiação 2D (Plano H) para o arranjo de antenas AKR2, AKR2 mini e AKQ0 (RODRIGUES, et al., 2016)

Desse resultado podemos observar que não houve deformações nos diagramas, onde todos se mostraram bastante diretivos, apenas os diagramas dos arranjos fractais AKR2 e AKR2 mini se mostraram bem semelhantes e com um leve deslocamento para a direita em relação ao diagrama do arranjo AKQ0. Todos os resultados se encontram resumidos na Tabela 2.

Parâmetro	Arranjo AKR2		Arranjo AKR2 mini	
	Simulado	Medido	Simulado	Medido
$F_r$	1,57 GHz	1,65 GHz	2,44 GHz	2,46 GHz
BW	15 MHz	25 MHz	60 MHz	80 MHz
BW%	0,95%	1,51%	1,43%	3,25%
VSWR	—	1,4	—	1,1
RL (dB)	-14,5	-16,0	-29,5	-26,5

Tabela 2 - Resultados dos arranjos AKR2 e AKR2 mini propostos

## 5 | CONCLUSÃO

Nesse artigo foi proposto um arranjo de antenas de microfita utilizando a curva fractal de Koch de nível 2, onde o arranjo foi projetado para operar na banda de 2,44 GHz (banda ISM, que corresponde ao intervalo compreendido entre 2,40 – 2,483 GHz). O arranjo proposto apresentou uma redução de 26,5% em relação a um arranjo

retangular com elementos convencionais (AKQ0) propostos em (RODRIGUES, et al., 2016). A estrutura apresentou boas características de radiação na frequência de 2,44 GHz, com baixas perdas de retorno.

Uma boa concordância entre os resultados simulados e medidos foi verificada, validando assim o projeto. O erro estimado entre os resultados simulados e medido para o arranjo ARK2 mini foi de 0,8% e para o arranjo AKR2 foi de 5%, o que pode ser considerável tolerável, devido às limitações no processo de construção dos mesmos, bem com as condições não ideais do setup de medição utilizado.

Um ponto negativo do projeto foi à estreita largura de banda para os arranjos, os quais os autores pretendem contornar em futuros trabalhos nessa linha de pesquisa. Como trabalho futuro, pretende-se investigar e construir novos dispositivos utilizando outras curvas fractais e comparar seus resultados com outros artigos relacionados, verificando sua contribuição em termos de eficiência e capacidade de redução de suas dimensões, além de melhoramento em sua largura de banda. Dessa forma, pretende-se investigar o comportamento desse arranjo, verificando a sua viabilidade e/ou aplicabilidade em sistemas de comunicação sem fio, potencializando assim suas características de radiação, que exijam baixo perfil e sem perda de desempenho.

## REFERÊNCIAS

ALAM, M. M.; RAHMAN, Md. M.; GONI, Md. O. **Design and Performance Analysis of Microstrip Array Antenna**, Progress, Electromagnetics Research Symposium Proceedings, pp. 1837-1842, 2009.

Ansoft Designer (part of ANSYS Inc.) Online Available: [www.ansoft.com](http://www.ansoft.com).

BALANIS, Constantine. **Antenna theory-analysis and design**, ed. 2, John Wiley & Sons, 1997.

COHEN, N. **Fractal antenna applications in wireless telecommunications**, Proceedings of Electronics Industries Forum of New England, pp. 43-49, 1997.

MANDELBROT, Benoit. *The Fractal Geometry of Nature*, New York: W. H. Freeman, 1983.

MANDELBROT, Benoit. **Objectos Fractais: forma, acaso e dimensão**, Gradiva Publicações, Lisboa, 1991.

PUENTE, C.; ROMEU, J.; POUS, R; CARDAMA, A. **On the Behavior of the Sierpinski Multiband Fractal Antenna**, IEEE Transactions on Antennas and Propagation, vol. 46, pp. 517-524, 1998.

RAMESH, M.; KB, Y. **Design Formula for inset fed Microstrip Patch Antenna**, Journal of Microwaves and Optoelectronics, vol. 3, pp. 5-10, 2003.

RODRIGUES, R. R. M. OLIVEIRA, E. E. C.; OLIVEIRA, M. A.; SILVA, P. H. F.; SILVA, R. C. F. **Projeto e Análise de um Arranjo Pré-Fractal de Antenas Tipo Patch em Microfita**, 17º SBMO - Simpósio Brasileiro de Micro-ondas e Optoeletrônica e 12º CBMag - Congresso Brasileiro de Eletromagnetismo, Porto Alegre, 2016.

## FINDING REPEATER PLACEMENT FOR P2P WIRELESS LINKS WITH NLOS IN EXTREMELY MOUNTAINOUS REGIONS

**Alvaro Javier Ortega**

Pontifical Catholic University of Rio de Janeiro  
(PUC-Rio)

Center for Studies in Telecommunications  
(CETUC)

Rio de Janeiro - RJ, Brazil,  
E-mail: javier.ortega@cetuc.puc-rio.br.

**ABSTRACT:** A backhaul network consisting of p2p microwave links is the most cost-effective solution for rural Internet access, however, it can be a challenging design because the extreme mountainous geographical conditions make the links have Non-line out sight (NLOS) and be necessary the use of repeaters. Many relays can be implemented to overcome this difficulty. To the best of my knowledge, it does not exist an approach to finding the suboptimal repeaters placement in a computational way. This paper proposes the form to construct a cost function that represents the problem described before, and a genetic algorithm is implemented to find its local minimum at different scenarios.

**KEYWORDS:** Wireless backhaul network, rural zones, repeaters placement, digital elevation map.

### 1 | INTRODUCTION

Developed and undeveloped countries are concerned about providing Internet in rural zones (MICROSOFT, [S.d.]; MINTIC, [S.d.]; TIMES, [S.d.]). In this kind of scenarios, a traditional optical fiber backhaul network is too costly (DE SOUZA e colab., 2012), also the required installation routes are long and intricate, and some sites may be very difficult to reach. For this reason, wireless alternatives are implemented. Many approaches to the wireless backhaul for rural Internet access have been proposed. Most of them use point to point radio (p2p) links based on WiFi (KRETSCHMER e colab., 2011; NIEPHAUS e colab., 2012; ZAMBRANO e colab., 2012), WiMAX (CHAI e colab., 2011), or proprietary technologies (DE SOUZA e colab., 2009). These works concentrate on describing specific implemented solutions, but they do not address the network design problems, nor optimization procedures. In (BERNARDI e colab., 2011) a tool for incremental planning of wireless networks is presented, but is limited to network expansion rather than initial design. Other solutions for rural scenarios are mesh networks, which can be implemented with WiFi or WiMAX standards (HENKEL e colab., 2011), and Satellite options can be considered as well.

In rural scenarios is difficult to design

wireless backhaul networks because line of sight (LOS) between nodes is not always achievable. LOS is reached in a radio link if there are no obstacles in the straight line between the antenna sites and the first Fresnel zone is cleared. Most of WiFi-based p2p links work in the 5 GHz U-NII bands in order to avoid interference at 2.4 GHz, and to take advantage of a smaller Fresnel zone clearance. In practice higher towers are used at the link endpoints to obtain the first zone clearance. However, it can be a challenge because the extreme mountainous geographical conditions (i.e., extremely hilly terrain characterized by strong changes in terrain height over small areas) make the links have Non-line of sight (NLOS) and be necessary the use of repeaters. Many relays can be implemented to overcome this difficulty.

The designer has to be concerned about in finding the repeaters placement such that minimize: relays number, tower heights and links distances. To the best of my knowledge, it does not exist an approach to find the repeaters placement in a computational way. This paper proposes the form to construct a cost function that represents the problem described before. A digital elevation map is used to simulate the rural scenarios. Clearly this problem cannot be solved deterministically therefore, with the aim to find a suitable solution a heuristic method is implemented in different scenarios. In addition, the complexity of this problem is proportional to the factorial number of candidate positions of repeaters, , it is discussed with more detail in the Section 2. The most popular heuristic to resolves NP-hard problems is the Genetic Algorithm (GA), for that reason it was chosen to deal this problem (DAVIS, 1991).

The remainder of this paper is organized as follows: Section 1 describes a criterion to establish if a link is feasible or not; Section 2 presents my approach to find the suboptimal repeaters placement; Section 3 contains numerical results at different scenarios; and finally, in Section 4 appropriate conclusions wrap up this paper.

## 2 | CRITERION OF LINK FEASIBILITY

Link feasibility can be approached in terms of antenna types, gain, opening angle, etc., however, it can be addressed in a simpler way. For establishing the link connectivity, a criterion of LOS is described in the subsection 2.1. On the other hand, the wireless link performance is related directly by the received signal power, therefore, in order to increase this power, the link distance has to be decreased, thus a distance constraint is defined in the subsection 2.2.

### 2.1 Criterion of LOS

LOS is reached in a radio link if there are no obstacles in the straight line between the antenna sites and the first Fresnel zone is cleared as stated before. In practice higher towers are used at the link endpoints to obtain the first zone clearance. The general equation for calculating the Fresnel zone radius at any point in between the

endpoints of the link is the following:

$$F_n(\mathbf{p}) = \sqrt{\frac{n\lambda d_{1-p}d_{p-2}}{d_{1-p}+d_{p-2}}} \quad (1)$$

where  $F_n(\mathbf{p})$  is the  $n$ -th Fresnel zone radius,  $d_{1-p}$  is the distance of one end from  $p$ ,  $d_{p-2}$  is the distance of  $p$  from the other end, and  $\lambda$  is the wavelength of the transmitted signal.

Let  $h_{n_1}$  and  $h_{n_2}$  be the towers heights at the link endpoints such that LOS is reached. To limit the search of these values a finite set of available tower heights is defined  $\mathcal{H} = \{h_1, h_2, \dots, h_m\}$  and the best option from it is found through exploring brute force.

Algorithm 1 presents a way to find  $h_{n_1}$  and  $h_{n_2}$  values through the description of the *towersheights* function, which depends on both (i) the link endpoints,  $\mathbf{n}_i \in \mathbb{R}^2$ ,  $i=1,2$ , which are given by its corresponding latitude and longitude; and (ii) the digital map  $\mathcal{M}$  (NASA, 2014), such as the one presented in the Figure 1. At the beginning of the algorithm a set of candidate heights of towers,  $\mathcal{O}$ , is defined as an empty set. Then, the *elevationprofile* function extracts the elevation profile,  $\mathbf{p} \in \mathbb{R}^{n_p}$ , between the endpoints,  $n_1$  and  $n_2$ , from the digital map  $\mathcal{M}$ , i.e., this function extracts the heights on a straight line connecting  $n_1$  and  $n_2$  from  $\mathcal{M}$ .  $n_p$  represents the number of heights samples in this straight line. The first Fresnel zone radius is computed at all points between the link endpoints according to (1) and a vector  $\mathbf{f} \in \mathbb{R}^{n_p}$ , is arranged with these values. Different straight lines,  $\mathbf{l} \in \mathbb{R}^{n_p}$ , are generated such that connects the heights of link endpoints plus his respectively height tower using  $\mathcal{H}$ . If the curve,  $\mathbf{c} = \mathbf{l} - \mathbf{f}$ , is greater than the elevation profile,  $\mathbf{p}$ , then the set  $\mathcal{O}$  is updated adding the new candidate of towers heights. Finally, the couple of towers heights is selected such that sum between them be minimum. If the set of candidate heights of towers is empty,  $h_{n_1}$  and  $h_{n_2}$  are set with absurdly high values, which is going to help to avoid these repeaters positions with NLOS in the optimization procedure.

Figure 2 presents the elevation profile of a link with LOS and NLOS using the Algorithm 1.

---

**Algorithm 1** Searching towers heights to obtain LOS using a finite set of tower heights available

---

1: *towersheights* function  
**Inputs:**  $\mathbf{n}_1, \mathbf{n}_2, \mathcal{M}$   
**Outputs:** cost  $h_{n_1}, h_{n_2}$   
2: Starting  $\mathcal{O} = \emptyset$   
3:  $\mathbf{p} = \text{elevationprofile}(\mathbf{n}_1, \mathbf{n}_2, \mathcal{M})$   
4: Compute  $\mathbf{f} \in \mathbb{R}^{n_p}, f_k = F_1(k), k = 1, \dots, n_p$   
5: **for**  $i = 1, \dots, m$   
6:     **for**  $j = 1, \dots, m$



```

7:       $l =$  straight line connecting  $p_1 + h_i$  and  $p_{n_p} + h_j$ 
8:       $c = l - f$ 
9:      if  $c_k > p_k, k = 1, \dots, n_p$ 
10:          $\mathcal{O} = \mathcal{O} \cup [h_i, h_j]$ 
11:      end
12:  end
13: end
14: if  $\mathcal{O} = \emptyset$ 
15:   $[h_{n_1}, h_{n_2}] = [10^3, 10^3]$ 
16: else
17:   $[h_{n_1}, h_{n_2}] = \arg \min_{h_i, h_j \in \mathcal{O}} (h_i + h_j)$ 
18: end
19: return  $h_{n_1}, h_{n_2}$ 

```

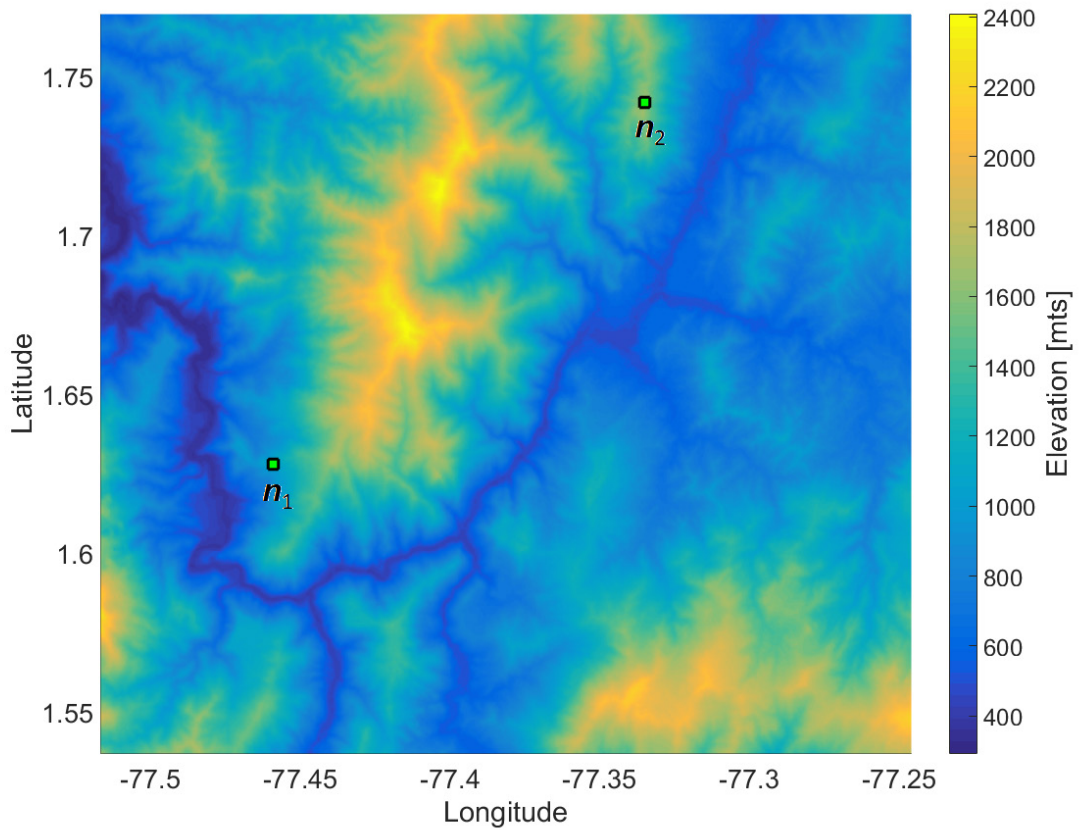
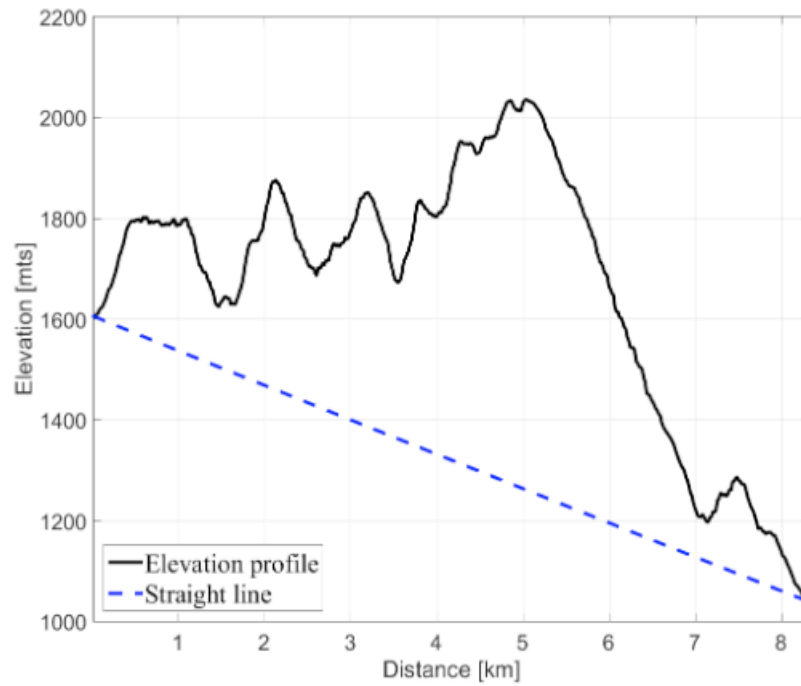
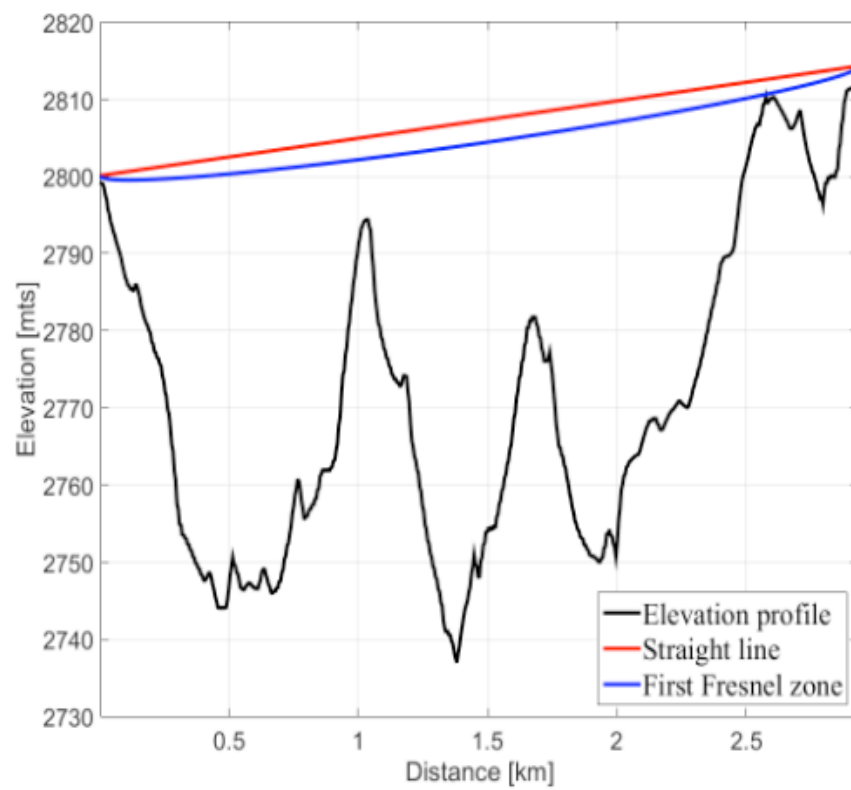


Figure 1. A scenario for backhaul network design in Nariño, Colombia. Altitude over the sea level (in meters) is represented by colors, showing the tremendous terrain elevation changes and the difficult for designing p2p links.





(a) Link with NLOS. (b)



(b) link with LOS,  $h_{n,1} = 1$  and  $h_{n,2} = 1$ .

Figure 2. Elevation profile of links with NLOS and LOS using the Algorithm:

## 2.2 Distance constraint

Other important constraint in designing of the wireless backhaul networks is link distance because the following reasons:

- The free-path loss is proportional to the square of distance and the commu-

nication quality depends directly on the received signal power.

- The access points used in p2p links cannot operate with high transmit power due to the constraint hardware and the policies of use of radio spectrum frequency.
- The alignment of the higher directivity antennas at long distance can be difficult to reach.
- In order to decrease the search space.

Using the link distance constraint,  $d_{max}$ , it is possible to limit the search space of the repeaters position to an elliptic region. Consider the geometry of an ellipse shown at the Figure 3 and an unfeasible link that just needs a repeater to obtain connection between the endpoints,  $n_1$  and  $n_2$ . Taking the link endpoints as the ellipse focus, we can obtain the following relations:

- $fe_i = n_i, i = 1,2$
- $r$  = repeater position
- $d_{max} = c+a$
- $d_c = d_1 + d_2$

Let us define the points into the ellipse as a set of candidate positions of repeaters,  $\mathcal{C} = \{c_1, c_2, \dots, c_{nc}\}$ . For the cases where more than one repeater are needed, we use the next approximation to construct the elliptical constraint,  $d_c \approx (n_r + 1)d_{max}$ , where the  $n_r$  is the number of repeaters to be found.

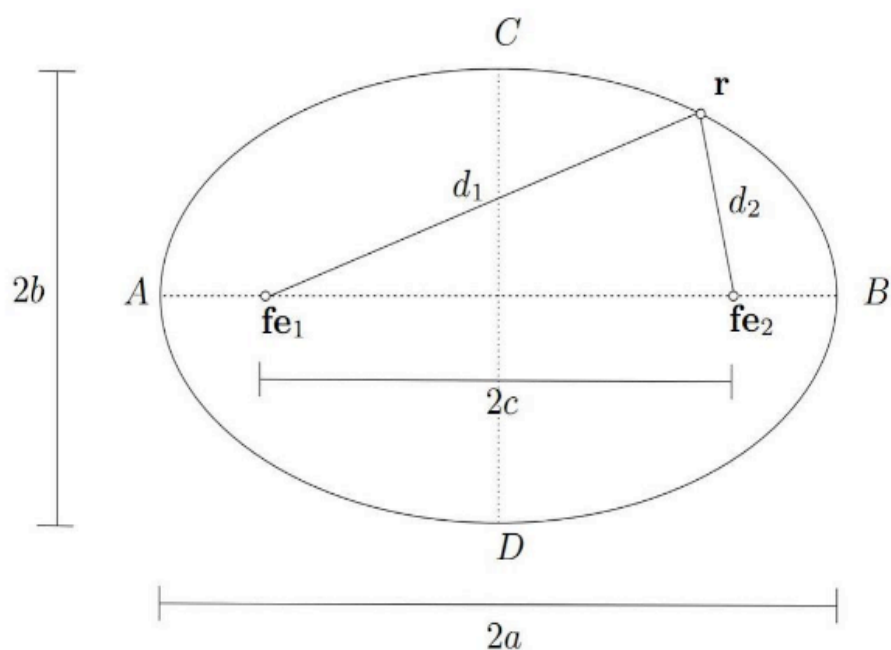


Figure 3. Geometric of an ellipse

### 3 | DESIGN OF THE PROBLEM FUNCTION

A backhaul network consistent of p2p microwave links is the most cost-effective solution for rural Internet access as stated before. This section describes a design methodology for such network in mountainous regions, motivated by the needing of many developing countries. Subsection 3.1 details the construction of the cost function and the implemented heuristic method is described in the Subsection 3.2.

#### 3.1 Cost function construction

Consider two nodes,  $\mathbf{n}_1$  and  $\mathbf{n}_2$ , be the endpoints of an unfeasible link in a rural scenario, which is defined by a digital map  $\mathcal{M}$  (NASA, 2014) (see Figure 1) . Algorithm 2 describes the cost function for the searching case of one repeater position represented by the vector  $\mathbf{r}$ . The cost of this function is given by the links distances and towers heights, thus, the next step is to apply a heuristic method to minimize it. Many scenarios cause that an unfeasible link does not realizable with just one repeater, therefore a similar algorithm to Algorithm 2 with more repeaters into account have to be implemented.

---

**Algorithm 1** Description of the cost function to a repeater searching

---

1: **Inputs:**  $\mathbf{r}$

**Outputs:**  $cost$

2: Find distances

$d_{n_1-r}$  = distance between  $n_1$  and  $r$  in km

$d_{r-n_2}$  = distance between  $r$  and  $n_2$  in km

$D_{total} = d_{n_1-r} + d_{r-n_2}$

3: Find tower heights

$[h_{n_1}, h_{r_1}] = towerheights(\mathbf{n}_1, \mathbf{r}, \mathcal{M})$

$[h_{r_2}, h_{n_2}] = towerheights(\mathbf{r}, \mathbf{n}_2, \mathcal{M})$

$h_r = \max(h_{r_1}, h_{r_2})$

$H_{total} = h_{n_1} + h_r + h_{n_2}$

4: Ideal values

$D_{ideal}$  = distance between  $n_1$  and  $n_2$  in km

5: Cost function value

$cost = \sqrt{(D_{total} - D_{ideal})^2 + (H_{total} - H_{ideal})^2}$

---

Note that if we want to find the optimum repeater position, we have to explore all elements of the set  $\mathcal{C}$ , i.e., test  $n_c$  candidate repeater positions. However, typically we need more than a relay, so, if we consider  $n_r$  repeaters to be found, we have to test all combinations of  $n_r$  elements into a set of  $n_c$  elements times the permutation of  $n_r$ , i.e.,  $\frac{n_c!}{(n_c - n_r)!} n_r!$ , which is computationally prohibitive. Therefore, heuristic techniques must be implemented.

### 3.2 Genetic algorithm

Genetic Algorithms (GAs) are search and optimization techniques inspired by two biological principles namely the process of natural selection and the mechanics of natural genetics. GAs manipulates not just one potential solution to a problem but a collection of potential solutions. This is known as population. The potential solution in the population is called “chromosomes”. These chromosomes are encoded representations of all the parameters of the solution (DAVIS, 1991). Each chromosome is compared to other chromosomes in the population through an awarded fitness rating that indicates how successful a chromosome is. The GA uses genetic operators or evolution operators such as crossover and mutation for the creation of new chromosomes from the existing ones in the population. The selection mechanism for parent chromosomes takes the parent fitness into account. This will ensure that the better solution will have a higher chance to procreate and donate their beneficial characteristic to their off spring. How well an individual performs a task is measured and assessed by the objective function. The objective function assigns to each individual a corresponding value called its fitness. The fitness of each chromosome is assessed and a survival of the fittest strategy is applied (DAVIS, 1991). In this project, the cost function described before is the objective function and the element index of the candidate set,  $\mathcal{C}$ , is taken as chromosome, i.e., each chromosome  $\in \{1, 2, \dots, n_c\}$ .

## 4 | NUMERICAL RESULTS

This section presents some simulations results of the proposed method. All implemented scenarios belong to Nariño-Colombia, because this region belongs to the most mountainous zone in all South America (i.e., the Andes mountains), therefore this gives a big challenge in the design. The number of repeaters is limited to 2,  $n_r=2$ . The frequency transmission is 5 GHz. The set of tower heights available is  $\mathcal{H}=\{0, 1, 2, \dots, 10\}$ . Table 1 resumes the main settings parameters of the GA to the two objective functions.

Figure 4 shows the found solution to the challenging scenario given in the Figure 1, where the blue markets represent the repeaters positions. To this scenario two repeaters were necessary. Another scenario with similar complexity is presented in the Figure 5. From all results is very easy to note that the algorithm searches effectively the shortest route to connect the links endpoints. Table 2 contains the description of interest points of the above designs in terms of latitude, longitude and needed tower height.

	a repeater	two repeaters
Population size	50	150
CrossoverFraction	0.7	0.7

Generations	1000	3000
StallGenLimit	$10^5$	$10^5$

Table 1. GA's main settings for the search of

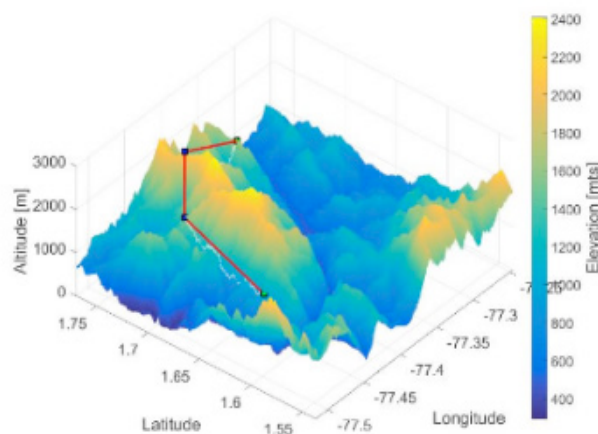
Interest points		Latitude	Longitude	Tower height
Design 1	$n_1$	1.6279	-77.4589	8
	$r_1$	1.7149	-77.4447	1
	$r_2$	1.7466	-77.3998	1
	$n_2$	1.7419	-77.335	8
Design2	$n_1$	1.2301	-77.2873	1
	$r_1$	1.2237	-77.3224	1
	$r_2$	1.2949	-77.2322	1
	$n_2$	1.3821	-77.1566	1

Table 2. Description of p2p wireless links designs

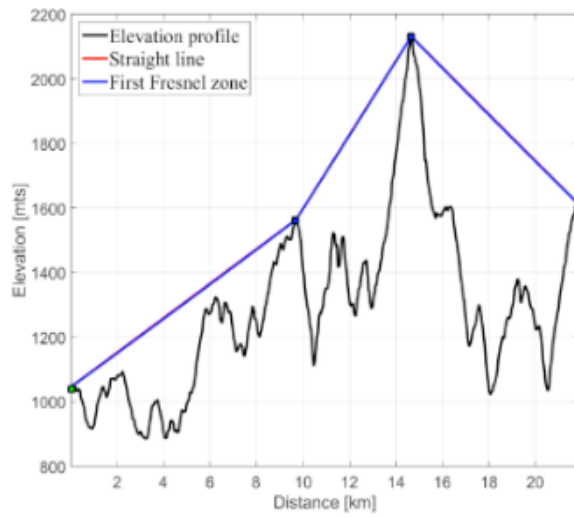
## 5 | CONCLUSIONS

This paper proposes a new and unique method to find the suboptimal repeaters placement to connect rural zones using p2p wireless links in a computational way. In the simulations were chosen some very mountainous scenarios to challenge the backhaul design.

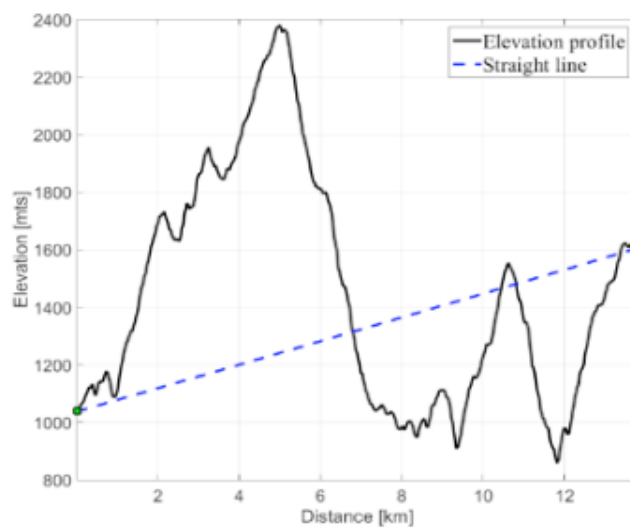
The results show that the proposed method can give good solutions to make workable an infeasible link; nevertheless, it does not mean that we can obtain the final design through this tool. We must be aware that the digital map just represents an approximation of the real scenario because of its limited resolution (the new data have been released with a 1 arc-second, or about 30 meters (NASA, 2014), which, although it can be improved using interpolation still remain inaccurate) and data acquisition time (many changes could have happened in the interest region e.g., landslides or growth of large trees which are not in the database).



(a) 3D visualization of the wireless links connection design

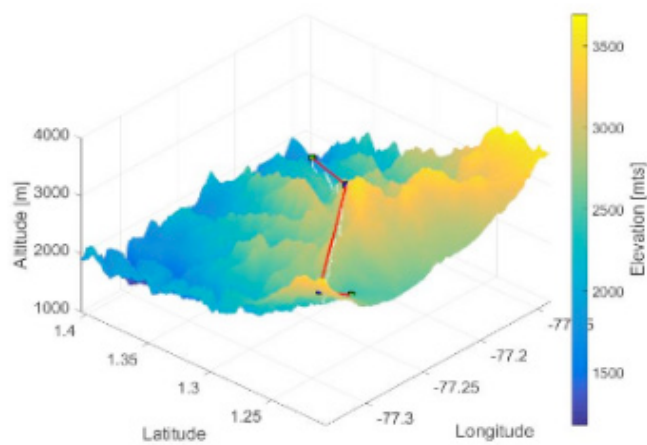


(b) Elevation profile from  $n_1$  to  $n_2$  with relays

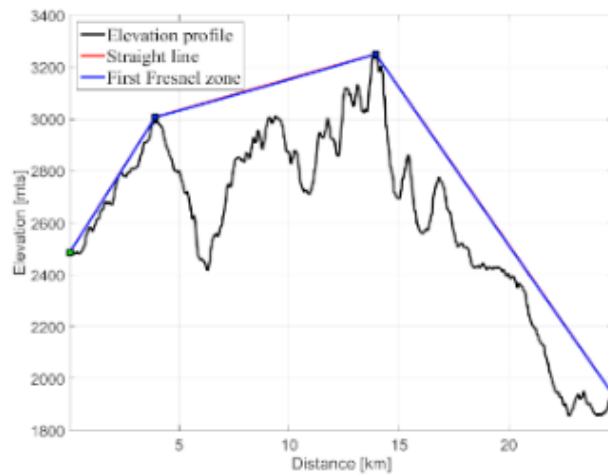


(c) Elevation profile from  $n_1$  to  $n_2$  without relays

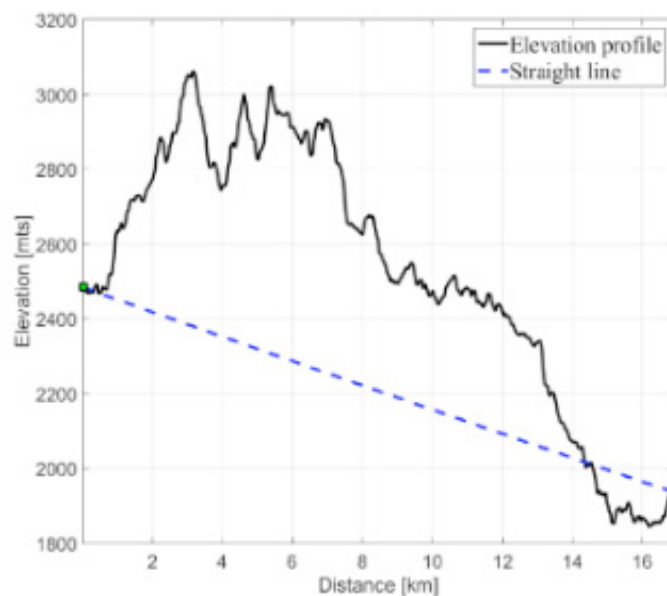
Figure 4. Design 1



(a) 3D visualization of the wireless links connection design



(b) Elevation profile from  $n_1$  to  $n_2$  with relays



(c) Elevation profile from  $n_1$  to  $n_2$  without relays

Figure 5. Design 2

On the other hand, to avoid prohibitive places as private zones or very difficult to reach, we must penalize this points adding a big value at his cost function result or removing them from the set .

Given that the GA difficultly can reproduce the same results, each execution of the algorithm gives a new possible design, this is very useful since several design alternatives can be obtained and thus, the network designer only have to select which is the better realizable backhaul from them.



## REFERENCES

- BERNARDI, G e colab. **IncrEase: A tool for incremental planning of rural fixed Broadband Wireless Access networks**. 2011, [S.l: s.n.], 2011. p. 1013–1018.
- CHAI, P S N e CHUNG, Kah-Seng e CHAN, K S. **Two-channel two-transceiver IEEE 802.16 wireless backhaul**. 2011, [S.l: s.n.], 2011. p. 1029–1033.
- DAVIS, Lawrence. **Handbook of genetic algorithms**. 1991, [S.l: s.n.], 1991.
- DE SOUZA, L V e colab. **Multimedia transmission on Amazon region using wireless broadband networks**. 2009, [S.l: s.n.], 2009. p. 1–4.
- DE SOUZA, L V e COSTA, J.C.W.A. e FRANCES, C R L. **Integrated solutions for broadband access in Brazilian Amazon Rural areas**. 2012, [S.l: s.n.], 2012. p. 36–39.
- HENKEL, D e colab. **Connecting the unconnected: Economic constraints and technical requirements towards a back-haul network for rural areas**. 2011, [S.l: s.n.], 2011. p. 1039–1044.
- KRETSCHMER, M e colab. **QoS-Aware Wireless Back-Haul Network for Rural Areas with Support for Broadcast Services in Practice**. 2011, [S.l: s.n.], 2011. p. 758–764.
- MICROSOFT. **A rural broadband strategy: connecting rural America to new opportunities**. [S.d.], [S.l: s.n.], [S.d.]. Available: <<https://goo.gl/Bf8XvA>>.
- MINTIC. **Inhabitants of rural areas of the Coffee Region released 31 Kiosks Live Digital**. [S.d.], [S.l: s.n.], [S.d.]. Available: <<https://goo.gl/XvJHYy>>.
- NASA. **Shuttle Radar Topography Mission**. 2014, [S.l: s.n.], 2014. Available: <[www2.jpl.nasa.gov/srtm](http://www2.jpl.nasa.gov/srtm)>.
- NIEPHAUS, C e KRETSCHMER, M e JONAS, K. **QoS-aware Wireless Back-haul network for rural areas in practice**. 2012, [S.l: s.n.], 2012. p. 24–29.
- PANIGRAHI, D e colab. **Minimum Cost Topology Construction for Rural Wireless Mesh Networks**. 2008, [S.l: s.n.], 2008. p. 771–779.
- QUADRI, A e colab. **Next generation communication technologies: Wireless mesh network for rural connectivity**. 2011, [S.l: s.n.], 2011. p. 1019–1023.
- TIMES, Finantial. **Broadband investors pour 500m into Britains countryside**. [S.d.], [S.l: s.n.], [S.d.]. Available: <<https://goo.gl/6zx1jK>>.
- ZAMBRANO, A e colab. **Municipal Communications Infrastructure for Rural Telemedicine in a Latin-American Country**. Latin America Transactions, IEEE (Revista IEEE America Latina), v. 10, n. 2, p. 1489–1495, 2012.

## NOVA ARQUITETURA DE DEMODULADOR $\pi/3$ -BPSK PARA OS SATÉLITES DO SISTEMA BRASILEIRO DE COLETA DE DADOS

**Flavia Vasconcelos Maia**

Instituto Federal de Educação, Ciência e Tecnologia do Ceará (IFCE)  
Fortaleza – CE

**Antonio Macilio Pereira de Lucena**

Instituto Nacional de Pesquisas Espaciais (INPE)  
Universidade de Fortaleza (UNIFOR)  
Fortaleza – CE

**Francisco de Assis Tavares Ferreira da Silva**

Instituto Nacional de Pesquisas Espaciais (INPE)  
Eusébio – CE

**RESUMO:** Este artigo apresenta o projeto de uma nova arquitetura de demodulador  $\pi/3$ -BPSK totalmente digital a ser utilizado como parte de um transponder regenerativo de bordo, o qual está sendo desenvolvido para modernização do Sistema Brasileiro de Coleta de Dados. Este demodulador apresenta características inovadoras, pois além de ser uma solução totalmente digital, com processamento sobre o sinal discreto, os circuitos de recuperação de portadora e de sincronização de símbolo são inéditos para esta aplicação. O trabalho também apresenta os resultados de desempenho, em termos de taxa de erros de bits e de tempo de aquisição, obtidos através de simulação computacional. As medidas demonstram que a arquitetura proposta atende com certa margem as especificações do sistema, tanto para o

tempo de sincronização como também para a taxa de erros de bits, ficando apenas 1 dB abaixo do limite teórico na condição em que a taxa de erros de bits é  $10^{-4}$ .

**PALAVRAS-CHAVE:** Demodulador  $\pi/3$ -BPSK, PCD, SBCD, sincronismo de símbolo, recuperação de portadora.

**ABSTRACT:** This paper presents the design of a new fully digital  $\pi/3$ -BPSK demodulator architecture to be used as part of an on-board regenerative transponder, which is being developed for the modernization of the Brazilian Data Collection System. This demodulator has innovative features because besides having an architecture implemented via discrete signal processing, the carrier recovery circuits and the symbol synchronization circuits are unprecedented for this application. The work also presents the performance results, in terms of bit error rate and acquisition time, obtained through computational simulation. The measurements demonstrate that the proposed architecture meets the system specifications for both the synchronization time as well as the bit error rate, being only 1 dB below the theoretical limit in the condition where the bit error rate is  $10^{-4}$ .

**KEYWORDS:** Demodulator  $\pi/3$ -BPSK, PCD, SBCD, symbol synchronization, carrier recover.

## 1 | INTRODUÇÃO

O Sistema Brasileiro de Coleta de Dados (SBCD), desenvolvido e operado pelo Instituto Nacional de Pesquisas Espaciais (INPE), é composto do segmento solo e do segmento espacial, e tem como função retransmitir informações ambientais relevantes do território brasileiro, tais como: dados de hidrologia, de meteorologia, de oceanografia, sobre a qualidade d'água, sobre a química de atmosfera, dentre outras. O sistema atualmente possui mais de 1.000 Plataformas de Coleta de Dados (PCDs), espalhadas pelo país, que transmitem as informações para os Satélites de Coleta de Dados (SCD-1 e SCD-2) e para o satélite CBERS (TIKAMI et al., 2013). Os transponders de coleta de dados atuais funcionam como repetidores analógicos e retransmitem os sinais para as estações de recepção em solo, de Cuiabá e de Alcântara, que por sua vez enviam os dados para processamento no Centro de Missão em Natal, RN (PÉCALA, 2005).

Devido à transmissão simultânea de vários sinais das PCDs, situadas em diversas localizações geográficas, os sinais chegam ao transponder no satélite em instantes de tempos aleatórios e com diferentes desvios em frequência, devido ao efeito Doppler. Desse modo, o enlace espacial do sistema pode ser modelado como um canal multiusuário com desvio Doppler e ruído AWGN (Do inglês, *Additive White Gaussian Noise*).

A proposta deste trabalho é o desenvolvimento de um demodulador  $\pi/3$ -BPSK totalmente digital como parte de um transponder regenerativo multicanal do SBCD, que deve implementar a demodulação e processamento dos sinais das PCDs a bordo do satélite.

O demodulador proposto utiliza um recuperador de portadora cuja arquitetura é inédita para aplicação com sinais  $\pi/3$ -BPSK. Por causa do curto intervalo de tempo especificado para recuperação dos bits, o sincronizador de símbolos utilizado é diferente das estruturas de sincronizadores consagradas na literatura para processamento discreto, tais como Gardner e Early-late (GARDNER, 1986; MENGALI; ANDREA, 2013). Devido a esta nova abordagem, o desempenho do demodulador em termos de taxa de erros de bits é superior às versões atuais em operação (Cisotto, 2012). A arquitetura do demodulador  $\pi/3$ -BPSK proposto está apresentada em termos de diagramas de blocos e equações.

O trabalho está organizado da seguinte maneira: na Seção II, descreve-se a arquitetura do demodulador, detalhando-se cada módulo funcional do sistema. Na Seção III, apresentam-se os resultados da avaliação de desempenho através de simulações computacionais, com algumas discussões. Na Seção IV, apresentam-se as conclusões do trabalho e as perspectivas futuras.

## 2 | DESCRIÇÃO DO SISTEMA E ANÁLISES

### 2.1 Visão Geral do Transponder Regenerativo

O Sistema Brasileiro de Coleta de Dados Ambientais é composto por satélites com funções de repetidores analógicos, no qual recebem os sinais das PCDs e retransmitem para as estações de recepção em solo. Há uma proposta de utilizar um transponder regenerativo capaz de realizar o processamento digital dos sinais recebidos das PCDs e o tratamento dos dados para envio às estações receptoras. A ideia é viabilizar o processamento dos sinais das PCDs a bordo do satélite. O modelo ainda está em fase de pesquisa no INPE.

Considerando que o enlace espacial do SBCDA é modelado como um canal multiusuário com desvio Doppler e ruído AWGN, temos que o transponder regenerativo deverá ser eficiente em corrigir o desvio de frequência, realizar uma demodulação digital e por fim, verificar a integridade dos dados das PCDs recebidas para transmiti-los ao segmento do solo. Neste contexto, a arquitetura do transponder é composta por um seletor de canais, sincronizador de frequência, demodulador BPSK e processador de dados. Apresenta-se na Figura 1 o diagrama de blocos simplificado da arquitetura do transponder regenerativo.

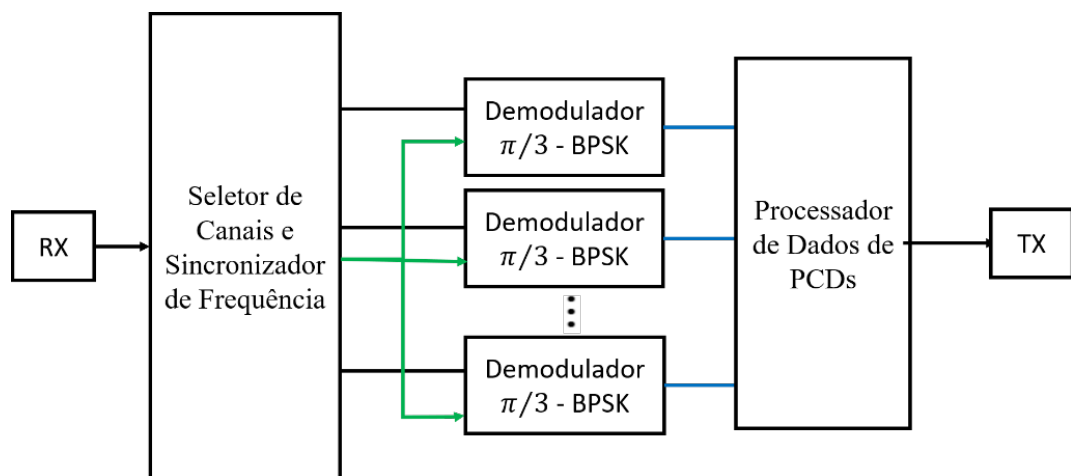


Figura 1 – Diagrama de blocos do transponder regenerativo

Fonte: Silva e Lucena, 2018.

Ao receber os sinais emitidos pelas PCDs (RX), o transponder iniciará o processo de seleção dos canais e estimativa das frequências de cada plataforma, para então serem alocados aos demoduladores. Estes processos de seleção e habilitação dos demoduladores são implementados pelo barramento de controle, em cor verde (Figura 1). Os demoduladores por sua vez, são responsáveis por completar a aquisição da portadora da PCD, sincronização dos símbolos recebidos e detecção de bits. As informações serão enviadas para o processador de dados para decodificação e checagem de erros através de um módulo de processamento de CRC (do inglês,

*Cyclic Redundancy Check*). Se os dados estiverem incorretos serão descartados, caso contrário, serão armazenados em um buffer, tipo FIFO (do inglês, *First In. First Out*), até que sejam coletadas todas as mensagens recebidas das PCDs, durante a visada corrente do satélite. Posteriormente, o processador formatará as mensagens armazenadas e enviará ao transmissor (TX) em formato de frames, detalhe do barramento de dados em cor azul, de modo a transmiti-las ao segmento solo (Francisco Tavares da Silva; Antonio Macilio Lucena, 2018).

## 2.2 Modelagem do Sinal

Os sinais transmitidos pelas PCDs consistem em uma portadora com modulação  $\pi/3$ -BPSK, taxa de bits 400 bps, frequência de 401,62 MHz e largura de banda de 60 kHz. Os dados antes da modulação são codificados no formato de *Biphase-L*, dobrando a taxa de *chip* original. O satélite recebe os sinais oriundos das PCDs com acréscimo de ruído, cuja densidade espectral de potência é igual a -173 dBm/Hz, deslocamento Doppler máximo de 9 kHz e potência na faixa de -108 a -126 dBm (PÉCALA, 2005).

O modelo discreto do sinal com modulação de cada PCD, na entrada do demodulador, pode ser expresso da seguinte maneira:

$$r[n] = \frac{\sqrt{3}A}{2} b[n] \cos(\omega_c n + \theta_c[n]) + \frac{A}{2} \sin(\omega_c n + \theta_c[n]) + w[n], \quad (1)$$

em que  $A$  é a amplitude do sinal recebido,  $\omega_c$  representa frequência discreta da portadora,  $b[n]$  é o sinal em banda base,  $\theta_c[n]$  corresponde a fase da portadora,  $w[n]$  indica o ruído AWGN com variância  $N_D$ . Por outro lado, o sinal em banda base é dado por

$$b[n] = \sum_{k=-\infty}^{\infty} a[k] P[n - kN_T - N_T], \quad (2)$$

em que  $a[k] \in \{-1,1\}$  são os símbolos BPSK transmitidos no  $k$ -ésimo período, representa o atraso de símbolo e  $P[n]$  é o pulso Manchester discreto definido por

$$\theta_c[n] = \theta_0 + \nu_D n \quad (4)$$

em que,  $\theta_0$  é uma fase desconhecida e  $\nu_D$  representa um resíduo da frequência Doppler, que corresponde a um valor máximo de  $\pm 5$  Hz. A maior parte do deslocamento Doppler foi corrigido pelo estimador de frequência do transponder regenerativo conforme descrito na seção anterior. Por outro lado, a fase desconhecida  $\theta_0$  decorre de um possível desvio de frequência e fase, provenientes do enlace espacial.

A estrutura da mensagem de uma PCD é representada na Tabela abaixo:

Portadora Pura	Portadora Modulada						
160 ms	A	B	C	D	E	F	G
Portadora Pura (Início)	160 ms						
Estrutura da Mensagem							
A – Sincronização de Bits	15 bits “1” = 37,5 ms						
B – Sincronização da Palavra	8 bits (“00010111”)						
C – Inicialização da Mensagem	1 bit “1”						
D – N° de Blocos de 32 bits ( $N_b$ )	4 bits ( $N_b = 1,2,\dots,8$ )						
E – Identificação da PCD	20 bits						
F – Dados dos Sensores	$32 \cdot N_b$						
G – CRC (opcional)	8 bits						

Tabela 1 Estrutura das mensagens das pcds (PÉCALA, 2005).

Os sinais recebidos pelo satélite possuem duração entre 360 a 920 ms, com período inicial de 160 ms de portadora pura, sem modulação, reservados para sincronização de frequência e fase. Destaca-se que os 37,5 ms seguintes da mensagem, conforme descrito na Tabela 1, são atribuídos para sincronismo de símbolo e corresponde a 15 bits “1”.

### 2.3 Arquitetura Proposta

O diagrama de blocos funcional do demodulador  $\pi/3$ -BPSK é apresentado na Figura 2.

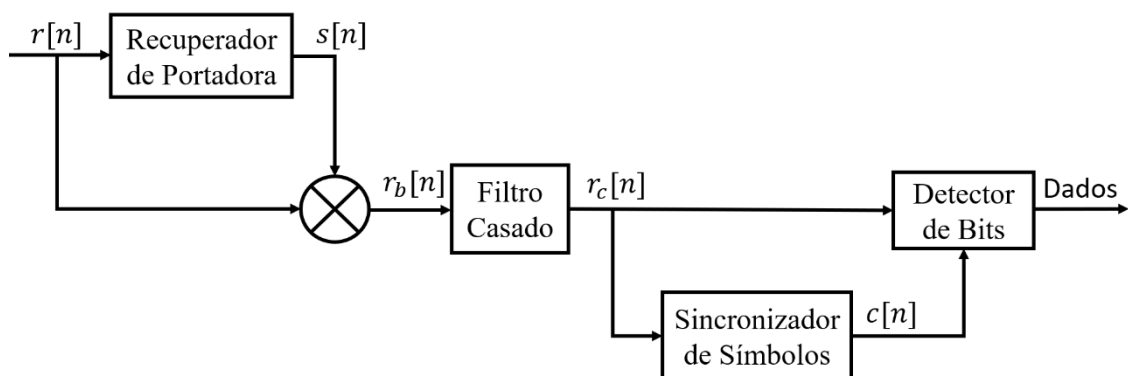


Figura 2 – Modelo do demodulador  $\pi/3$ -BPSK completamente digital.

O módulo recuperador de portadora deriva uma réplica da portadora recebida,  $s[n]$ , a partir do sinal  $r[n]$ .

O multiplicador que segue o recuperador de portadora efetua a conversão do sinal recebido, derivando o sinal  $r_b[n]$ , que contém o sinal em banda base e a portadora

modulada na frequência  $2\omega_c$ , a qual será filtrada no estágio seguinte.

Para maximizar a relação sinal-ruído, foi utilizado um filtro casado para o sinal  $r_b[n]$ . O sinal de saída é denotado por  $r_c[n]$ .

O sincronizador de símbolo recebe  $r_c[n]$  do filtro casado e gera o sinal de *clock* (neste caso uma onda quadrada)  $c[n]$ , que está sincronizado com o início e término dos símbolos.

Os sinais  $r_c[n]$  e  $c[n]$  são entregues ao módulo detector de bit que, a partir destes, faz a detecção dos bits e entrega a sequência de dados ao processador do transponder.

## 2.4 Recuperador de Portadora

O módulo recuperador de portadora, indicado na Figura 3, tem como função estimar a fase da portadora e eventualmente algum *offset* de frequência e, a partir das estimativas, gerar uma réplica da portadora recebida, representada por  $s[n]$ .

É interessante destacar que a malha de Costas clássica para sincronização de portadora com modulação BPSK (PROAKIS; SALEHI, 2008) não funciona para modulação  $\pi/3$ -BPSK. A arquitetura adotada é uma variante de uma malha de Costas discreta modificada, concebida para recuperação de portadoras com modulação QPSK (TYTGAT; STEYAERT; REYNAERT, 2012). Nossa contribuição, para o funcionamento da malha com sinais  $\pi/3$ -BPSK, foi o acréscimo dos ganhos  $g_3$  e  $g_3$  na saída dos limitadores nos braços I e Q da malha, respectivamente.

O sinal de entrada  $r[n]$  é multiplicado em ambos os *mixers*, e os sinais convertidos são filtrados nos braços I e Q da malha pelos filtros passa-baixas, para supressão da segunda harmônica. Os filtros passa-baixas são *filtros discretos com resposta infinita e frequência de corte equivalente a 800 Hz*.

Os sinais resultantes  $I[n]$  e  $Q[n]$ , nas saídas destes filtros, são utilizados para gerar o sinal de erro  $e[n]$ , que depois de filtrado pelo filtro de malha, controla a fase e a frequência das formas de onda seno e cosseno geradas pelo oscilador controlado numericamente (NCO) e utilizadas para fechar a malha por meio dos dois *mixers* de entrada. A portadora recuperada é expressa por:

$$s[n] = \cos(\omega_c n + \hat{\theta}_c[n]), \quad (5)$$

em que  $\hat{\theta}_c[n]$  representa a estimativa de fase e frequência.

### 2.4.1 Determinação da curva S da malha

Considerando desprezível a parcela de ruído  $w[n]$  na Equação (1), os sinais  $I[n]$  e  $Q[n]$  na saída dos filtros passa-baixas são dados por



$$I[n] = \frac{A}{4} \text{sen}(\phi) + \frac{\sqrt{3}A}{4} b[n] \cos(\phi), \quad (6)$$

$$Q[n] = -\frac{A}{4} \cos(\phi) + \frac{\sqrt{3}A}{4} b[n] \text{sen}(\phi), \quad (7)$$

em que  $\phi = \{\theta_c[n] - \hat{\theta}_c[n]\}$  é o erro de fase da malha. Supondo  $\phi$  menor do que  $\pi/10$ , nas saídas dos limitadores, em cada braço da malha, tem-se os seguintes sinais:

$$\text{Sgn}\{I[n]\} = b[n] \text{ e } \text{Sgn}\{Q[n]\} = -1 \quad (8)$$

em que  $\text{Sgn}\{x\}$  é igual 1 se  $x \geq 0$  e é igual a -1 se  $x < 0$ .

Portanto, conforme indicado na Figura 3, o sinal de erro pode ser expresso por

$$e[n] = g_0 g_3 \text{Sgn}\{I[n]\} Q[n] - g_0 g_4 \text{Sgn}\{Q[n]\} I[n], \quad (9)$$

que resulta em

$$e[n] = \frac{g_0 A (\sqrt{3} g_3 + g_4)}{4} \text{sen}(\phi) + \frac{g_0 A (\sqrt{3} g_4 - g_3)}{4} b[n] \cos(\phi), \quad (10)$$

substituindo (6), (7) e (8) em (9).

A curva S é definida como sendo (MENGALI; ANDREA, 2013).

$$S(\phi) = E\{e[n]|\phi\}, \quad (11)$$

Em que  $E\{e[n]|\phi\}$  representa a média estatística de  $e[n]$  dado  $\phi$ . Atribuindo-se  $g_3 = \sqrt{3}$  e  $g_4 = 1$ , obtém-se

$$S(\phi) = g_0 A \text{sen}(\phi). \quad (12)$$

Reconhecendo ainda que  $\phi \ll 1$ , na condição de rastreamento da malha, e ajustando-se o ganho  $g_0$  para normalizar a inclinação da curva S, resulta  $S(\phi) = \phi$ .

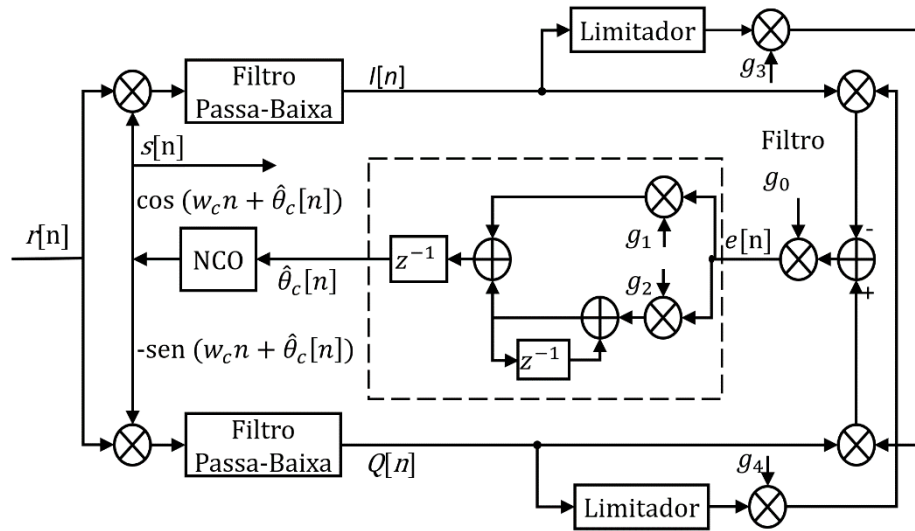


Figura 3 - Diagrama de blocos do recuperador de portadora.

### 2.4.2 Projeto da malha de Costa

A malha discreta representada na Figura 3 pode ser aproximada por um sistema linear de segunda ordem cuja função do sistema é (MENGALI; ANDREA, 2013; LINDSEY; CHIE, 1981)

$$H(z) = \frac{\hat{\theta}_c[z]}{\theta_c[z]} = \frac{(g_1 + g_2)z - g_1}{z^2 - (g_1 + g_2 - 2)z + (1 - g_1)}, \quad (13)$$

em que  $g_1$  e  $g_2$  são os ganhos do filtro de malha que definem a frequência natural  $\omega_n$ , e o fator de amortecimento  $\zeta$  da malha analógica equivalente.

Adotando como tempo de regime  $t_s = 140$  ms e  $\zeta = 0,707$ , e usando-se a seguinte aproximação (GARDNER, 2005):

$$t_s \cong \frac{4}{\omega_n \zeta}, \quad (14)$$

determina-se  $\omega_n = 40,41$ . Os ganhos  $g_1$  e  $g_2$  são determinados pelas seguintes equações (LI; MEINERS, 2000):

$$g_1 = 1 - e^{2\zeta\omega_n T_s} \quad (15)$$

$$g_2 = 1 + e^{2\zeta\omega_n T_s} - e^{2\zeta\omega_n T_s} \cos\left(\omega_n T_s \sqrt{1 - \zeta^2}\right), \quad (16)$$

que resulta em  $g_1 = 2,85 \times 10^{-4}$  e  $g_2 = 4,08 \times 10^{-8}$ .

## 2.5 Conversor para banda base

Este módulo é implementado por um simples multiplicador. O sinal de saída,  $r_b[n]$ , é o resultado do produto dos sinais  $r[n]$  e  $s[n]$ . Supondo que a portadora recuperada esteja perfeitamente sincronizada, o sinal  $r_b[n]$  pode ser descrito por:

$$r_b[n] = \frac{\sqrt{3}A}{4} b[n] + \frac{\sqrt{3}A}{4} b[n] \cos(2\omega_c n + 2\theta_c[n]) - \frac{A}{4} \sin(2\omega_c n + 2\theta_c[n]) + w[n] \cos(\omega_c n + \theta_c[n]). \quad (17)$$

Na equação acima, observa-se que o primeiro termo representa o sinal em banda base, e contém as informações transmitidas, o segundo e terceiro termos estão centrados na frequência  $2\omega_c$  e precisam ser eliminados antes da detecção de bits, e finalmente, o último termo corresponde a parcela de ruído Gaussiano.

## 2.6 Filtro casado

A resposta impulsiva do filtro casado é um pulso Manchester discreto, exatamente como está definido pela Equação 3. Ele maximiza a relação sinal-ruído e também elimina os sinais em torno de  $2\omega_c$  presentes no sinal  $r_b[n]$ . O sinal de saída pode ser representado por

$$r_c[n] = \frac{\sqrt{3}A}{4} b[n] * P[n] + z[n], \quad (18)$$

em que o asterisco representa a convolução discreta,  $P[n]$  é a resposta impulsiva do filtro casado (dada pela Eq. 3) e  $z[n]$  é o ruído Gaussiano. Neste ponto do sistema, a taxa de amostragem ainda continua sendo  $F_s = 1/T_s$ , de forma que o sinal  $r_b[n]$  contém  $N_T$  amostras por símbolo.

## 2.7 Sincronizador de símbolo

O módulo sincronizador de símbolo recebe o sinal  $r_c[n]$  para gerar um sinal de *clock*  $c[n]$ , sincronizado com os símbolos recebidos. O *clock* é repassado ao módulo detector de bits para completar o processo de detecção das informações transmitidas.

O diagrama de blocos do sincronizador de símbolo está mostrado na Figura 4. A solução adotada é uma versão modificada do sincronizador quadrático (D'ANDREA; MENGALI, 1986; PROAKIS; SALEHI, 2008) em que a não-linearidade (NL)  $x^2$  é substituída por  $|x|$ . O sinal  $r_b[m]$ , ao passar pela NL, gera uma raia espectral na frequência  $w_T$ , que corresponde a frequência analógica  $1/T$ . O sinal resultante  $c_N[n]$ , contendo a raia espectral, é filtrado por um filtro passa-faixa (FPB) estreito, centrado em  $w_T$  com largura de banda  $B_T = w_T/10$ . Como o FPB é estreito, apenas a raia espectral desejada e parte do ruído contínuo, em torno de  $w_T$  e presente em  $c_N[n]$ , passarão pelo

filtro. O sinal  $c_F[n]$  na saída do filtro é uma cossenoide com a mesma frequência e fase da raia espectral que pode ser expressa por

$$c_F[n] = V[n]\cos(w_T n + \varphi[n]), \quad (19)$$

em que  $V[n]$  e  $\varphi[n]$  são a amplitude e a fase de  $c_F[n]$ , respectivamente. A amplitude e a fase da raia recuperada não são constantes, sofrem alguma flutuação em decorrência da parcela de ruído contida em  $c_N[n]$ . A determinação da estatística destas flutuações ainda é matéria de estudos nesta pesquisa. A cossenóide  $c_F[n]$  é transformada em uma onda quadrada  $c[n]$  pela ação do módulo limitador.



Figura 4 - Diagrama de blocos do sincronizador de símbolo.

Nesta arquitetura de sincronizador, por causa do fator de sobreamostragem elevado ( $N_T \gg 1$ ), não é necessário o uso de interpoladores para determinação da melhor amostra para decisão, como é feito em (MENGALI; ANDREA, 2013).

O diagrama de blocos do módulo detector de bits está mostrado na Figura 5. O sinal de entrada  $r_c[n]$  passa por um subamostrador que escolhe uma amostra por símbolo, sob o controle da subida do *clock*  $c[n]$ , para gerar o sinal  $r_c[k]$ . Note que a subida de  $c[n]$  indica o início e o final de cada símbolo. A decisão sobre o bit recebido é feita pelo limitador. O bit estimado será considerado “1” quando  $r_c[k]$  for maior ou igual a zero, e será “0” se  $r_c[k]$  for menor do que zero.

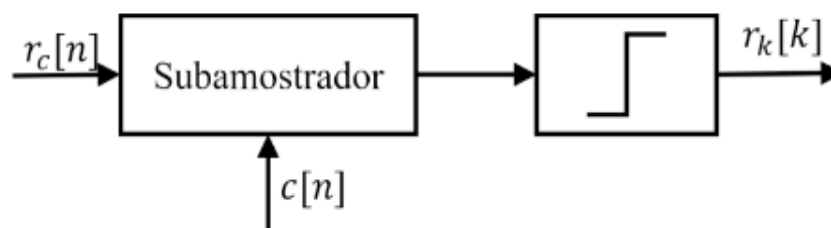


Figura 5 - Diagrama de blocos do detector de bits.

A arquitetura de demodulador proposta implementa o receptor ótimo para o canal AWGN. Entretanto, por causas das imperfeições nas estimativas dos parâmetros de sincronismo, como será apresentado nos resultados, há uma pequena perda, em termos de taxa de erro de bits, com relação ao receptor ideal.

### 3 | RESULTADOS E DISCUSSÕES

Nesta seção, apresentam-se alguns resultados de desempenho do demodulador  $\pi/3$ -BPSK proposto, obtidos via simulação computacional. Foram avaliados a taxa de erro de bits e o tempo de aquisição para o sincronismo de portadora e de símbolo.

Resumem-se abaixo os valores dos principais parâmetros utilizados no modelo:

- Frequência da portadora:  $f_c = 50$  kHz;
- Taxa de símbolo:  $1/T=400$  bps;
- Frequência de amostragem:  $F_s=200$  kHz ;
- Energia de bit por densidade de ruído:  $5 \text{ dB} \leq E_b/N_0 \leq 30 \text{ dB}$ ;
- Offset de fase: **0 a  $2\pi$** ;
- Offset de atraso: 0 a T;
- Offset da frequência da subportadora:  $\pm 5$  Hz .

#### 3.1 Taxa de erro de bit

Algumas curvas da taxa de erro de bit (BER) do demodulador estão mostradas Figura 6. A Curva 1 indica o desempenho teórico da modulação BPSK. A Curva 2 corresponde a condição de *offset* nulo para frequência, fase e atraso de símbolo. Por fim, a Curva 3 mostra a BER do demodulador com *offset* de frequência igual a 5 Hz e atraso de símbolo igual a T/2. Na condição de taxa de erros de bits igual a  $10^{-4}$  e sem desvio Doppler, o demodulador proposto tem um ganho de 1,8 dB em relação ao sistema atual em operação (Cisotto, 2012).

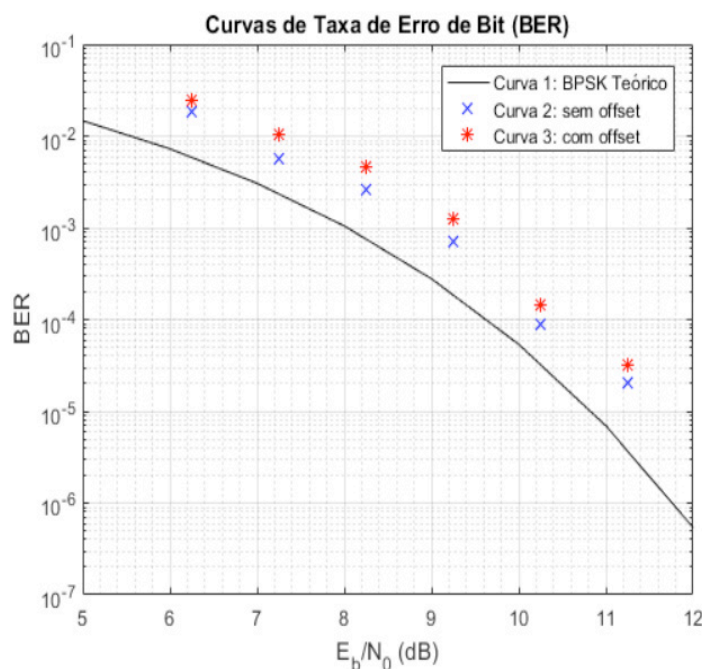


Figura 6 - Curvas da taxa de erro de bit do demodulador

### 3.2 Tempo de Aquisição

A Fig. 7 mostra a curva do transitório de fase da malha de Costas modificada para um degrau de fase de  $\pi/3$  na condição em que  $E_b/N_0 = 30$  dB. Observa-se que o tempo de regime ( $t_s$ ) é inferior a 140 ms, de modo que a especificação de projeto para aquisição de portadora é cumprida (Tab. 1).

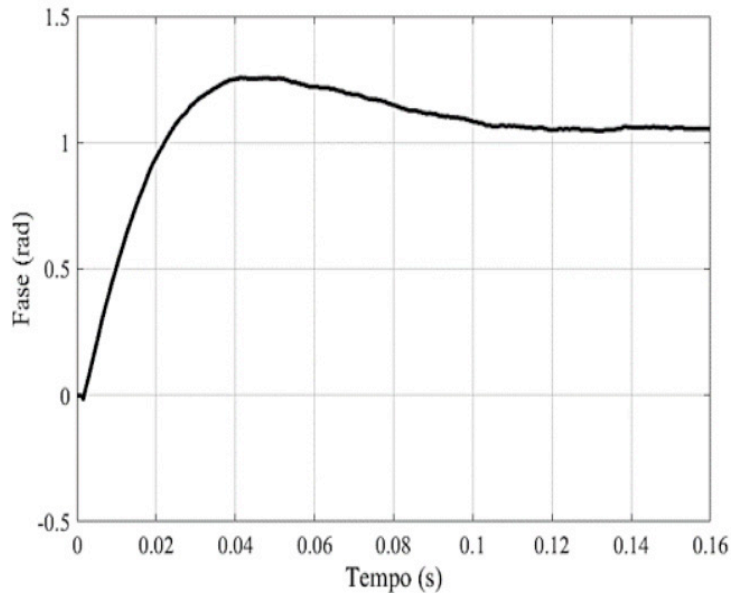


Figura 7 - Resposta transitória da malha para um degrau de fase

Medidas do transitório da saída do filtro passa-faixa do sincronizador de símbolo e a forma de onda do sinal de *clock*  $c[n]$  indicam que  $c[n]$  se estabiliza em um intervalo de tempo menor do que 4 bits. Portanto a sincronização ocorre em um tempo bem menor do que os 16 bits especificados.

## 4 | CONCLUSÃO

Neste artigo, foi apresentado o projeto de um demodulador  $\pi/3$ -BPSK coerente, com arquitetura completamente digital, destinado à aplicação nos satélites de coleta de dados do INPE, como parte do novo transponder regenerativo de bordo. Os diagramas de blocos e equações de todos os módulos funcionais do demodulador, juntamente com algumas análises matemáticas, foram apresentados para explicar a funcionalidade dos módulos.

Os resultados da simulação computacional demonstram que as soluções adotadas para a recuperação da subportadora e o sincronismo de símbolo permitem a sincronização do sistema em intervalo e tempo inferior ao que foi especificado, além de provocar pouco impacto na relação  $E_b/N_0$ .

Na pior situação operacional, quando existe offset de frequência e atraso de símbolo, o desempenho do demodulador posto em termos de BER, é apenas 1,0 dB inferior ao limite teórico, na condição de BER =  $10^{-4}$ .

Embora este trabalho demonstre que arquitetura proposta apresenta boa performance, novas estudos estão sendo desenvolvidas para determinação analítica do desempenho de cada módulo e do sistema como um todo.

## REFERÊNCIAS

- D'ANDREA, A. N.; MENGALI, U. Performance Analysis of the Delay-Line Clock Regenerator. **IEEE Transactions on Communications**, v. 34, n. 4, p. 321 – 328, Abril 1986.
- GARDNER, F. M. A BPSK/QPSK timing-error detector for sampled receivers. **IEEE Transactions on Communications**, v. 34, n. 5, p. 423 – 429, Maio 1986.
- GARDNER, F. M. **Phaselock Techniques**. 3. ed. New Jersey, USA: John Wiley & Sons, 2005. ISBN 978-0471430636.
- LI, W.; MEINERS, J. Introduction to phase-locked loop system modeling. **Analog and Mixed-Signal Products**, Texas Instruments Incorporated, Maio 2000.
- LINDSEY, W. C.; CHIE, C. M. A Survey of Digital Phase-Locked loops. **PROCEEDINGS OF THE IEEE**, v. 69, n. 4, p. 410 – 431, Abril 1981.
- MENGALI, U.; ANDREA, A. N. D. **Synchronization techniques for digital receivers**. [S.l.]: Springer Science & Business Media, 2013.
- M. Oerder. e H. Meyr, “**Digital filter and square timing recovery**”, IEEE Transactions on communications, vol. 36, nº 5, pp. 605-612, 1988.
- M. V. Cisotto, “**Proposta de simulador/emulador para avaliação para avaliação e testes do processador do Sistema Brasileiro de Coleta de Dados**”, Dissertação de Mestrado, Instituto Nacional de Pesquisas Espaciais (INPE), 2012.
- PÉCALA, J. C. **Detector de Sinais para os Satélites do Sistema Brasileiro de Coleta de Dados usando Análise Espectral Digital**. 2005. 123 p. Dissertação (Engenharia Eletrônica e Computação, Área Telecomunicação) — Instituto Tecnológico de Aeronáutica.
- PROAKIS, J. G.; SALEHI, M. **Digital Communication**. 5. ed. ed. [S.l.]: McGraw-Hill Education, 2008.
- SILVA, F. T. da; LUCENA, A. M. **Proposta de novo transponder regenerativo baseado em FPGA**. 2018.
- TIKAMI, A. et al. **O desempenho do processador de coleta de dados versão III nas estações de recepção do Sistema Brasileiro de Coleta de Dados Ambientais**. Anais XVI Simpósio Brasileiro de Sensoriamento Remoto - SBSR, Foz do Iguaçu, p. 9071 – 9077, Abril 2013.
- TYTGAT, M.; STEYAERT, M.; REYNAERT, P. **Time Domain Model for Costas Loop Based QPSK Receiver**. In: 8th Conference on Ph.D. Research in Microelectronics & Electronics. Aachen, Germany, Germany: VDE, 2012. ISBN 978-3-8007-3442-9.



## PROPOSTA DE UM NOVO ALGORITMO QoS-AWARE PARA O ESCALONAMENTO *DOWNLINK* LTE-A EM CENÁRIOS DE TRÁFEGO MISTO: UMA COMPARAÇÃO DE DESEMPENHO

**Júnio Moreira**

Instituto Federal do Triângulo Mineiro (IFTM) –  
Campus Patrocínio

Patrocínio – Minas Gerais

**Éderson Rosa da Silva**

Universidade Federal de Uberlândia (UFU)

Uberlândia – Minas Gerais

**Paulo Roberto Guardieiro**

Universidade Federal de Uberlândia (UFU)

Uberlândia – Minas Gerais

**RESUMO:** As aplicações de vídeo são responsáveis pela maior parte do tráfego nas redes móveis celulares. Em vista disso, propõe-se neste artigo um algoritmo *QoS-aware* para o escalonamento *downlink* LTE-A, denominado de “*enhanced-PSS*” (e-PSS), desenvolvido a partir do algoritmo PSS com o objetivo de melhor atender aos requisitos de QoS das aplicações de videoconferência. Esta proposta surgiu a partir de um estudo comparativo de desempenho entre diversos algoritmos de escalonamento. A avaliação de desempenho do algoritmo proposto foi realizada por meio de modelagem e simulação. Os resultados das simulações demonstraram que o escalonador proposto pode alcançar um melhor desempenho em termos de *throughput*, atraso médio, perdas de pacotes e justiça, atendendo aos requisitos das aplicações de videoconferência.

**PALAVRAS-CHAVE:** *QoS-aware*, Escalonamento, *Downlink*, LTEA, Videoconferência.

**ABSTRACT:** Video applications account for most of the traffic on mobile cellular networks. In this paper, we propose a QoS-aware algorithm for the LTE-A downlink scheduling, known as “enhanced PSS” (e-PSS), developed from the PSS algorithm with the goal of better meeting the QoS requirements of videoconference applications. This proposal arose from a comparative performance study among several scheduling algorithms. The performance evaluation of the proposed algorithm was performed through modeling and simulation. The results of the simulations demonstrated that the proposed scheduler can achieve better performance in terms of throughput, average delay, packet loss and fairness, meeting the requirements of videoconference applications.

**KEYWORDS:** QoS-aware, Scheduling, Downlink, LTE-A, Videoconference.

### 1 | INTRODUÇÃO

Atualmente, as aplicações de vídeo nas redes móveis celulares representam a maior demanda de tráfego e espera-se que as mesmas continuem crescendo nos próximos anos. De acordo com a Cisco (2017), o volume de tráfego de vídeo deverá ser 70% dos dados

móveis em 2018, aumentando a capacidade em treze vezes, entre 2014 e 2019.

Um estudo recente apresentado pela Telecoms (2015) indica que as operadoras consideram a transmissão de conteúdo de vídeo como um dos serviços do *Long-Term Evolution* (LTE) mais lucrativos. Assim, as operadoras estão preocupadas em usar os recursos de rádio de forma eficiente para alcançar um alto desempenho do sistema e garantir uma *Quality of Service* (QoS) adequada para aplicações multimídia aos usuários. No entanto, esta tarefa torna-se mais desafiadora com o crescente número de usuários do sistema (WANG et al., 2014), e com os requisitos restritos de aplicações, como por exemplo, de videoconferência.

Para enfrentar esses desafios, o *Third Generation Partnership Project* (3GPP) introduziu as especificações do *Long-Term Evolution Advanced* (LTE-Advanced) (HOLMA e TOSKALA, 2012). A rede LTEA melhora substancialmente o *throughput* do usuário final e reduz a latência, além do provisionamento de recursos e *Quality of Experience* (QoE) aprimorados para várias aplicações de multimídia em tempo real.

Para atender aos requisitos dessas aplicações e suportar as limitações dos enlaces de radiofrequência, escalonadores adequadamente desenvolvidos são requeridos. Em vista disso, os escalonadores procuram otimizar o uso de recursos de rádio, bem como maximizar o *throughput* do sistema. No padrão LTE não é definido nenhum tipo específico de escalonador, deixando assim espaço para inovação e diferenciação.

Diversos trabalhos apresentaram propostas de escalonadores *downlink* LTE. Por exemplo, em [5], os autores apresentaram a implementação de vários escalonadores *downlink*, nomeadamente, *maximum throughput* (MT), *throughput to average* (TTAS), *blind equal throughput* (BET), *token bank fair queue* (TBFQ) e *priority set* (PSS). O estudo compara os resultados de simulação produzidos pelos escalonadores com o respectivo desempenho teórico em alguns cenários de referência. No presente trabalho propõe-se a extensão do algoritmo PSS, apresentado no trabalho em Zhou et al. (2013), visando melhor desempenho de QoS. Já no trabalho descrito em Grondalen et al. (2017), comparou-se o desempenho de vários algoritmos de escalonamento *downlink* para cenários de tráfego UDP e TCP considerando os domínios de tempo e frequência.

Outros trabalhos propuseram diferentes abordagens para atender aos requisitos de QoS. Em Wang e Hsieh (2015), os autores desenvolveram um mecanismo eficiente de gerenciamento de recursos para suportar QoS e alocar os *Resource Blocks* (RB's) *downlink* para os *User Equipments* (UEs) para fluxos de tráfego *Guaranteed Bit Rate* (GBR). Os resultados mostraram que a proposta pode reduzir a perda de pacotes e o atraso dos fluxos de tráfego GBR. No estudo de Skondras et al. (2015), um escalonador *QoS-aware* para o LTE *downlink* é proposto, com o objetivo de priorizar o tráfego em tempo real.

Com base na literatura, até onde é do conhecimento dos autores deste artigo, não foram encontrados estudos comparativos dos principais escalonadores *QoS-aware* no *downlink* LTE-A com foco na aplicação de videoconferência. Dessa forma,

as principais contribuições deste artigo são:

- i. Uma análise de desempenho de diferentes algoritmos de escalonamento do estado da arte (CQAS, PFS, PSS, e-PSS e TBFQS) da transmissão *downlink* em redes móveis *Orthogonal frequency-division multiple access* (OFDMA) para tráfego misto.
- ii. Uma proposta de um escalonador QoS-aware no *downlink* LTE-A em cenário de tráfego misto, chamado enhanced-PSS (e-PSS), a fim de obter um melhor desempenho em termos de índice de justiça, atraso médio, *Packet Loss Ratio* (PLR) e *throughput* quando comparado ao PSS.

O restante deste artigo está organizado da seguinte forma. As principais informações dos algoritmos de escalonamento de pacotes para sistemas LTE que são usados na simulação são apresentadas na Seção 2. A proposta de melhoria do escalonador PSS para e-PSS é descrita na Seção 3. A Seção 4 apresenta o ambiente de simulação. Os resultados da simulação e desempenho do escalonador proposto são analisados na Seção 5. Conclusões e sugestões de trabalhos futuros são apresentadas na Seção 6.

## 2 | ALGORITMOS DE ESCALONAMENTO EM REDES LTE

Apresenta-se, a seguir, os algoritmos de escalonamento *downlink* comparados neste trabalho e a Tabela I, adaptada de Grondalen et al. (2017), resume as principais características de cada algoritmo.

Escalonadores	CQAS	PFS	PSS	TBFQS
<b>Objetivos</b>	Maximizar eficiência espectral com controle de latência	Maximizar o equilíbrio entre eficiência espectral e justiça	Maximizar <i>throughput</i> com controle de justiça	Mecanismo <i>leaky bucket</i> e qualidade do canal
<b>Características</b>	Híbrido	FD ou TD	Híbrido	Híbrido
<b>Parâmetros</b>	Capacidade atual dos RB's & <i>throughput</i> médio dos usuários & atrasos de pacotes HoL	Capacidade atual dos RB's & <i>throughput</i> médio dos usuários	Capacidade atual dos RB's & <i>throughput</i> médio dos usuários	Capacidade atual dos RB's & SINR

TABELA I: Principais Características de Diferentes Algoritmos de Escalonamento

### 2.1 Channel and QoS Aware Scheduler (CQAS)

O escalonador CQAS proposto por Bojovic e Baldo (2014) utiliza parâmetros de QoS, tais como, atraso, perda de pacotes e taxa GBR para a definição de sua métrica. Em cada *subframe*  $j$  (2 RB's), o escalonador CQAS agrupa os usuários de acordo com

a seguinte equação:

$$i_{td}(j) = \left\lfloor \frac{d_i(j)}{g} \right\rfloor \quad (1)$$

na qual  $d_i(j)$  é o atraso devido ao efeito *Head of the Line* (HoL) do usuário  $i$  no *subframe*  $j$ , e  $g$  é uma constante. Estes grupos de usuários são escalonados em ordem decrescente de acordo com o  $i_{td}(j)$ .

O agrupamento é usado para selecionar o fluxo com a maior urgência, ou seja, com o maior valor de atraso HoL e para garantir que o mecanismo de escalonamento considere os fluxos no domínio da frequência. Um valor baixo da variável  $g$  reduz a diversidade dos usuários e os ganhos de escalonamento, portanto, privilegia o  $d_i(j)$  dos usuários. Por outro lado, um valor alto da variável  $g$  aumenta a diversidade dos usuários e os ganhos de escalonamento, porém o  $d_i(j)$  possui um impacto menor nas decisões de escalonamento (BOJOVIC e Baldo, 2014). Os UEs em cada grupo são atribuídos aos *Resource Block Group* (RBGs) nesse *subframe*. Assim, para cada  $i$ -ésimo RBG, os UEs são ordenados conforme uma segunda métrica, definida como:

$$i_{fd,CQA}(j, l) = d_i(j) \cdot \frac{GBR_i}{\zeta_j(j)} i_{ca}(j, l) \quad (2)$$

em que  $GBR_i$  é a taxa de bits que a rede deve suportar para o usuário  $i$ , a variável  $\zeta_j(j)$  representa o *throughput* médio, enquanto que  $i_{ca}(j, l)$  leva em conta o RBG alocado para o  $i$ -ésimo usuário, sendo calculada como:

$$i_{ca,PF}(j, l) = \frac{r_i(j, l)}{\zeta_j(j)} \quad (3)$$

na qual  $r_i(j, l)$  é a taxa de dados em que a EnodeB pode transmitir para o usuário  $i$  no  $i$ -ésimo RBG do *subframe*  $j$ .

## 2.2 Proportional Fair Scheduler (PFS)

O PFS tenta encontrar um equilíbrio entre eficiência espectral e justiça, ajustando o índice de escalonamento de acordo com a qualidade instantânea de canal (CAPOZZI et al., 2013), em relação ao seu desempenho ao longo do tempo. A métrica de escalonamento do PFS é, portanto, dada por:

$$i_{pfs}(j) = \arg \max_{1 \leq i \leq N} \left( \frac{r_i(j)}{\zeta_i(j)} \right). \quad (4)$$

O objetivo é utilizar o *throughput* médio  $\zeta_i(j)$  para atuar como um fator de ponderação da taxa de dados  $r_i(j)$ , para que os usuários em condição de qualidade baixa de canal sejam atendidos dentro de um determinado período de tempo.

### 2.3 Priority Set Scheduler (PSS)

O PSS é um escalonador *QoS-aware* que combina operações de escalonamento no domínio do tempo (TD) e domínio da frequência (FD). De acordo com Monghal et al. (2008), no primeiro estágio, o PSS extrai  $N_{max}$  usuários para serem servidos no segundo estágio, de acordo com a abordagem FD. Para cada *subframe*  $j$ , o escalonador divide os usuários em dois grupos, baseado em seu *throughput* médio  $\zeta_i(j)$ , se o valor for menor (grupo 1), ou se o valor for maior (grupo 2) do que o *target bit rate* (TBR). Os UEs do grupo 1, tem prioridade sobre o grupo 2. Os usuários  $N_{max}$  com maior prioridade são selecionados para o segundo estágio. Assim, os  $i$ -ésimos RBGs no *subframe*  $j$  são atribuídos para o UE  $i$  com a máxima métrica do escalonador *Proportional Fair* (PFS), ou seja:

$$i_{pss}(j, l) = \arg \max_{1 \leq i \leq N} \left( \frac{r_i(j, l)}{\zeta_{sch,i}(j)} \right). \quad (5)$$

na qual  $\zeta_{sch,i}(j)$  considera o *throughput* médio do passado para o EU  $i$ , ao contrário de  $r_i(j, l)$ , ao qual é atualizado apenas quando o UE é realmente escalonado.

### 2.4 Token Bank Fair Queue (TBFQS)

OTBFQS é um escalonador *QoS-aware* que combina operações de escalonamento no domínio do tempo (TD) e domínio da frequência (FD). O escalonador sempre seleciona o UE com a métrica mais alta e aloca o RBG com o maior *Channel Quality Indicator* (CQI) de sub-banda até que não haja nenhum pacote dentro do *buffer Radio Link Control* (RLC) do UE ou todos os RBGs sejam alocados (BOKHARI et al., 2008). No TD-TBFQ, depois de selecionar o UE com a métrica máxima, ele aloca todos os RBGs para este UE usando o CQI. O TBFQS é baseado no mecanismo de *leaky bucket* que controla os fluxos para um determinado perfil de tráfego. Cada fluxo  $k$  é caracterizado pelos seguintes parâmetros:

- $\Lambda_k$ : Taxa de chegada de pacotes
- $r_k$ : Taxa de geração de *token*
- $p_k$ : Tamanho do *pool token*
- $e_k$ : Contador do número de *tokens* pelo fluxo  $k$

A alocação é baseada na SINR máxima, que é utilizada para selecionar o usuário:

$$i_{TBFQS}(j) = \arg \max_{1 \leq i \leq N} (Y_{ij}(t_j)). \quad (6)$$

na qual  $Y_{ij}(t_j)$  é a SINR do usuário  $i$  selecionado no *subframe*  $j$ . Este é um algoritmo oportunista, o que significa que a política *Adaptive Modulation and Coding* (AMC) explora ao máximo a diversidade multiusuário em uma célula com intervalo de tempo em que apenas um usuário é escalonado.

### 3 | PROPOSTA DE MELHORIA DO ALGORITMO QOS-AWARE (PSS): ENHANCED PRIORITY SET SCHEDULER (E-PSS)

Neste trabalho, o objetivo é priorizar o tráfego de videoconferência, que é uma aplicação multimídia em tempo real com requisitos de QoS.

Como algoritmo de escalonamento base para a proposta apresentada no presente trabalho, escolheu-se o PSS devido a três de suas características, que são: simplicidade, permitir operações de escalonamento nos domínios do tempo (TD) e da frequência (FD), e ser um algoritmo *QoS-aware*. Para isso, propõe-se modificar o PSS de forma a atender aos requisitos de QoS da aplicação de videoconferência em cenários de tráfego misto. Como a referida aplicação é sensível ao atraso e requer um *throughput* mínimo, acredita-se ser importante priorizar usuários que estão há mais tempo esperando por uma transmissão. Nesse sentido, a opção escolhida para melhorar o desempenho do PSS foi alterar a sua métrica derivada da Equação 5, visando manter suas características básicas. Este novo algoritmo de escalonamento foi denominado de «*enhanced-PSS*» (e-PSS), cuja métrica é apresentada a seguir:

$$i_{e-PSS}(j, l) = \arg \max_{1 \leq i \leq N} d_i(j) \cdot \left( \frac{r_i(j, l)}{\zeta_{sch,i}(j)} \right). \quad (7)$$

Neste caso,  $i_{e-PSS}(j, l)$  atua diminuindo o atraso na rede de acesso LTE, levando em conta o atraso HoL sofrido pelos pacotes.

### 4 | AMBIENTE DE SIMULAÇÃO

Para avaliar o desempenho do e-PSS, um ambiente de simulação no NS-3.28 (NS3, 2017) foi modelado com *software* do projeto LENA (LENA, 2018). A Tabela II apresenta os principais parâmetros de simulação, os quais são baseados no trabalho descrito em Mata e Guardieiro (2017).

Modelo de antena da EnodeB	Modelo de antena isotrópico
Potência EnodeB TX	46 dBm
Potência UE TX	23 dBm



Largura de banda	25 RB's (5MHz)
Esquema AMC	PiroEW2010 (PIRO et al., 2010)
Modo RLC	UM (tamanho do buffer: 10 MB)
Modelo perda de percurso	COST 231
Modelo perda de desvanecimento	Pedestrian EPA model 3 km/h
Modelo de mobilidade do usuário	Steady state random waypoint
	Velocidade Min.: 0,8 m/s
	Velocidade Max.: 0,83 m/s
	Pausa Min.: 0s / Pausa Max.: 0,1s
	Retângulo: 500m x 500m
Tempo simulado	30 segundos
Sequência de vídeo	akiyo (300 frames)
Resolução do vídeo	QCIF (176 x 144)
Informação de vídeo	Taxa de Bits: 128 kbps
	Taxa do Frame: 25 fps
	Grupo de imagens: 30
	MTU: 1460 Bytes
	Encoder: ffmpeg
Distribuição de tráfego	[Background:VoIP:Video]
	[0,20:0,40:0,40]
Rodadas de simulação	30
Intervalo de confiança	95%

TABELA II: Parâmetros de Simulação

O modelo utilizado considera o tráfego de videoconferência compartilhando os recursos de rede com um tráfego de *background*, do tipo VoIP e transferência de arquivos. Considera-se um cenário de célula única, com um conjunto de UEs conectados a uma EnodeB. Esses UEs foram distribuídos aleatoriamente em uma área quadrada de 500 m<sup>2</sup> ao redor da EnodeB. A modelagem do enlace de radiofrequência segue o COST 231 [9] para perdas de percurso e Pedestrian EPA model 3 km/h para perdas de desvanecimento. Os usuários movem-se dentro da área de cobertura de acordo com o modelo *Steady State Random Waypoint* (NAVIDI e CAMP, 2004).

Para a modelagem do tráfego de vídeo gerado, utilizou-se a sequência de vídeo *akiyo* através do *framework* Evalvid (EVALVID, 2018). Essa sequência apresenta características típicas das videoconferências e parâmetros de simulação próximos aos usados em um sistema real. A sequência de vídeo *akiyo* tem duração de 12 segundos e o tempo simulado é ajustado para 30 segundos. Os resultados de simulação são baseados em médias obtidas a partir de 30 rodadas e intervalo de confiança de 95% foram incluídos.

## 5 | APRESENTAÇÃO E ANÁLISE DE RESULTADOS

Para O desempenho do algoritmo de escalonamento proposto (e-PSS) foi



comparado com o desempenho dos seguintes algoritmos: CQAS, PFS, PSS e TBFQS. As métricas de desempenho utilizadas para a análise foram: índice de justiça, atraso médio, PLR e *throughput*.

## 5.1 Índice de Justiça

A Figura 1 apresenta o índice de justiça em relação ao número de usuários na rede.

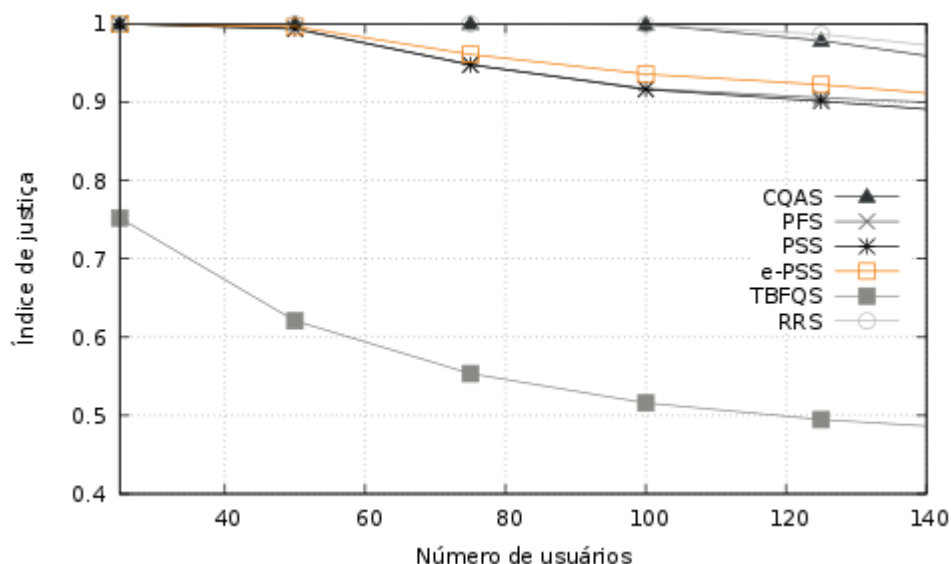


Figura 1: Índice de justiça para usuários no cenário de tráfego misto.

Com exceção do TBFQS, que possui o menor índice de justiça em relação aos demais algoritmos, o restante dos algoritmos possuem uma leve diferenciação no nível de justiça. Para fins de comparação, apresenta-se também a curva de desempenho do algoritmo *Round-robin Scheduling* (RRS) (MOHAMMAD et al., 2012). Com relação ao algoritmo proposto, e-PSS, nota-se um bom desempenho, o qual se justifica por usar técnicas baseadas no algoritmo PFS.

## 5.2 Taxa de Perda de Pacotes

Uma característica importante para avaliar a qualidade da transmissão de vídeo é a PLR. Os buffers de RLC nos UEs foram configurados para 10 MB para assegurar que as perdas sejam causadas pelo processo de descarte e não afetadas por sobrecargas de buffer. A Figura 2 mostra a PLR média em função do número de usuários na rede. Ao dar maiores prioridades para o tráfego de videoconferência, o algoritmo CQAS produz os melhores resultados. Com relação ao algoritmo proposto, e-PSS, observa-se que o desempenho apresentado é superior ao algoritmo PSS.

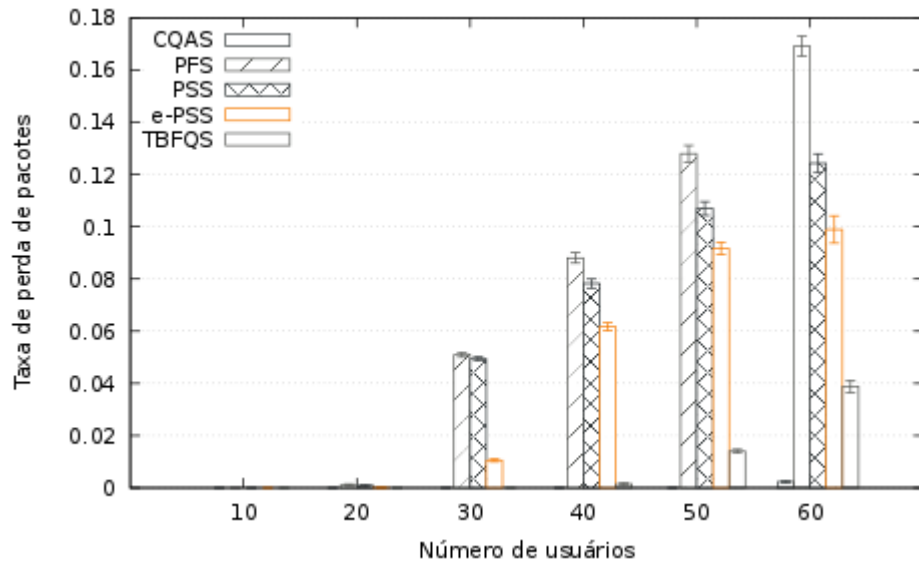


Figura 2: Taxa de perda de pacotes para usuários da aplicação de videoconferência.

### 5.3 Atraso Médio

A Figura 3 apresenta o atraso médio em função do número de usuários na rede para a aplicação de videoconferência. Ao dar maiores prioridades para o tráfego de videoconferência, o algoritmo CQAS produz os melhores resultados. Além disso, o algoritmo e-PSS apresenta um melhor desempenho em comparação ao PSS, o que comprova a eficácia da inserção da métrica HoL.

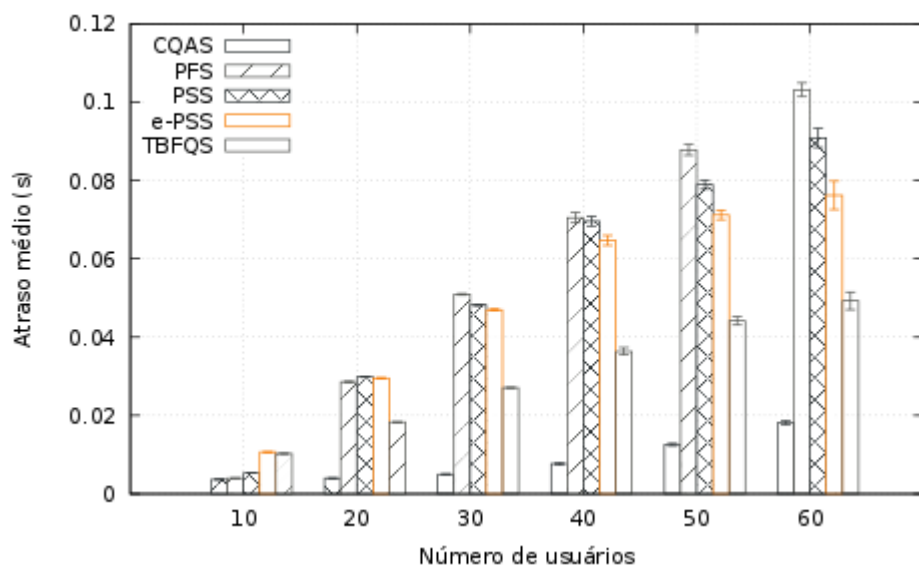


Figura 3: Atraso médio para a aplicação de videoconferência.

### 5.4 Throughput

A Figura 4 mostra a Cumulative Distribution Function (CDF) para o throughput por usuário para a aplicação de videoconferência, considerando 150 usuários.

O tráfego de vídeo pertence à classe GBR e, portanto, requer uma taxa de bits

mínima. A fonte de tráfego de videoconferência usada na simulação opera a 128 kbps. Portanto, o gráfico da Figura 4 mostra que somente os algoritmos TBFQS, CQAS e e-PSS atendem aos requisitos de QoS da aplicação. Além disso, conclui-se que o algoritmo proposto e-PSS apresenta desempenho superior ao algoritmo PSS.

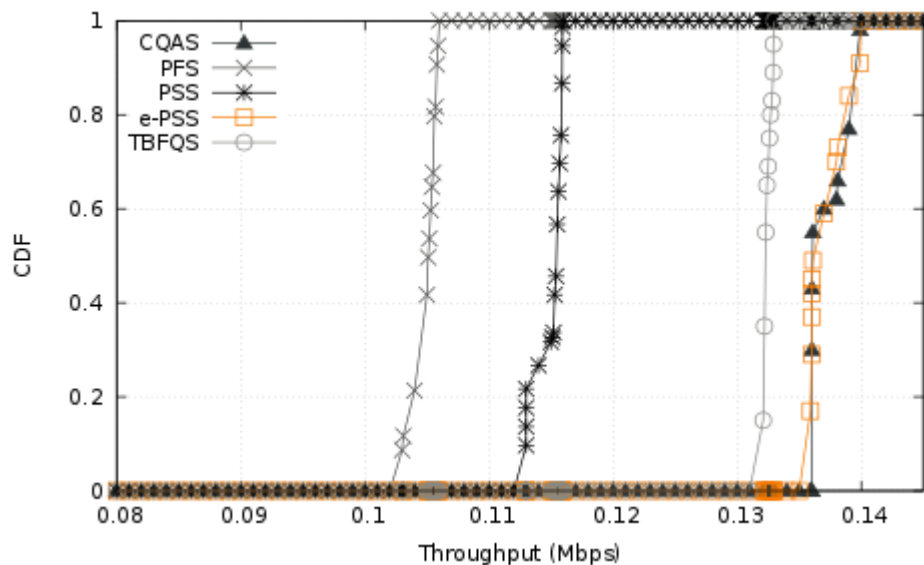


Figura 4: CDF para o Throughput por usuário para a aplicação de videoconferência, considerando 150 usuários.

## 6 | CONCLUSÕES

Neste artigo propõe-se uma otimização do algoritmo *QoS-aware* PSS, para o *downlink* LTE-A, visando priorizar o tráfego da aplicação de videoconferência. Para isso, considerou-se o valor do atraso HoL na definição da métrica do algoritmo proposto e-PSS.

Os resultados de simulação mostraram que o algoritmo e-PSS apresenta um desempenho superior ao algoritmo PSS. Além disso, o e-PSS atende satisfatoriamente aos requisitos de QoS da aplicação de videoconferência. Para a continuidade deste trabalho destaca-se a necessidade de um estudo aprofundado do algoritmo de escalonamento proposto em um cenário de redes heterogêneas.

## 7 | AGRADECIMENTOS

Os autores agradecem à Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) pelo apoio financeiro fornecido através do processo TEC-APQ-03279-17.

## REFERÊNCIAS

A. R. H. A. M. J. S. Mohammad T. Kawser, Hasib M. A. B. Farid and I. K. Razu, “**Performance comparison between round robin and proportional fair scheduling methods for lte,**” in International Journal of Information and Electronics Engineering, vol. 2, no. 5, September 2012, pp.

B. Bojovic and N. Baldo, “**A new channel and qos aware scheduler to enhance the capacity of voice over lte systems**,” in 2014 IEEE 11th International Multi-Conference on Systems, Signals Devices (SSD14), Feb 2014, pp. 1–6.

“**Cisco visual networking index: Global mobile data traffic forecast update, 2016-2021**,” Cisco, San Jose, CA, USA, Rep., 2017. [Online]. Available: <https://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/mobile-white-paper-c11-520862.html?tid=ossdc000283>

D. Zhou, N. Baldo, and M. Miozzo, “**Implementation and validation of lte downlink schedulers for ns-3**,” in Proceedings of the 6th International ICST Conference on Simulation Tools and Techniques, ser. SimuTools '13. ICST, Brussels, Belgium, Belgium: ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2013, pp. 211–218. [Online]. Available: <http://dl.acm.org/citation.cfm?id=2512734.2512763>

(2018) **EvalVid Framework**. [Online]. Available: <http://www.tkn.tu-berlin.de/menue/research/evalvid>

E. Skondras, A. Michalas, A. Sgora, and D. D. Vergados, “**A downlink scheduler supporting real time services in lte cellular networks**,” in 2015 6th International Conference on Information, Intelligence, Systems and Applications (IISA), July 2015, pp. 1–6.

F. A. Bokhari, W. K. Wong, and H. Yanikomeroglu, “**Adaptive token bank fair queuing scheduling in the downlink of 4g wireless multicarrier networks**,” in VTC Spring 2008 - IEEE Vehicular Technology Conference, May 2008, pp. 1995–2000.

F. Capozzi, G. Piro, L. A. Grieco, G. Boggia, and P. Camarda, “**Downlink packet scheduling in lte cellular networks: Key design issues and a survey**,” IEEE Communications Surveys Tutorials, vol. 15, no. 2, pp. 678–700, Second 2013.

G. Monghal, K. I. Pedersen, I. Z. Kovacs, and P. E. Mogensen, “**Qos oriented time and frequency domain packet schedulers for the utran long term evolution**,” in VTC Spring 2008 - IEEE Vehicular Technology Conference, May 2008, pp. 2532–2536.

G. Piro, L. A. Grieco, G. Boggia, and P. Camarda, “**A two-level scheduling algorithm for qos support in the downlink of lte cellular networks**,” in 2010 European Wireless Conference (EW), April 2010, pp. 246–253.

H. Holma and A. Toskala, **LTE Advanced: 3GPP Solution for IMTAdvanced**. New York, NY, USA: Wiley, 2012.

(2018) **LENA Module**. [Online]. Available: <http://lena.cttc.es/manual/lte.html>

(2017) **Network Simulator 3 Homepage**. [Online]. Available: <http://www.nsnam.org>

O. Grondalen, A. Zanella, K. Mahmood, M. Carpin, J. Rasool, and O. N. Osterbo, “**Scheduling policies in time and frequency domains for lte downlink channel: A performance comparison**,” IEEE Transactions on Vehicular Technology, vol. 66, no. 4, pp. 3345–3360, April 2017.

S. H. da Mata and P. R. Guardieiro, “**Resource allocation for the lte uplink based on genetic algorithms in mixed traffic environments**,” Computer Communications, vol. 107, pp. 125 – 137, 2017.

“**Telecoms.com intelligence annual industry survey 2015**,” Telecoms.com, 2015. [Online]. Available: <http://telecoms.com/intelligence/telecoms-com-annual-industry-survey-2015/>

W. Navidi and T. Camp, “**Stationary Distributions for the Random Waypoint Mobility Model**,” IEEE Transactions on Mobile Computing, vol. 3, no. 1, pp. 99–108, Jan. 2004.

Y. C. Wang and S. Y. Hsieh, “**Qos-provisioning downlink resource management in 4g cellular systems**,” in 2015 International Wireless Communications and Mobile Computing Conference (IWCMC), Aug 2015, pp. 67–72.

Y. Wang, J. Xu, and L. Jiang, “**Challenges of system-level simulations and performance evaluation for 5g wireless networks**,” IEEE Access, vol. 2, pp. 1553–1561, 2014.

## SERVIÇO DE L2VPN EM REDES DE *BACKBONE* IP: ESTUDO DE CASO DA REDECOMEP-RIO

### **Pedro Henrique Diniz da Silva**

Centro Brasileiro de Pesquisas Físicas,  
Coordenação de Desenvolvimento Tecnológico  
RedeRio/FAPERJ, Centro de Engenharia e  
Operações  
Rede Nacional de Ensino e Pesquisa, Ponto de  
Presença do Rio de Janeiro  
Rio de Janeiro - RJ

### **Natália Castro Fernandes**

Universidade Federal Fluminense, Departamento  
de Engenharia de Telecomunicações  
Niterói - RJ

### **Nilton Alves Jr.**

Centro Brasileiro de Pesquisas Físicas,  
Coordenação de Desenvolvimento Tecnológico  
RedeRio/FAPERJ, Centro de Engenharia e  
Operações  
Rio de Janeiro - RJ

### **Márcio Portes de Albuquerque**

Centro Brasileiro de Pesquisas Físicas,  
Coordenação de Desenvolvimento Tecnológico  
RedeRio/FAPERJ, Centro de Engenharia e  
Operações  
Rio de Janeiro - RJ

**RESUMO:** As redes OpenFlow introduzem o conceito de controladores centralizados para gerenciamento do comportamento do encaminhamento dos elementos de rede. Esse conceito permite que diversos serviços em

redes de *backbone*, atualmente implementados através de soluções proprietárias, distribuídas pelos diversos elementos e de grande complexidade de operação, sejam simplificados facilitando a operação de rede. Esse trabalho propõe uma nova aplicação para provimento de conexões de Redes Privadas Virtuais de Camada 2 (L2VPN, do inglês *Layer 2 Virtual Private Network*) em redes de *backbone*, para caso de uso do *backbone* acadêmico, de pesquisa e de governo da cidade do Rio de Janeiro (REDECOMEP-Rio). São apresentados os algoritmos implementados através do protocolo OpenFlow e controlador Ryu. A solução proposta é avaliada através de cenários emulados por meio do emulador Mininet e de um estudo de caso executado na rede de produção da REDECOMEP-Rio.

**PALAVRAS-CHAVE:** L2VPN; Redes Definidas por Software; OpenFlow; Controlador Ryu.

**ABSTRACT:** OpenFlow networks introduce the concept of centralized controllers for managing the forwarding behavior of network elements. This concept allows multiple services in backbone networks, currently implemented through proprietary solutions, distributed by the various elements, and of highly complex operation, to be simplified to facilitate network operation. This paper proposes a new application for provision of Layer 2 Virtual Private Networks (L2VPN) in

backbone networks for use case of academic, research, and government backbone network of the city of Rio de Janeiro (REDECOMEP-Rio). We present the algorithms implemented using the OpenFlow protocol and Ryu controller. We also evaluate the proposed solution through emulated scenarios using Mininet emulator and a case study implemented in the production network of Redecomep-Rio.

**KEYWORDS:** L2VPN; Software Defined Networks; OpenFlow; Ryu Controller.

## 1 | INTRODUÇÃO

A RedeRio Metropolitana (REDECOMEP-Rio) (Moraes *et al.*, 2015) é uma infraestrutura de fibras óticas próprias que formam uma rede de alta velocidade para as instituições de ensino, ciência, tecnologia, inovação e governo na cidade do Rio de Janeiro. Apesar de ser um *backbone* que atende instituições acadêmicas, a REDECOMEP-Rio, que opera como uma rede de núcleo IP tradicional puramente roteada, também atende a diversos outros tipos de instituições, as quais incluem empresas ligadas à prefeitura, ao estado e ao governo federal. Além de serviços de conectividade de alta velocidade à Internet e de alta confiabilidade, essas instituições também carecem de serviços que simulem redes privadas em cima dessa rede de núcleo compartilhada entre os diversos afiliados. Esse tipo de rede privada é também conhecido como Rede Privada Virtual (VPN, do inglês *Virtual Private Network*).

Vários tipos de VPNs estão disponíveis para a comunidade atualmente e assim sendo existem alguns tipos de classificação para tais. Um tipo de classificação é a divisão em dois modelos de serviços (Knight e Lewis, 2004): VPNs de camada 2 do modelo de referência TCP/IP (L2VPNs do inglês *Layer 2 Virtual Private Networks*); e VPNs de camada 3 (L3VPNs do inglês *Layer 3 Virtual Private Networks*). As VPNs mais comuns são as de camada 3 (Knight e Lewis, 2004) capazes de fazer com que dois elementos de rede pareçam estar diretamente conectados via uma rede roteada, ainda que existam diversos elementos de rede (roteadores) no meio do caminho. Já as VPNs de camada 2 são conexões ponto-a-ponto que simulam um circuito físico de camada 2. Sua principal vantagem é a transparência do enlace virtual formado ponta-a-ponta, permitindo que qualquer protocolo, e não somente o IP, possa trafegar.

Apesar de cada uma das tecnologias apresentadas oferecer diversas vantagens e desvantagens uma em relação a outra, lidar com a variedade de protocolos para a criação de VPNs torna a gerência e operação da infraestrutura de rede mais complexa para os seus operadores. Essas tecnologias são difíceis de implementar, ou operacionalmente quase impossíveis de realizar em larga escala em redes IP puramente roteadas (Osborne e Simha, 2002), como o caso da REDECOMEP-Rio, devido principalmente a complexidade de configuração dos equipamentos envolvidos.

Com o enfoque em simplificar e otimizar o processo de operação de uma rede de núcleo puramente roteada, propõe-se uma solução de estabelecimento de circuitos L2VPN, denominada de *Virtual Circuits Flow* (VCFlow), que evita a complexidade de



ter protocolos de encapsulamento adicionais para o estabelecimento das conexões ponto-a-ponto, sem necessidade de configurações adicionais nos equipamentos envolvidos e centraliza a configuração das conexões em um ponto único da rede. Essa solução escala naturalmente ao passo que a mesma depende, basicamente, do processo de descoberta de topologia do protocolo OpenFlow.

A aplicação desenvolvida para o estabelecimento de L2VPNs instala regras de fluxos OpenFlow reativamente a cada circuito que passa pelo *backbone*, baseado em um arquivo de configuração centralizado. Essa aplicação visa ser implementada na operação da REDECOMEP-Rio como um serviço na rede de produção.

Na próxima seção, são apresentados os trabalhos relacionados. Na seção subsequente é apresentada a arquitetura proposta de estabelecimento de L2VPNs. Apresenta-se, também, na Seção 4, uma avaliação do protótipo e um estudo de caso na REDECOMEP-Rio. O artigo é concluído na Seção 5.

## 2 | TRABALHOS RELACIONADOS

Nos últimos anos, a comunidade de Redes de Educação e Pesquisa (REN, do inglês *Research and Education Networks*) tem investido no desenvolvimento de diversas arquiteturas, protocolos e *softwares* controladores para suportar os serviços de provisionamento de circuitos virtuais (Rao *et al.*, 2005; Zheng *et al.*, 2005; Bobyshev *et al.*, 2006; Guok *et al.*, 2006; Yang *et al.*, 2006; Katramatos *et al.*, 2007), com enfoque principalmente voltado a projetos de *grids* computacionais para física de altas energias. O projeto *Lambda Station* (Bobyshev *et al.*, 2006), idealizado pelo *Fermi National Accelerator Laboratory* (FermiLab) e o *California Institute of Technology* (Caltech), foi uma das primeiras iniciativas para provisionamento dinâmico de circuitos para aplicações científicas e utilizava técnicas de *Policy Based Routing* (PBR) para encaminhamento de tráfego entre *clusters* de computadores. Já o projeto *Terapaths* (Katramatos *et al.*, 2007), financiado pelo Departamento de Energia dos Estados Unidos, propôs a criação de caminhos virtuais fim a fim com garantias de banda, combinando o uso de redes locais (LAN) baseadas em técnicas de *DiffServ* e redes de longa distância (WAN) baseadas em túneis *Layer 2 Virtual Private Network* (L2VPN) *Pseudowire* em redes *Multi Protocol Label Switching* (MPLS). Essa solução se baseava na utilização de técnicas de estabelecimento de circuitos conhecidas como L2VPN, que permitem a criação de túneis virtuais sobrepostos a uma rede IP por exemplo, onde para o seu funcionamento é necessário a utilização de protocolos de encapsulamento. Por consequência, esse tipo de protocolos opera através de cabeçalhos adicionais incorporados aos pacotes tradicionais, o que faz com que mais recursos de rede sejam consumidos para o estabelecimento dos circuitos virtuais fim a fim. Soluções como UltraScience (Rao *et al.*, 2005), também financiado pelo Departamento de Energia dos Estados Unidos, e CHEETAH (Zheng *et al.*, 2005), proposto por pesquisadores da Universidade da Virginia e do Laboratório Nacional de Oak Ridge, se baseavam

no estabelecimento de canais óticos dedicados, também conhecidos como L1VPNs óticos, em redes *Synchronous Optical Networking* (SONET) através do protocolo *Generalized Multi Protocol Label Switching* (GMPLS). Entretanto, esse tipo de implementação carece também de cabeçalhos adicionais para encapsulamento, além de um conjunto de *softwares* e protocolos de sinalização e controle (*e.g.* OSPF-TE e RSVP-TE) que possuem grande complexidade para serem implementados para permitir o provisionamento de canais dedicados de forma dinâmica.

No caso de redes IP convencionais, que não implementam nenhuma técnica para provisionamento de circuitos dinâmicos ou VPNs, seja uma rede de *campus* ou até mesmo um provedor de serviços que não implementa MPLS, muitas das vezes um *software* de controle tem de ser empregado para exercer funções de monitoramento e configuração dos *switches*, roteadores e enlaces, para prover algum tipo de circuito estático. Tipicamente, esse tipo de *software* é implementado utilizando tecnologias proprietárias, permitindo acesso somente aos operadores de rede do provedor, além de carecer muitas das vezes de configurações manuais em vários *switches* e roteadores no caminho (dispositivos esses, em geral, produzidos pelo mesmo fabricante). Além disso, cada tipo de técnica para provisionamento de circuitos virtuais de maneira estática demanda um conjunto de procedimentos específicos para o fornecimento do mesmo serviço, como o caso de protocolos *Layer 2 Tunneling Protocol* (L2TP) e túneis *Generic Routing Encapsulation* (GRE). Isso acaba por aumentar e muito a sua complexidade de implantação.

Atualmente, os serviços de provisionamento de circuitos existentes e em operação, como OSCARS (Guok *et al.*, 2006; Esnet, 2016), adotado pelo *backbone* da *Energy Sciences Network* (ESnet) do Departamento de Energia dos Estados Unidos, e AutoBAHN (Geant, 2017), adotado pela rede de educação e pesquisas da Europa denominada GEANT, se baseiam na utilização de policiamento de tráfego para prover Qualidade de Serviço (QoS) em redes que operam por meio de diferentes tipos de tecnologia, como *Ethernet*, SONET e MPLS por exemplo. Ambas as soluções se baseiam, basicamente, na utilização de protocolos abertos como *Inter Domain Controller Protocol* (IDCP) e *Network Services Interface* (NSI), mas cada um desses serviços disponibiliza um conjunto de *Application Programming Interfaces* (API) desenvolvidas para cada cenário com suas especificidades, o que dificulta sua utilização em outras redes.

Entretanto, novas soluções para provisionamento de circuitos virtuais fim a fim têm surgido mais recentemente a partir da introdução do novo paradigma de Redes Definidas por *Software* (SDN, do inglês *Software Defined Networks*). Esse paradigma reduz a complexidade para implantação de serviços de provisionamento de circuitos virtuais dinâmicos já que permite o desacoplamento do *software* de plano de controle do dispositivo de encaminhamento. Isso faz com que esse *software* de controle, que era integrado fortemente nos equipamentos individuais de rede, possa ser implementado em controladores SDN externos (logicamente centralizados) em vez de nos próprios

*switches*, permitindo a abstração da infraestrutura de rede para as aplicações e serviços de rede (Chaves Filho, 2015). Esse mecanismo de funcionamento permite que os controladores SDN tenham uma visão centralizada da rede e regras de encaminhamento sejam aplicadas aos *switches*, baseando-se em uma classificação de tráfego em fluxos de rede a partir somente das informações dos protocolos tradicionalmente utilizados em LAN e WAN, como *Ethernet*, IP e TCP. No momento atual, o OpenFlow é o padrão de interfaces de SDN para conexão entre controladores e dispositivos de encaminhamento mais amplamente aceito e implementado por fabricantes de dispositivos de rede, que disponibilizam recursos de SDN. Esse padrão prove, basicamente, a especificação para o canal de comunicação entre *switches* e controladores, e é baseado nele que a maioria das soluções/aplicações de rede disponibilizadas pela comunidade se baseiam (Kreutz *et al.*, 2015).

Essa abordagem de serviços de provisionamento de circuitos baseada em SDN/OpenFlow traz como uma de suas principais vantagens a não carência de protocolos de encapsulamento adicionais para estabelecimento de circuitos, visto que podem fazer uso de técnicas de modificação dos cabeçalhos já utilizados por padrão em redes locais, como é o caso do cabeçalho de VLAN. Outra vantagem também muito importante é a possibilidade de que quando esse serviço de provisionamento for utilizado, por exemplo para transferências de arquivos, as perdas de pacotes se mantenham baixas, visto que os recursos computacionais podem ser estritamente atribuídos a fluxos de dados específicos, levando a disponibilização de serviços de *Bandwidth on Demand* (BoD). Exemplos desse tipo de solução de provisionamento baseado em SDN/OpenFlow são o OESS, (Tepsuporn *et al.*, 2015; Globalnoc, 2016), utilizado inicialmente no *backbone* de educação e pesquisas da Internet2 nos Estados Unidos, e o DynPaC (Mendiola *et al.*, 2015), desenvolvido para utilização no *backbone* da GEANT.

Contudo, as soluções baseadas em SDN/OpenFlow como OESS e DynPaC, visam o provisionamento desse tipo de circuitos em redes **integralmente** SDN/OpenFlow. Isso é um problema visto que para o estabelecimento desse tipo de rede é, em geral, necessário a aquisição de novos recursos, tanto de novos enlaces para operar em paralelo a rede de produção IP tradicional quanto de novos equipamentos (*switches*) que suportem tecnologias de SDN/OpenFlow. Uma solução para esse problema é a utilização de redes SDN/OpenFlow **híbridas**, onde o tráfego IP tradicional de propósitos gerais roteado tradicionalmente pode ser encaminhado em paralelo ao tráfego destinado a uma aplicação específica, mas que se baseia nas regras de decisão de SDN/OpenFlow. Esse tipo de solução permite a utilização de equipamentos que já estejam em operação, mas que proveem algumas funcionalidades de SDN/OpenFlow.

Este trabalho, então, visa o desenvolvimento de uma solução para serviços de provisionamento de circuitos virtuais dinâmicos do tipo L2VPN *Ethernet/VLAN* em redes de *backbone* SDN/OpenFlow híbridas já em operação, em que o encaminhamento dos pacotes dos circuitos virtuais compartilha os mesmos recursos de rede já

disponíveis, como enlaces físicos e roteadores/*switches*. Essa solução é uma solução aberta destinada a redes de *backbone* de provedores de serviço de rede em geral (não somente voltada a redes de educação e pesquisa) ou até mesmo para redes de *campus*, que inclui como principal vantagem um serviço de conectividade privada sem custos adicionais e de alta capacidade baseada na agregação das vantagens oferecidas por uma tecnologia bem compreendida, como o *Ethernet*, aliada ao novo paradigma de SDN/OpenFlow.

### 3 | CARACTERÍSTICAS PRINCIPAIS DA APLICAÇÃO PROPOSTA

Uma rede de *backbone* IP consiste de múltiplos roteadores do tipo *Provider Edge* (PE), que conectam o roteador de borda *Customer Edge* (CE) à rede de núcleo do provedor, e roteadores do tipo *Provider* (P), os quais funcionam como um roteador de trânsito na rede de núcleo entre os PE, conforme mostrado na Figura 1.

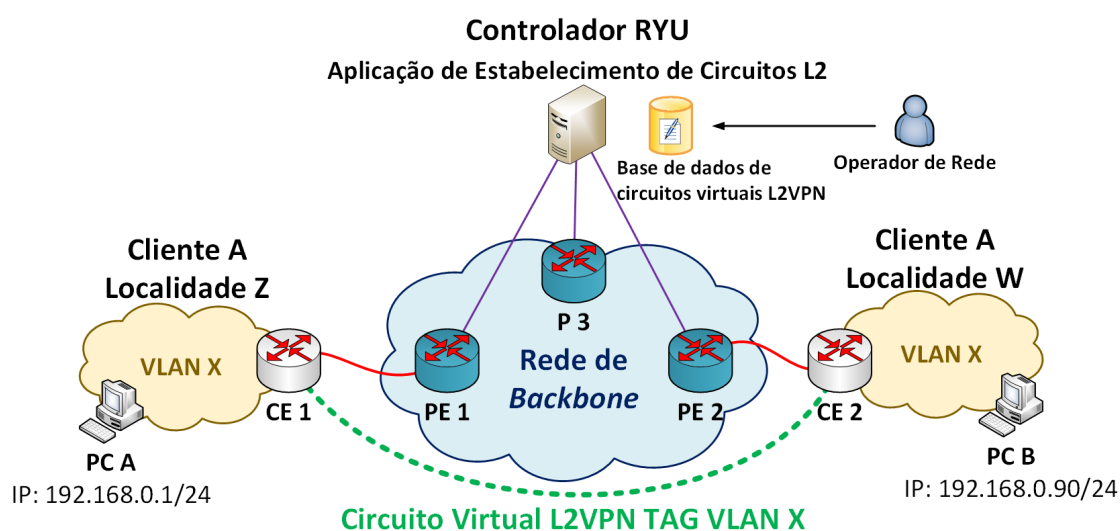


Figura 1: Modelo básico da visão da aplicação de estabelecimento de circuitos L2.

Na solução proposta, os administradores de rede do cliente informam aos operadores de rede do provedor as *tags* de VLAN 802.1q (IEEE (Institute of Electrical and Electronics Engineers), 2014) utilizadas em cada localidade nas quais desejam estabelecer um circuito L2VPN, e os operadores somente precisam armazenar as informações do circuito (como, por exemplo, *tag* de VLAN e roteadores PE envolvidos) em uma base de dados. A aplicação, então, é responsável por direcionar o tráfego de maneira adequada através da rede de núcleo, baseando-se nas informações obtidas através da base de dados e utilizando o algoritmo de escolha de menor caminho *Shortest Path First* (SPF) (Misa e Frana, 2010).

Nesta seção, primeiramente, são descritas as características principais do protocolo OpenFlow utilizadas nessa solução. Em seguida, é descrito como o operador de rede configura sua base de dados de estabelecimento de circuitos L2VPN. Em sequência, é descrito a utilização do algoritmo de menor caminho SPF

para estabelecimento entre os roteadores PE.

### **A. Características Principais do Protocolo OpenFlow**

O OpenFlow é um protocolo que permite que tabelas de fluxos em *switches* e roteadores sejam remotamente gerenciadas por um controlador. O protocolo define um fluxo como uma tupla com valores tais como: porta física de entrada, endereço MAC de origem e destino, campo tipo do cabeçalho *Ethernet*, identificador de VLAN, endereços IP de origem e de destino, campo protocolo do cabeçalho IP, porta de origem e destino do protocolo de camada de transporte. O *pipeline* de tabelas de fluxos em um *switch* OpenFlow mapeia a definição de um fluxo através dessa tupla em uma ação a ser tomada nos pacotes que pertencem a esse fluxo. Algumas dessas ações podem ser descartar os pacotes, encaminhá-los em uma porta física específica ou em um conjunto delas, ou enviar os pacotes para o controlador. Em caso de um pacote não corresponder a nenhuma das entradas da tabela de fluxos, o *switch* pode ser configurado para armazenar em um *buffer*, encapsular e enviar o pacote ao controlador para inspeção. Quando o controlador toma a decisão referente ao que fazer com todos os pacotes com aquela característica descrita pela tupla de campos do cabeçalho, ele adiciona uma entrada para esse fluxo na tabela de fluxos para armazenar essa decisão. Além disso, existe também a opção do controlador instalar entradas pró-ativamente para que esse processo não seja novamente repetido. Essas entradas são também conhecidas como regras.

O OpenFlow também implementa um conjunto de mensagens de sinalização e controle trocadas entre o controlador e o *switch*. Essas mensagens são responsáveis por realizar diversas ações como: verificação de características, configuração, modificação de estados, leitura de estados, envio de pacotes e mensagens de barreira. Dentre os tipos de mensagens mais relevantes estão o *packet\_in*, o *packet\_out* e o *flow\_mod*. O *packet\_in* é uma mensagem assíncrona enviada do *switch* para o controlador para notificar a chegada de um fluxo não classificado. O *packet\_out* é uma mensagem enviada do controlador para o *switch*, em resposta a um *packet\_in*, indicando qual ação deve ser tomada para aquele pacote. Já o *flow\_mod* é também enviado do controlador para o *switch* para modificar o estado do mesmo, podendo realizar diversos comandos como adição e modificação de entradas na tabela de fluxos.

Em nossa solução de L2VPN, o controlador realiza algumas funções principais como: o procedimento de descoberta de topologia; verificação dos *endpoints* (*switches* PE) do circuito virtual L2VPN a ser estabelecido; escolha do menor caminho entre todos os *switches* na rede e encaminhamento do pacote na porta do *switch* associada ao caminho entre os *endpoints*; realização do VLAN *stitching*, garantindo a consistência do circuito fim-a-fim. Essas funções serão descritas em mais detalhes adiante.

### **B. Verificação dos Endpoints e VLAN Stitching**

A aplicação proposta se baseia nos dados armazenados em duas bases de



dados: (i) gerenciada pelo operador de rede para armazenar a configuração do circuito virtual; (ii) utilizada para o processo de VLAN *stitching* no núcleo da rede. No caso da base de configuração a cada cinco minutos, cada entrada de configuração é checada em sequência em busca de modificações.

Na base de dados de configuração, cada entrada é representada por uma tupla de seis valores: porta *switch* PE 1, *switch* PE 1, VLAN ID de acesso no *switch* PE 1, porta *switch* PE 2, *switch* PE 2 e VLAN ID de acesso no *switch* PE 2.

Para lidar com as situações em que um mesmo cliente utiliza *tags* de VLANs distintas nos locais em que é atendido, e para decisão de qual *tag* é utilizada no núcleo da rede sem que haja sobreposição, outra base de identificação é utilizada. Nela, cada circuito passa a ser identificado por uma tupla de quatro valores: *switch* PE 1, VLAN ID de acesso no *switch* PE 1, *switch* PE 2, VLAN ID de acesso no *switch* PE 2. A partir desses dados, seleciona-se sequencialmente em função das *tags* já em uso no núcleo, qual a próxima a ser utilizada e armazena-se essa informação na base, conforme ilustrado na Figura 2.

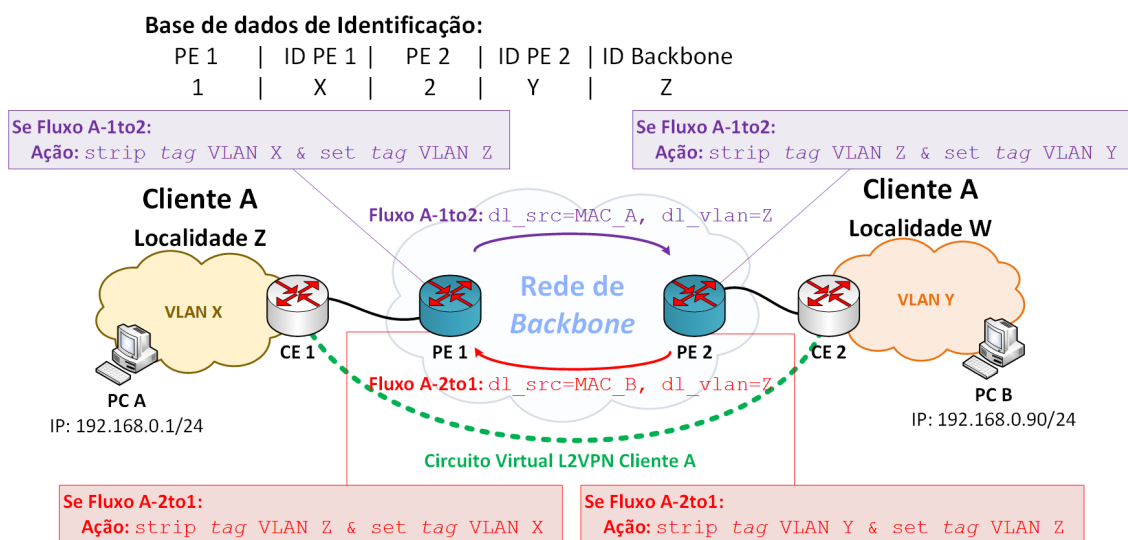


Figura 2: Modelo básico da visão da aplicação de estabelecimento de circuitos L2 para processo de VLAN *stitching*.

### C. Escolha do Menor Caminho e Estabelecimento do Circuito

A cada pacote do tipo *packet\_in*, o controlador analisa as informações dos cabeçalhos do pacote, verificando para cada entrada de configuração se: (i) o switch de entrada do pacote é um dos *endpoints* do circuito; (ii) a porta de entrada do *switch* é referente a um dos circuitos; (iii) a *tag* de VLAN do pacote é igual ao VLAN ID de entrada de algum circuito. Caso essas informações se confirmem, o *switch* realiza as ações de associar o endereço MAC de origem do pacote ao *endpoint*, de substituir a *tag* de entrada pela ID de *backbone* do circuito e de encaminhar o pacote. Ao chegar no *endpoint* de destino a operação inversa é realizada. Caso o *switch* que recebe o pacote não seja um *endpoint*, mas a *tag* seja associada a um ID de circuito no *backbone* e o MAC de origem esteja vinculado a um dos *endpoints*, o pacote é encaminhado

para o outro *endpoint*. Então, para estabelecer o caminho entre o *switch* que envia a mensagem de *packet\_in* e o *switch* PE de destino do circuito, a aplicação utiliza o algoritmo de *Shortest Path First (Misa e Frana, 2010)*. Dessa forma, a aplicação é capaz de estabelecer o circuito L2VPN fim-a-fim reativamente a cada mensagem de *packet\_in*, encaminhando o fluxo à porta de saída de cada *switch*, obtida em função do algoritmo SPF em conjunto com as bases de dados de configuração e de identificação de circuitos. Para evitar que ocorram *loops* na topologia utilizou-se também o protocolo *Spanning Tree (STP)*.

A aplicação proposta instala uma regra com *soft-timeout* de 60 segundos para poder lidar com o processo de alteração da configuração de um circuito. Então, após uma mudança por parte do operador de rede, há a necessidade de aguardar: (a) um período de 60 segundos de inatividade para que a entrada da tabela de fluxo de cada *switch* no caminho seja deletada; (b) um período de 5 minutos para que a configuração possa ser verificada novamente e o novo circuito possa ser estabelecido a partir de novas mensagens de *packet\_in* - podendo ambos os intervalos (a) e (b) ocorrerem concorrentemente.

#### 4 | IMPLEMENTAÇÃO E AVALIAÇÃO DO FUNCIONAMENTO DA APLICAÇÃO

Nesta seção, buscou-se avaliar o tempo de convergência da aplicação visto pelo *switch* CE, definido como o intervalo de tempo entre o início do funcionamento da aplicação e o momento em que os protocolos STP e SPF já convergiram, ou seja, quando os *switches* PE OpenFlow iniciam o encaminhamento de pacotes baseados nas mensagens de *packet\_in* e *packet\_out*. Para tal, os resultados foram subdivididos em dois cenários: topologias emuladas e estudo de caso da implantação no *Backbone REDECOMEP-Rio*. A avaliação de desempenho ilustra como o procedimento para o estabelecimento do circuito é eficaz e como o sistema escala adequadamente, aparentando ter pouco *overhead* em função da complexidade da rede. A complexidade, nesse caso, é considerada como o número total de nós, isto é, o número total de *switches* mais *hosts*.

Para validar a proposta nos cenários de topologias emuladas, construiu-se um protótipo utilizando o OpenVswitch (um *switch* OpenFlow via *software*) e o controlador OpenFlow Ryu, por meio do emulador de redes Mininet (Lantz *et al.*, 2010). A aplicação Ryu instala as regras nos *switches* do caminho do circuito L2VPN por meio do protocolo OpenFlow v1.3. Nesse protótipo, todos os procedimentos de verificação do circuito e de escolha de menor caminho foram desenvolvidos conforme descritos nas Seções 3.b e 3.c.

Para avaliar o funcionamento da aplicação, foi medido o tempo de convergência na rede, considerando o uso do algoritmo STP, disponibilizado por padrão no controlador Ryu, com o algoritmo SPF, desenvolvido para a aplicação, e com algoritmo proposto de estabelecimento de conexões L2VPN. Emulou-se o processo de estabelecimento

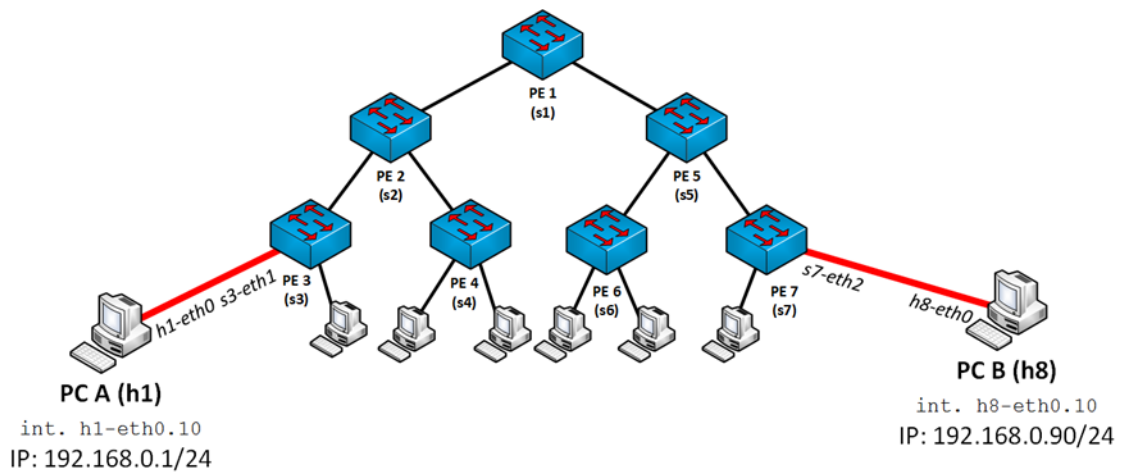


de uma conexão entre dois *hosts* de um mesmo cliente, assumindo que o circuito já estava previamente descrito no arquivo de configuração. Foi medido o intervalo de tempo entre o primeiro pacote ARP *Request* enviado e sem resposta e o primeiro pacote ARP *Request* que recebe resposta. Os resultados são apresentados com um intervalo de confiança de 95%.

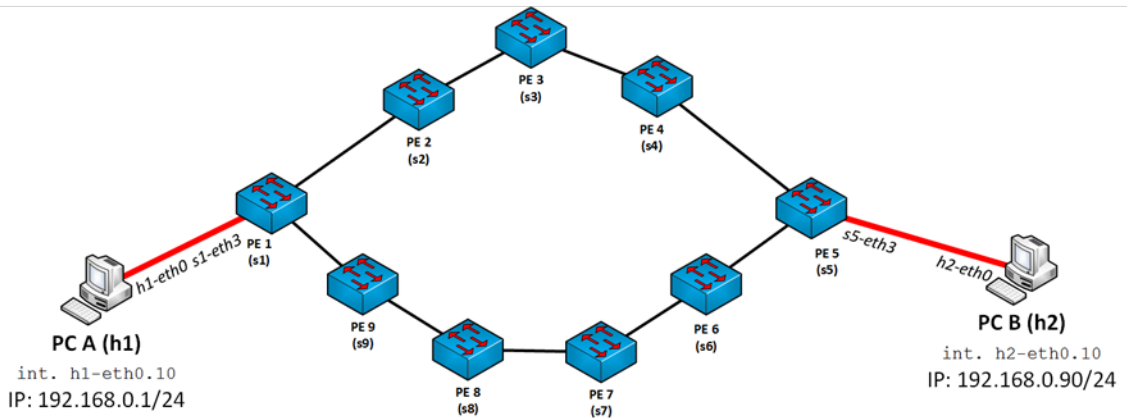
### A. Cenários Emulados

Para realizar a avaliação do funcionamento da aplicação, emulou-se três diferentes topologias. Utilizou-se duas topologias padrão do Mininet e uma customizada, conforme descritas em detalhes adiante. Para geração dos pacotes para estabelecimento de conexão e geração dos pacotes ARP, utilizamos a ferramenta FPING (Schemers, 2007), com intervalos entre mensagens ICMP *Echo Request* de 10 ms. Os intervalos entre os pacotes ARP foram checados através da ferramenta *tcpdump* (Fuentes e Kar, 2005), capturando pacotes na interface do *host* de origem.

**1. Topologia Linear com 2 Switches:** Como avaliação inicial, utiliza-se a topologia mais simples disponível no emulador com dois *switches*. Os *switches* são conectados diretamente e dois *hosts* são conectados cada um a um dos *switches*. **2. Topologia Fat-tree com Três Níveis e Sete Switches:** A topologia *fat-tree* é um tipo de topologia voltada para *data centers*. Realizou-se a avaliação de funcionamento em uma topologia de três níveis com sete *switches* e oito *hosts* conectados de acordo com a Figura 3(a). **3. Topologia Emulada do Backbone REDECOMEP-Rio:** Como a aplicação visa ser implementada na rede em produção da REDECOMEP-Rio, optou-se por realizar uma emulação em uma infraestrutura similar à sua topologia real. A topologia de *backbone* da REDECOMEP-Rio, atualmente, é composta por nove Pontos de Presença (PoP, do inglês *Point of Presence*) espalhados por uma topologia em anel com múltiplas redundâncias. A Figura 3(b) exibe o *layout* da topologia da rede em produção simplificada. Foram conectados nove *switches* em anel com dois *hosts*. Nesse cenário, cada *host* faz o papel dos *switches* CE que devem ser conectados aos *switches* PE de *backbone*. Vale ressaltar que como a topologia é em anel, ocorrem *loops* do *Spanning Tree*.



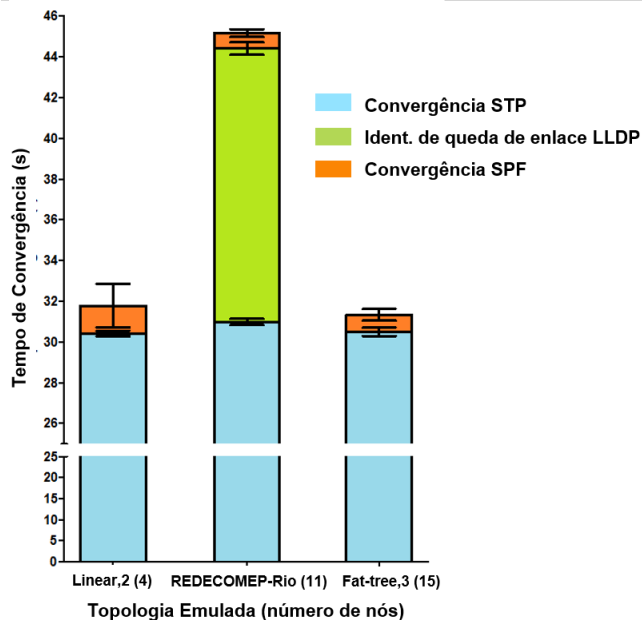
(a) Fat-tree com sete switches e oito hosts.



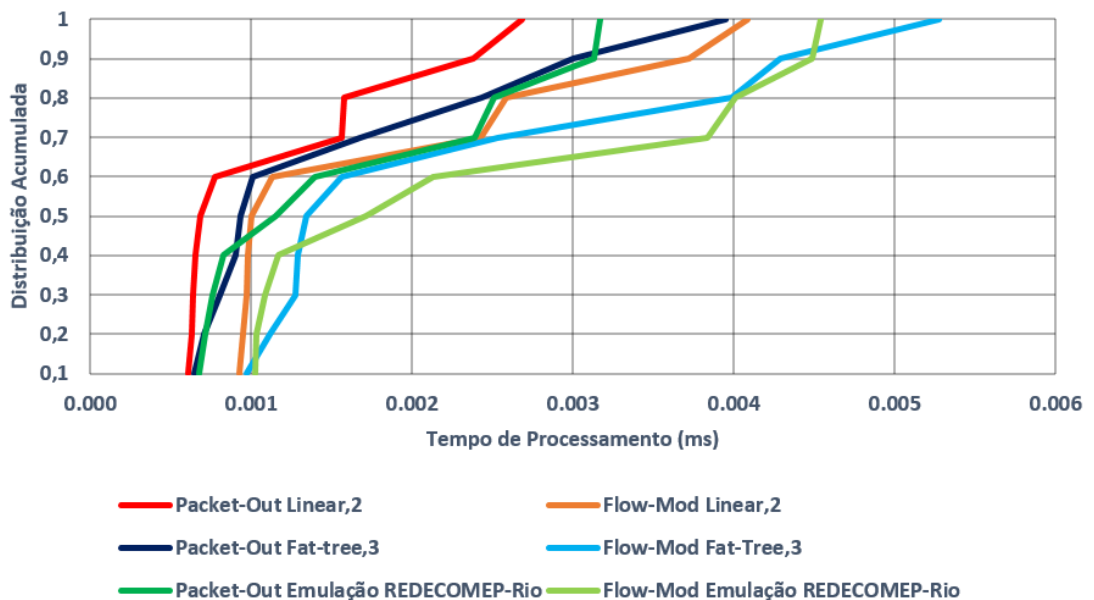
(a) Emulação da REDECOMEP-Rio com nove switches e dois hosts.

Figura 3: Layout das topologias emuladas.

Os resultados referentes aos cenários emulados podem ser visualizados na Figura 4(a). O tempo de convergência é em grande parte dominado pelo tempo de convergência do STP, que para todos os casos apresentados é de 30s. Isso ocorre de acordo com o apresentado em (Ryu, 2016). A partir do instante em que o STP converge a aplicação demora somente 1,5s para iniciar o encaminhamento de pacotes. Assim como os casos em que não ocorrem *loops*, o tempo de convergência é em grande parte dominado pelo tempo de convergência do STP, cerca de 30s. Porém, após uma porta entrar em estado BLOCKED pelo STP, é necessário mais 13,5s para que o LLDP, implementado por padrão no Ryu, identifique que o enlace está bloqueado e a aplicação, a partir de então, possa recalculer os caminhos. Desde o momento em que o enlace passa a estar bloqueado, a aplicação leva cerca de 1s para encaminhar os pacotes.



(a) Tempo total de convergência em relação ao número total de nós na rede.



(b) CDF do tempo de processamento das mensagens de *packet\_out* e *flow\_mod*.

Figura 4: Resultados obtidos para os cenários emulados.

Cabe ainda ressaltar que a partir do momento em que a aplicação já convergiu, o estabelecimento do circuito depende somente da latência entre os *switches* e do tempo em que as mensagens de *packet\_out* e *flow\_mod* demoram para serem processadas e enviadas pelo controlador. Como pode ser observado na Figura 4(b), a função de distribuição acumulada (CDF) dos tempos de processamento dessas mensagens demonstra que esse intervalo representa cerca de 0,017% do tempo total de convergência para o pior dos casos de processamento de mensagens *flow\_mod*, ou seja, a topologia emulada *fat-tree*. Logo, isso mostra que a partir da convergência da aplicação, o tempo para estabelecimento do circuito é muito pequeno.

## B. Estudo de Caso da Implantação Inicial no Backbone da REDECOMEP-Rio

Apresenta-se, a seguir, a avaliação do sistema protótipo na REDECOMEP-Rio. O estudo de caso foi executado em uma configuração com equipamentos híbridos, ou seja, realizando o encaminhamento dos pacotes da rede em produção no mesmo enlace que os pacotes SDN/OpenFlow. O cenário descrito foi avaliado utilizando os equipamentos: 2 (PE) x Cisco ASR9000 rodando IOS XR 5.1.3 equipados com pelo menos 6 GB DRAM; 1 (CE) x Cisco Catalyst WS-C3560G-24TS rodando IOS 12.2(50)SE5; 1 (CE) x Cisco Catalyst WS-C3750G-24TS-1U rodando 12.2(25)SEE2. A associação entre os roteadores Cisco ASR9000 e o controlador ocorre *in-band*. Todos os dispositivos são conectados com enlaces de 1 Gbps de capacidade, conforme mostrado na Figura 5. A conexão entre o *switch* CE e o roteador PE ocorre através de uma interface do tipo tronco com permissão de tráfego das VLANs 802.1q 10 e 100.

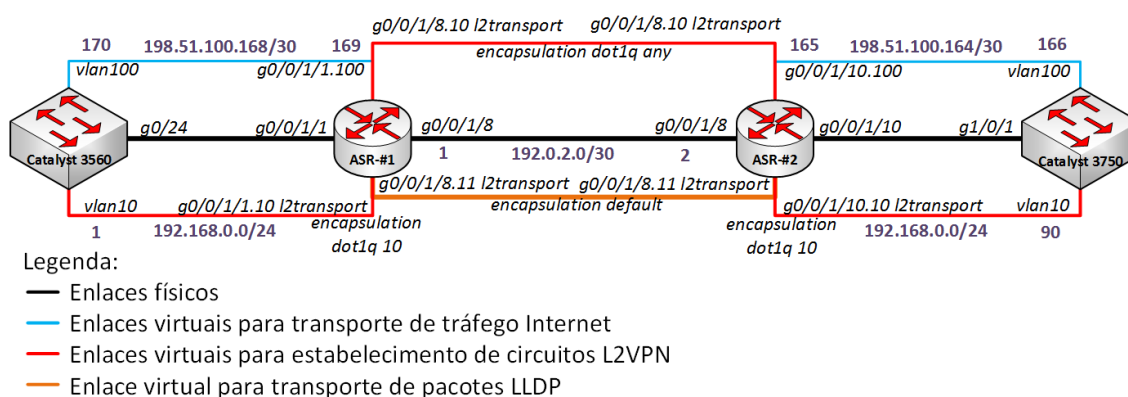


Figura 5: Cenário de estudo de caso da avaliação do funcionamento da aplicação de estabelecimento de circuitos L2VPN no *backbone* da REDECOMEP-Rio.

De acordo com o que pode ser observado na Figura 5, fez-se necessário o estabelecimento de dois enlaces virtuais entre os roteadores PE para o correto funcionamento da aplicação. Além do transporte dos pacotes com *tags* de VLAN 802.1q para o estabelecimento de circuitos, também é necessário o transporte dos pacotes LLDP para o mapeamento da topologia. Desse modo, para cada tipo de enlace virtual foi permitido um determinado tipo de encapsulamento de pacotes. Isto é, para o transporte dos pacotes dos circuitos com encapsulamento dot1q configurou-se uma interface com *encapsulation dot1q any*, em vermelho na Figura 5, enquanto que para os pacotes de controle LLDP sem encapsulamento foi configurado outro tipo de interface com *encapsulation default*, em laranja na Figura 5.

Assim como o exposto na Seção 4.a., o tempo de convergência da aplicação também foi avaliado, *i.e.*, o intervalo de tempo entre o pacotes ARP *Request* inicial e o que recebe resposta foi determinado. Entretanto, em vez de utilizar a ferramenta FPING foi utilizado o utilitário *ping* disponível por padrão nos *switches*, também com intervalo de 10ms entre ICMP *Echo Request*, e para análise dos resultados a ferramenta *tcpdump*. As requisições foram enviadas a partir do *switch* denominado Catalyst 3560 na Figura 5.

A Figura 6 exibe o resultado comparativo entre as CDFs dos tempos de convergência do cenário emulado de topologia linear descrito na Seção 4.a. e do estudo de caso descrito na presente seção. Pode-se observar uma diferença de até 8s entre o menor valor para o cenário emulado e o maior valor do estudo de caso. Essa diferença ocorre, principalmente, devido ao intervalo entre tentativas de associação entre controlador e roteador ASR9000 serem fixadas em 8s, enquanto que o intervalo referente ao OpenVswitch ser de apenas 1s.

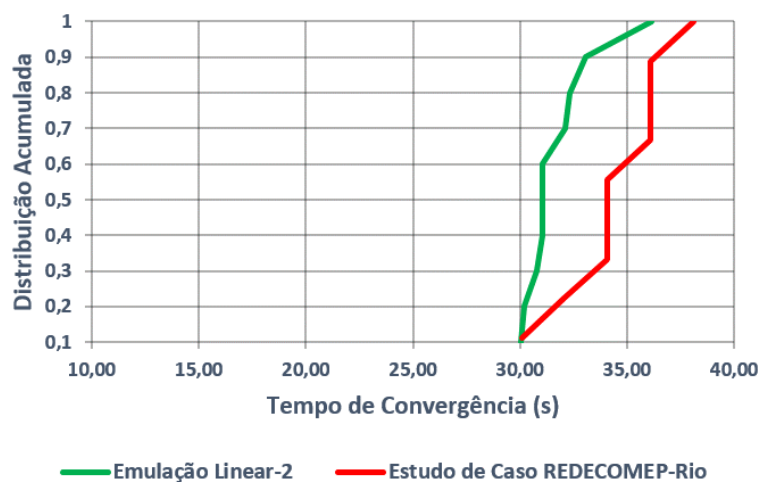


Figura 6: CDF do tempo total de convergência da aplicação para os cenários emulado linear e de estudo de caso.

## 5 | CONCLUSÃO

Redes de *backbone* têm como uma de suas principais demandas o provimento de serviços de estabelecimento de redes virtuais privadas. Um dos principais tipos de redes são as VPNs de camada 2 ou L2VPNs. A solução proposta de estabelecimento desse tipo de circuitos instala regras reativamente nos *switches* da rede utilizando o controlador OpenFlow Ryu. Essa solução se difere, principalmente, das disponíveis no mercado atualmente, pois não é necessário em nenhum momento que o operador tenha conhecimento do funcionamento do protocolo, nem sequer tenha que alterar as configurações dos equipamentos. As avaliações mostram que o sistema escala bem para redes com até quinze nós de *backbone*, dependendo em grande parte do tempo de convergência do protocolo *Spanning Tree*, e se há ou não *loops* na topologia. Esse efeito ocorre devido ao fato de que o protocolo LLDP leva cerca de 13s após a convergência do STP para detectar o bloqueio de um enlace. Além disso, a diferença encontrada entre o cenário emulado e o implantado, mostra que o tempo de associação entre o controlador e o *switch* OpenFlow tem um pequeno impacto no tempo total de convergência da aplicação. Ademais, o impacto dos atrasos entre os *switches*, e entre eles e o controlador no tempo de estabelecimento de circuito é pequeno, devido ao fato dos tempos de processamento de mensagens de *packet\_out* e *flow\_mod* somados a esses atrasos serem muito pequenos em relação ao tempo

total de convergência. Assim, após o período de inicialização da rede, o tempo para estabelecimento automático da L2VPN é inferior a 10ms, o que é um excelente tempo quando comparado ao processo tradicional para a configuração e estabelecimento desse tipo de circuito.

Portanto, em função dos resultados obtidos, a implementação da aplicação em todo o *backbone* em produção da REDECOMEP-Rio está em fase de finalização, onde o impacto de sua utilização nos usuários da rede deve ser avaliado, posteriormente.

## REFERÊNCIAS

- BOBYSHEV, A. et al. **Lambda Station: On-demand Flow Based Routing for Data Intensive Grid Applications over Multitopology Networks**. 2006 3rd International Conference on Broadband Communications, Networks and Systems, 2006, IEEE. p.1-9.
- CHAVES FILHO, G. D. A. **Comutação baseada em caminhos: uma solução SDN para problema de migração de máquinas virtuais em Data Centers**. 2015. (Mestrado em Informática). Centro de Ciências Exatas e Tecnologia, Universidade Federal do Estado do Rio de Janeiro, Rio de Janeiro, RJ, Brasil.
- ESNET. **OSCARS: On-Demand Secure Circuits and Advance Reservation System**. 2016. Disponível em: < <http://www.es.net/engineering-services/oscars/> >. Acesso em: 20 de fevereiro de 2019.
- FUENTES, F.; KAR, D. C. **Ethereal vs. Tcpcdump: a comparative study on packet sniffing tools for educational purpose**. Journal of Computing Sciences in Colleges, v. 20, n. 4, p. 169-176, 2005. ISSN 1937-4771.
- GEANT. **About AutoBAHN**. 2017. Disponível em: < [http://geant3.archive.geant.net/service/autobahn/about\\_autoBAHN/Pages/About\\_autoBAHN.aspx](http://geant3.archive.geant.net/service/autobahn/about_autoBAHN/Pages/About_autoBAHN.aspx) >. Acesso em: 20 de fevereiro de 2019.
- GLOBALNOC. **OESS: Open Exchange Software Suite**. 2016. Disponível em: < <http://globalnoc.iu.edu/sdn/oess.html> >. Acesso em: 20 de fevereiro de 2019.
- GUOK, C. et al. **Intra and Interdomain Circuit Provisioning Using the OSCARS Reservation System**. 2006 3rd International Conference on Broadband Communications, Networks and Systems, 2006, IEEE. p.1-8.
- IEEE (INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS). **IEEE 802: Local and Metropolitan Area Network Standards**. IEEE Standard 802.1q 2014.
- KATRAMATOS, D. et al. **The TeraPaths Testbed: Exploring End-to-End Network QoS**. Testbeds and Research Infrastructure for the Development of Networks and Communities, 2007. TridentCom 2007. 3rd International Conference on, 2007, IEEE. p.1-7.
- KNIGHT, P.; LEWIS, C. **Layer 2 and 3 Virtual Private Networks: Taxonomy, Technology, and Standardization Efforts**. IEEE Communications Magazine, v. 42, n. 6, p. 124-131, 2004. ISSN 0163-6804.
- KREUTZ, D. et al. **Software-Defined Networking: A Comprehensive Survey**. Proceedings of the IEEE, v. 103, n. 1, p. 14-76, 2015. ISSN 0018-9219.
- LANTZ, B.; HELLER, B.; MCKEOWN, N. **A network in a laptop: rapid prototyping for software-defined networks**. Proceedings of the 9th ACM SIGCOMM Workshop on Hot Topics in Networks, 2010, ACM. p.19.



- MENDIOLA, A. et al. **DynPaC: A Path Computation Framework for SDN**. 2015 Fourth European Workshop on Software Defined Networks, 2015, IEEE. p.119-120.
- MISA, T. J.; FRANA, P. L. **An interview with Edsger W. Dijkstra**. Communications of the ACM, v. 53, n. 8, p. 41-47, 2010. ISSN 0001-0782.
- MORAES, L. F. M. D.; ALBUQUERQUE, M. P. D.; RIBEIRO FILHO, J. L. **Infraestrutura Redes de Alta Velocidade no Rio de Janeiro: história e estado da arte**. In: (Ed.). A História da Telessaúde da Cidade para o Estado do Rio de Janeiro. Rio de Janeiro: EdUERJ, 2015. ISBN 978-85-7511395-0.
- OSBORNE, E. D.; SIMHA, A. **Traffic engineering with MPLS**. Cisco Press, 2002. ISBN 1587050315.
- RAO, N. S. et al. **UltraScienceNet: Network Testbed for Large-Scale Science Applications**. IEEE Communications Magazine, v. 43, n. 11, p. S12-S17, 2005. ISSN 0163-6804.
- RYU. **Framework Community: Ryu SDN controller**. 2016. Disponível em: < <https://osrg.github.io/ryu/> >. Acesso em: 20 de fevereiro de 2019.
- SCHEMERS, R. J. **FPING Man Page**. 2007. Disponível em: < <http://fping.sourceforge.net/man/> >. Acesso em: 20 de fevereiro de 2019.
- SHARAFAT, A. R. et al. **MPLS-TE and MPLS VPNs with OpenFlow**. ACM SIGCOMM Computer Communication Review, 2011, ACM. p.452-453.
- TEPSUPORN, S. et al. **A multi-domain SDN for dynamic layer-2 path service**. Proceedings of the Fifth International Workshop on Network-Aware Data Management, 2015, ACM. p.2.
- YANG, X. et al. **GMPLS-based Dynamic Provisioning and Traffic Engineering of High-Capacity Ethernet Circuits in Hybrid Optical/Packet Networks**. Proceedings IEEE INFOCOM 2006. 25TH IEEE International Conference on Computer Communications, 2006, IEEE. p.1-5.
- ZHENG, X. et al. **CHEETAH: Circuit-switched High-speed End-to-End Transport Architecture testbed**. IEEE Communications Magazine, v. 43, n. 10, p. S11-S17, 2005.



## SISTEMA DISTRIBUÍDO PARA DETECÇÃO DE AMEAÇAS EM REDES UTILIZANDO DEEP LEARNING

### Fábio César Schuartz

Universidade Tecnológica Federal do Paraná,  
CPGEI  
Curitiba - Paraná

### Mauro Sérgio Pereira Fonseca

Universidade Tecnológica Federal do Paraná,  
CPGEI  
Curitiba - Paraná

### Anelise Munaretto

Universidade Tecnológica Federal do Paraná,  
CPGEI  
Curitiba - Paraná

**RESUMO:** A detecção de ameaças na Internet é um fator essencial para manter a segurança de dados e informações. Um sistema de detecção de ameaças tenta prevenir que esses ataques ocorram através da análise de padrões e do comportamento do fluxo de dados na rede. Este trabalho apresenta uma extensão para a plataforma distribuída de detecção e análise de dados em grande fluxo, através do uso de deep learning para redução do espaço de características. A avaliação do sistema se baseia através da acurácia, do número de falsos positivos e de falsos negativos, onde cada classificador apresentou melhor acurácia ao utilizar 5 e 13 atributos. Ainda, o sistema apresentou menor número de falsos positivos e negativos, permitindo a detecção de ameaças

em tempo real sobre um grande volume de dados, com maior precisão.

**PALAVRAS-CHAVE:** Aprendizado de máquinas, aprendizagem profunda, grande volume de dados, sistema de detecção de ameaças, tempo real.

**ABSTRACT:** Detecting threats on the Internet is a key factor in maintaining data and information security. An intrusion detection system tries to prevent such attacks from occurring through the analysis of patterns and behavior of the data flow in the network. This paper presents an extension to the distributed large data flow detection and analysis platform through the use of deep learning to reduce the feature space. The evaluation of the system is based on accuracy, number of false positives and false negatives, where each classifier presented better accuracy using 5 and 13 attributes, besides having fewer false positives and negatives, allowing the detection of real-time threats over a large volume of data with greater accuracy.

**KEYWORDS:** Big data, deep learning, intrusion detection system, machine learning, real-time.

### 1 | INTRODUÇÃO

O crescimento no uso da Internet criou uma necessidade maior em proteger os dados

e informações guardadas em servidores centralizados e distribuídos, principalmente em sistemas acessados através de uma rede pública. Pessoas ganham benefícios e companhias geram lucro gerenciando seus recursos e transações através da rede, criando maiores oportunidades para usuários maliciosos roubarem informações pessoais e secretas. Segundo Leu et al. (2015), nos últimos anos diversas estatísticas mostram um número crescente de invasões reportadas no Symantec Global Internet Security Threat Report.

Um sistema de detecção de intrusão (IDS - Intrusion Detection System) é um sistema utilizado para monitorar as atividades de outro sistema ou de uma rede, procurando por atividades maliciosas e produzindo mensagens de alerta para a estação de controle, conforme Tan et al. (2014). Um IDS consiste de dois componentes: detecção por assinaturas e detecção por anomalias. A detecção por assinaturas é utilizada para procurar por ataques baseado em padrões extraídos de invasões conhecidas, enquanto a detecção por anomalias tenta descobrir ataques baseado no comportamento do tráfego que difere do padrão normal de comportamento.

Uma maneira de estudar os ataques ocorridos na rede é utilizar técnicas de aprendizagem de máquina, procurando por padrões no tráfego da rede. Um classificador pode estudar diversos exemplos de entradas que produzem resultados conhecidos, através de um aprendizado supervisionado, conforme Harbi e Bahri (2013). Porém, os ataques evoluem e podem não seguir os padrões de ataques anteriores. Assim, as técnicas de aprendizado não supervisionado podem apresentar melhores resultados, pois procuram por variações em padrões conhecidos, ao invés de classificar o tráfego em apenas uma categoria.

O aumento no volume, velocidade e variedade dos dados nas redes atuais demanda uma infraestrutura robusta de segurança. Monitorar e processar dados em altas taxas sem desperdiçar recursos é um enorme desafio atual, segundo Lopez et al. (2018). Uma área de pesquisa atual recebendo grande atenção é o de deep learning. Esta é uma sub-área avançada da aprendizagem de máquina, que permite sobrepujar as limitações do aprendizado superficial. Sua característica superior de camadas de aprendizagem pode resultar em desempenho superior ou equivalente, comparado às técnicas de aprendizado superficiais.

Este trabalho propõe uma expansão do sistema distribuído para detecção de ameaças em tempo real através da análise de grandes volumes de dados, pelo processamento por fluxo, apresentado no trabalho de Schuartz, Fonseca e Munaretto (2017). O objetivo é utilizar o método de deep learning na camada de pré-processamento de dados para obter uma redução no espaço de características dos dados, resultando assim, em maior acurácia na detecção de ameaças, com menor índice de falsos-positivos e falsos-negativos. Foram realizadas simulações para três casos: utilização dos 41 atributos existentes na base de dados (sem uso de deep learning), utilização da redução do espaço das características para 5 atributos e para 13 atributos. Resultados avaliados mostram que utilizando o método de deep learning

é possível obter resultados com maior acurácia e menor incidência de falsos-positivos e falsos-negativos. Para avaliação do sistema, utilizou-se um conjunto de dados com as classes marcadas, contendo dados normais e ataques, proveniente de uma base de dados de teste KDD'99, transformado em fluxos.

Este trabalho é dividido em cinco seções, onde a seção 2 apresenta os trabalhos relacionados. A seção 3 apresenta o sistema proposto. Os resultados obtidos são apresentados na seção 4. Por fim, a seção 5 finaliza o trabalho.

## 2 | TRABALHOS RELACIONADOS

Em 2017, um sistema de detecção de intrusão em redes (NIDS - Network Intrusion Detection System) baseado em anomalias foi construído por Van, Thinh e Sach (2017) utilizando técnicas de aprendizagem profunda (deep learning). Essa técnicas mostraram o grande poder dos modelos generativos com boa classificação, capazes de deduzir parte do seu conhecimento de dados incompletos e adaptar-se. O trabalho foi capaz de detectar intrusões baseadas em anomalias e classificá-las em cinco grupos, com acurácia baseada nas fontes de dados de rede.

Ainda em 2017, no trabalho de Kim e Aminanto (2017) são mostradas algumas limitações em IDSs anteriores que utilizam aprendizagem de máquina clássicas e introduzem o aprendizado de características, incluindo a construção, extração e seleção de características para superar os desafios. Também discutem algumas técnicas de deep learning e suas aplicações para utilização em IDS.

Em 2017, Alom e Taha (2017) apresentaram um trabalho sobre IDS utilizando técnicas de deep learning não-supervisionadas. As amostras de entrada são codificadas numericamente para a aplicação de técnicas não-supervisionadas, como Auto-Encoder (AE) e Restricted Boltzmann Machine (RBM), para extração de características e redução de dimensionalidade. Então é aplicado agrupamento iterativo k-means para aglomeração em um espaço dimensional menor com apenas 3 atributos.

Em 2017, Schuartz, Fonseca e Munaretto (2017) apresentaram uma proposta de uma plataforma distribuída para detecção de ameaças em tempo real, utilizando big data. A plataforma proposta é capaz de detectar diversos tipos de ataques com precisão acima de 90% e baixo número de falsos-positivos e falsos-negativos. O sistema utiliza 41 atributos da base de dados para o treinamento e detecção de ameaças, resultando em falha na detecção de alguns ataques em específicos, devido ao treinamento imparcial dos diversos tipos de ataques e pela baixa representatividade de tais ataques dentro da base de dados.

Em 2018, Shone et al. (2018) propôs uma técnica de deep learning para detecção de intrusões utilizando um auto-codificador não-simétrico de profundidade (NDAE - Nonsymmetric Deep Autoencoder) para aprendizado de características não-supervisionadas. O modelo de classificação proposto foi desenvolvido utilizando o TensorFlow e uma unidade de processamento gráfica (GPU - Graphics Processing

Unit) e foi avaliada através das bases de dados KDD Cup '99 e NSL-KDD.

Embora existam diversas técnicas e propostas na literatura para a detecção de intrusão, a maioria delas ainda não são eficientes o suficiente para trabalhar com um grande fluxo de dados (Big Data) em tempo real, enquanto outras propostas não apresentam acurácia suficiente. Este trabalho propõe a utilização do método deep learning para a redução do espaço de características dos fluxos de dados, permitindo assim o treinamento e a detecção de ameaças de maneira mais eficiente e precisos.

### 3 | ESQUEMA PROPOSTO

O sistema proposto visa expandir a plataforma aberta para coleta, distribuição, análise e processamento de dados proveniente de um fluxo de dados, apresentada em (SCHUARTZ; FONSECA; MUNARETTO, 2017). O processo consiste em, durante a fase de normalização das informações coletadas, utilizar o método de deep learning para reduzir o espaço de características dos dados, fornecendo maior acurácia na detecção de ameaças, maior rapidez no treinamento e detecção de anomalias, com redução no número de falsos-positivos e falsos-negativos.

O objetivo do método deep learning é aprender atributos hierárquicos de menor nível em atributos de maior nível. O método pode aprender características independentemente em múltiplos níveis de abstração e então descobrir funções de mapeamento complexas entre a entrada e a saída diretamente dos dados puros sem depender de atributos customizados por especialistas. Em abstrações de alto nível, os humanos normalmente não identificam as relações e conexões de uma entrada sensorial pura. Assim, a habilidade em aprender características complexas, também chamadas de extração de características, se torna necessária em vista do aumento na quantidade de dados (BENGIO et al., 2009).

A extração de características através do codificador automático empilhado (SAE - Stacked Auto Encoder) é capaz de reduzir a complexidade das características originais do conjunto de dados. Entretanto, além de ser um extrator de características, o SAE também pode ser utilizado para tarefas de classificação e aglomeração. É possível melhorar o processo de aprendizagem de características através da combinação da extração de características empilhadas com seleção de características com pesos. A extração de características do SAE é capaz de transformar as características originais em uma representação mais significativa ao reconstruir seus dados de entrada e fornecendo um meio de verificar a informação relevante nos dados capturados. O SAE pode ser eficientemente utilizado no aprendizado não-supervisionado em um conjunto de dados complexos.

A estrutura do SAE é apresentada na Figura 1. O primeiro auto-encoder recebe uma entrada e realiza uma redução nas características, procurando obter como saída, através de um decoder, a entrada original, pelo método de retropropagação (back-propagation).  $H_1$  representa a camada escondida (hidden layer) do auto-encoder 1. Na

sequência,  $H_1$  será a entrada do auto-encoder 2, que irá gerar  $H_2$  através de processo semelhante e que será a entrada do auto-encoder 3, resultando em  $H_3$ . Observando de uma percepção multi-camadas (multilayer), será o equivalente da entrada passar por 3 camadas escondidas, gerando a saída reduzida.

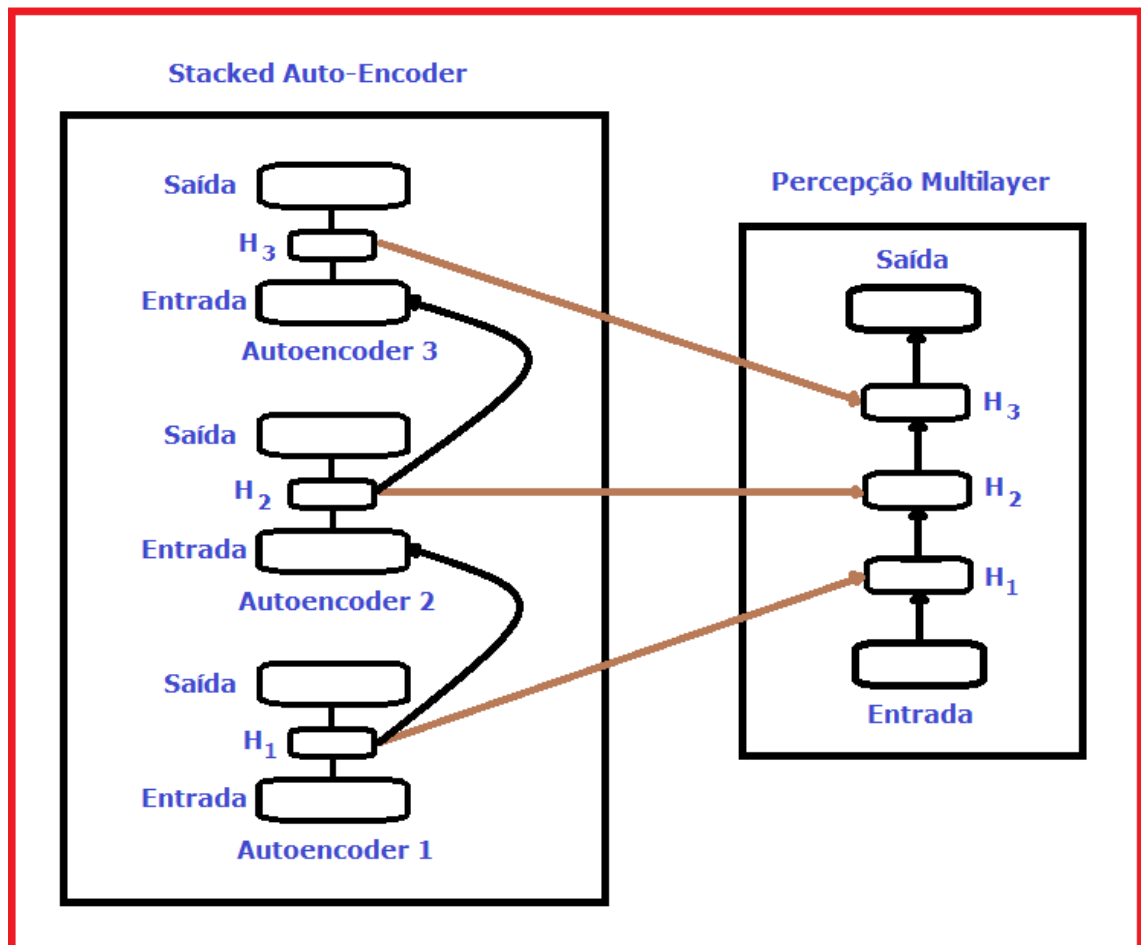


Fig. 1. Exemplo de uma estrutura Stacked Auto-Encoder, para três camadas. Fonte: autoria própria.

### 3.1 Auto-Encoder

Um auto-encoder é uma estratégia de rede neural profunda utilizada no aprendizado de características não-supervisionadas com codificação eficiente. O objetivo principal do AE é o aprendizado e representação (codificação) do dado, tipicamente com a finalidade da redução da dimensionalidade do dado. Esta técnica de AE consiste em duas partes: o codificador e o decodificador. Na fase de codificação, os mapas de amostras de entradas são reduzidas para um espaço de características de dimensão menor, contendo as características construtivas mais importantes. Este processo pode ser repetido até alcançar o espaço dimensional de característica desejada. Na fase de decodificação, são reconstruídas as características originais a partir de uma dimensão menor de características, através do processo de reversão. O diagrama conceitual do AE é mostrado na Figura 2. As transições de codificação e decodificação podem ser representadas por  $\Phi$  e  $\varphi$ :

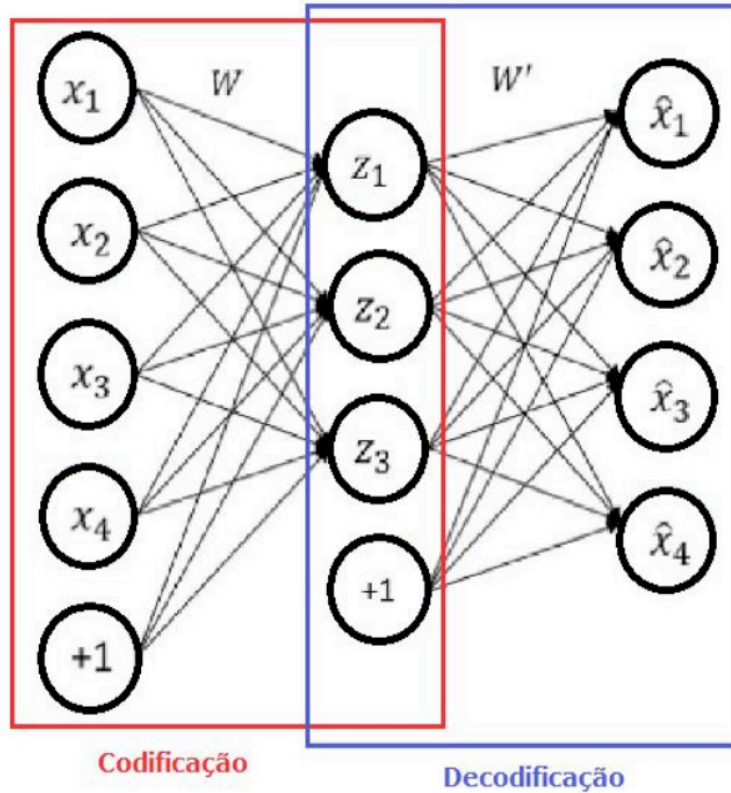


Fig. 2. Diagrama para um auto-encoder com as fases de codificação e decodificação. Fonte: autoria própria.

$$\phi: X \rightarrow F \quad (1)$$

$$\varphi: F \rightarrow X \quad (2)$$

$$\phi, \varphi = \underset{\phi, \varphi}{\operatorname{argmin}} \|X - (\phi, \varphi)\|^2 \quad (3)$$

Se considerarmos o AE mais simples com apenas uma camada escondida, onde a entrada é  $x \in \mathbb{R}^d = X$ , o qual é mapeada em  $X \in \mathbb{R}^p = F$ , então a expressão pode ser escrita na seguinte operação:

$$z = \sigma_1(W * x + b) \quad (4)$$

onde  $W$  é a matriz de peso e  $b$  é o bias.  $\sigma_1$  representa a função de ativação, tal como uma sigmóide ou uma unidade linear retificada (RLU). Considerando  $z$  novamente mapeado ou reconstruído em  $x'$ , com a mesma dimensão de  $x$ , a reconstrução pode ser expressa como:

$$x' = \sigma_2(W' * z + b') \quad (5)$$

Estas técnicas podem ser treinadas com o mínimo de erros de reconstrução:



$$L = \|x - x'\|^2 \equiv \|x - \sigma_2(W' * \sigma_1(W * x + b) + b')\|^2 \quad (6)$$

Normalmente o espaço de características de  $F$  possui a menor dimensão do espaço de características de entrada  $X$ , que pode ser associada como a representação comprimida da amostra de entrada. No caso de um AE de multicamadas, a mesma operação será incorporada a medida que for necessária nas fases de codificação e decodificação.

### 3.2 Conjunto de Dados

A base de dados KDD'99 (LEE; STOLFO; MOK, 1999) é um conjunto de dados de referência que foi simulado em ambiente de rede militar em 1998 e derivado em conjunto de atributos em 1999. O pacote da base de dados foi reunido e pré-processado em 41 atributos. A base de dados contém quatro categorias de ataques diferentes (DoS, R2L, U2R e probing), sendo um total de 22 tipos de ataques na base de treinamento e 14 tipos de ataques extras na base de teste que não existem na base de treinamento.

### 3.3 Protótipo do Sistema Proposto

Inicialmente, a coleta de dados é feita em um único ponto da rede. Esses dados foram extraídos de uma base de dados KDD'99, amplamente utilizada e testada na comunidade. Os dados são pré-processados e caracterizados em fluxos compostos de 41 atributos, utilizados para detectar 22 tipos de ataques. Estes fluxos de dados são encaminhados para o SAE (Stacked Auto-Encoder), composto por três camadas escondidas que irão realizar a redução no espaço de características de entrada de 41 atributos para 5 e 13 atributos principais. Esses mapas reduzidos serão então publicados na rede.

Em outro ponto da rede, três unidades distintas de processamento de fluxos irão receber as informações publicadas na rede, agindo, cada um, como um assinante para o mesmo fluxo de dados. Cada unidade irá, então, alimentar esse fluxo para a topologia criada no Apache Storm. Dentro do Storm, é criado um classificador utilizando a ferramenta Weka. Os classificadores foram treinados utilizando um conjunto de dados de treinamento do KDD'99. Cada unidade possuirá um algoritmo de aprendizagem de máquina diferente e irá processar o fluxo de dados recebido, caracterizando o mesmo em um tipo específico de ataque ou em um fluxo normal de dados. Por último, os resultados obtidos por cada unidade de processamento são enviados para uma única interface de visualização, onde serão exibidos. Cada unidade realizará três modelos de treinamento e classificação. Inicialmente será alimentado um fluxo contendo os 41 atributos, sem o uso de deep learning. Em seguida, o processo será repetido, porém com um fluxo contendo 5 atributos. E por último, o processo será feito com um fluxo de dados contendo 13 atributos, sendo esses dois últimos processos utilizando o método



de deep learning para redução do espaço de características na base de dados.

Neste protótipo, foram escolhidos três classificadores presentes na ferramenta Weka:

#### a) Árvores de Decisão

Uma árvore de decisão, ou árvore de classificação, é um sistema de suporte à decisão que utiliza um gráfico na forma de árvore para a tomada de decisões e seus possíveis efeitos posteriores. O algoritmo é usado para aprender uma função de classificação que decide o valor de um atributo dependente (uma variável), considerando os valores dos atributos independentes de entrada, conforme Bhargava et al. (2013).

A árvore de decisão J48 é a implementação do algoritmo ID3 (Iterative Dichotomiser 3) pelo WEKA. Maiores detalhes do algoritmo pode ser encontrado no trabalho de Quinlan (2014).

#### b) Naive Bayes

Naive Bayes é uma técnica probabilística para construção de classificadores baseado no teorema de Bayes, onde assume-se uma forte independência entre os atributos.

Classificadores Naive Bayes são escalonáveis e podem processar um grande número de variáveis lineares (parâmetros) em uma tarefa de aprendizagem. Em uma única iteração dos dados de treinamento, o algoritmo calcula a probabilidade de distribuição condicional de cada atributo de um determinado rótulo, seguido pela aplicação do teorema de Bayes para determinar a distribuição da probabilidade condicional do rótulo, usado para a previsão do resultado, de acordo com Aggarwal (2014).

Maiores informações da implementação do Naive Bayes pelo WEKA pode ser encontrado no trabalho de John e Langley (1995).

#### c) Tabelas de Decisão

Uma tabela de decisões é uma tabela que associa condições com ações a serem tomadas, apresentando um resultado após seguir uma série de decisões relacionadas. Ela permite modelar um conjunto complexo de regras com suas ações correspondentes.

No trabalho de Kohavi (1995), podem ser encontradas maiores informações da Tabela de Decisão implementada pelo WEKA.

Neste protótipo, o objetivo é a prova de conceito da caracterização dos fluxos de dados em tempo real. Assim, não será utilizado o processamento em lotes e o armazenamento das informações em um banco de dados, que permitem a realimentação de parâmetros para os algoritmos se adaptarem em tempo real. Os parâmetros calculados no processamento off-line com os dados históricos servem

para ajustar o modelo de processamento em tempo real, dando ao sistema uma característica adaptativa, pois os parâmetros podem ser atualizados, se ajustando a novos padrões de uso.

#### 4 | RESULTADOS NUMÉRICOS

O desempenho do sistema será avaliado através de duas métricas: a acurácia (percentual de acerto da classificação sobre a base de teste), e o número de falsos positivos e falsos negativos de cada classe.

A acurácia é a relação entre o número de amostras classificadas corretamente pelo número total de amostras. O número de falsos positivos indica quantas amostras normais foram classificadas como ataque e o número de falsos negativos indica quantas amostras de ataque foram classificadas como uma amostra normal.

Cada classificador será treinado e testado utilizando inicialmente os 41 atributos da base de dados KDD'99, sem o uso do deep learning para redução do espaço das características, servindo assim como a base de comparação. Depois, será repetido o treino e classificação utilizando 5 atributos e depois com 13 atributos. Os resultados serão, então, comparados pelas métricas estabelecidas.

Os resultados apresentam uma maior acurácia para a classificação de ataques em cada uma das unidades de processamento de fluxos criada quando se utilizam 13 atributos. O uso de apenas 5 atributos teve sua acurácia reduzida devido aos ataques do tipo R2L e U2L, pois o modelo SAE requer quantidades maiores de dados para realizar o aprendizado corretamente. Infelizmente, devido ao pequeno número de dados de treinamento disponíveis para esses ataques, os resultados obtidos não foram satisfatórios. Entretanto, no geral, mesmo utilizando 5 atributos, obteve-se melhores resultados comparados a não utilização de deep learning. A Tabela I apresenta a comparação entre os distintos algoritmos de classificação em termos de acurácia para os casos de 41, 5 e 13 atributos. Pode-se observar que com a redução dos atributos obtêm-se uma maior acurácia, em geral, na classificação de ataques.

<b>Algoritmo Classificador</b>	<b>41 Atrib.</b>	<b>5 Atrib.</b>	<b>13 Atrib.</b>
Árvores de Decisão	96,4980	96,7387	98,2162
Naive Bayes	90,2766	90,9497	96,6401
Tabelas de Decisão	95,8408	96,1994	98,0171

TABELA I

Comparação de acurácia (%) entre os algoritmos de classificação, utilizado 41, 5 e 13 atributos.

Os falsos positivos e falsos negativos são mostrados na Tabela II, a qual compara os diversos algoritmos utilizados pelas unidades de processamento de fluxo considerando 41, 5 e 13 atributos. As colunas da tabela mostram o número de ataques classificados corretamente (amostras classificadas como ataque, porém de

tipo diferente não são consideradas), o número de falsos positivos (conexão normal classificada como um ataque) e o número de falsos negativos (ataque classificado como conexão normal). Observa-se que, para um mesmo fluxo de dados, o algoritmo Naive Bayes apresenta um número muito maior de falsos-positivos, embora sua acurácia seja próxima dos outros algoritmos de classificação. Ainda, a classificação utilizando-se 5 atributos sofre novamente com a falta de dados para o treinamento correto dos ataques R2L e U2L, aumentando assim o número de falsos-positivos e falso-negativos comparado ao uso de 13 atributos.

<b>Classificador</b>	<b>Ataques Certos</b>	<b>FP</b>	<b>FN</b>
Árvores: 41 atrib.	3.768.413	684	1.489
Árvores: 5 atrib.	3.901.872	598	1.249
Árvores: 13 atrib.	4.429.091	323	881
Bayes: 41 atrib.	3.710.219	460.861	1.676
Bayes: 5 atrib.	3.881.231	442.893	1.393
Bayes: 13 atrib.	4.399.182	329.122	901
Tabelas: 41 atrib.	3.763.595	782	2.439
Tabelas: 5 atrib.	3.891.125	633	2.219
Tabelas: 13 atrib.	4.511.297	391	1.128

TABELA II

Resultados obtidos pelos algoritmos de classificação, em um total de 4.898.431 entradas, para 41, 5 e 13 atributos selecionados.

Para testar o desempenho do sistema proposto em tempo real, criou-se um ambiente virtual em uma máquina com processador Intel Core i7-4770S e 8 GB de memória RAM, executando o sistema operacional Linux Ubuntu 16. A medida de desempenho foi realizada pela média de fluxos processados por minuto pelas três unidades de processamento. O sistema foi capaz de processar aproximadamente 630 mil fluxos por minuto, com incerteza de 25 mil, dentro de um intervalo de confiança de 95%. Assim, cada ataque teve um tempo de detecção aproximado de 95 microssegundos.

## 5 | CONCLUSÕES

Este trabalho propõe um sistema de detecção de intrusão distribuído em tempo real, através do processamento de fluxos, utilizando deep learning para redução do espaço de características. O sistema apresentado utiliza ferramentas de código aberto que permite o processamento paralelo de diversos algoritmos de aprendizado de máquina, permitindo analisar um grande volume de dados em tempo real. A avaliação do sistema proposto é sobre a base de dados KDD'99, vastamente utilizada na comunidade, que foi transformada em um fluxo rotulado. Para a classificação dos dados, foram utilizados três unidades de processamento de fluxos, cada um com um algoritmo distinto de aprendizado de máquina - árvore de decisão, Naive Bayes e tabelas de decisão. Os resultados obtidos por cada algoritmo são, então, encaminhados para um visualizador.

Cada unidade de processamento de fluxos classificou 3 fluxos de dados diferentes. Inicialmente com 41 atributos (sem a utilização de deep learning, seguido por um fluxo de dados com 5 atributos e, por último, um fluxo de dados com 13 atributos. Utilizando-se deep learning para reduzir o espaço das características apresentou um melhor resultado em termos de acurácia e redução no número de falsos-positivos e falsos-negativos, embora ao utilizar-se apenas 5 atributos, notou-se uma falha na detecção de ameaças do tipo R2L e U2L devido a baixa representatividade desses ataques na base de dados, prejudicando o aprendizado. O sistema incorporando o método deep learning apresenta melhor acurácia em relação ao sistema sem redução de características, fornecendo uma solução para a detecção de ameaças em tempo real com maior desempenho.

Em trabalhos futuros, será proposta a comunicação entre as unidades de processamento de fluxos para troca de informações e a inclusão de diferentes fontes de fluxo de dados, sendo assim possível avisar os diferentes pontos da rede sobre um possível ataque à rede caso uma ameaça seja detectada por uma unidade. Será utilizado o deep learning não apenas para redução do espaço de características, mas também para a detecção de anomalias através da distribuição Gaussiana dos dados trafegados pela rede, permitindo a detecção de ataques desconhecidos.

## REFERÊNCIAS

AGGARWAL, C. C. **Data Classification: Algorithms and Applications**. 1st ed. Chapman & Hall/CRC, 2014.

ALOM, M. Z.; TAHA, T. M. **Network intrusion detection for cyber security using unsupervised deep learning approaches**. 2017 IEEE National Aerospace and Electronics Conference (NAECON), Junho 2017, pp. 63–69.

BENGIO, Y. et al. **Learning deep architectures for ai**. Foundations and trends in Machine Learning, vol. 2, no. 1, pp. 1–127, 2009.

BHARGAVA, N. et al. **Decision tree analysis on j48 algorithm for data mining**. Proceedings of International Journal of Advanced Research in Computer Science and Software Engineering, vol. 3, no. 6, 2013.

HARBI, N.; BAHRI, E. **Real detection intrusion using supervised and unsupervised learning**. 2013 International Conference on Soft Computing and Pattern Recognition (SoCPaR), Dez. 2013, pp. 321–326.

JOHN, G. H.; LANGLEY, P. **Estimating continuous distributions in bayesian classifiers**. Proceedings of the Eleventh conference on Uncertainty in artificial intelligence. Morgan Kaufmann Publishers Inc., 1995, pp. 338–345.

KIM, K.; AMINANTO, M. E. **Deep learning in intrusion detection perspective: Overview and further challenges**. 2017 International Workshop on Big Data and Information Security (IWBIS), Set. 2017, pp. 5–10.

KOHAVI, R. **The power of decision tables**. Proceedings of the 8<sup>th</sup> European Conference on Machine Learning, ser. ECML '95. London, UK, UK: Springer-Verlag, 1995, pp. 174–189. [Online]. Disponível

em: <http://dl.acm.org/citation.cfm?id=645324.649649>.

LEE, W.; STOLFO, S. J.; MOK, K. W. **Mining in a data-flow environment: Experience in network intrusion detection**. Proceedings of the fifth ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, 1999, pp. 114–124.

LEU, F. Y. et al. **An internal intrusion detection and protection system by using data mining and forensic techniques**. IEEE Systems Journal, vol. PP, no. 99, pp. 1–12, 2015.

LOPEZ, M. A. et al. **An evaluation of a virtual network function for real-time threat detection using stream processing**. 2018 Fourth International Conference on Mobile and Secure Services (MobiSecServ), Fev. 2018, pp. 1–5.

QUINLAN, J. R. **C4. 5: programs for machine learning**. Elsevier, 2014.

SCHUARTZ, F. C.; FONSECA, M. S. P.; MUNARETTO, A. **Sistema distribuído para detecção de ameaças em tempo real utilizando big data**. XXXV Simpósio Brasileiro de Telecomunicações e Processamento de Sinais - SBrT 2017, Set. 2017, pp. 472–476.

SHONE, N. et al. **A deep learning approach to network intrusion detection**. IEEE Transactions on Emerging Topics in Computational Intelligence, vol. 2, no. 1, pp. 41–50, Fev. 2018.

TAN, Z. et al. **Enhancing big data security with collaborative intrusion detection**. IEEE Cloud Computing, vol. 1, no. 3, pp. 27–33, Set. 2014.

VAN, N. T.; THINH, T. N.; SACH, L. T. **An anomaly-based network intrusion detection system using deep learning**. 2017 International Conference on System Science and Engineering (ICSSE), Julho 2017, pp. 210–214.

## UM MÓDULO DE DEFESA PARA ATAQUES DDOS NA CAMADA DE APLICAÇÃO USANDO ESTRATÉGIAS SELETIVAS

**Túlio Albuquerque Pascoal**  
**João Henrique Gonçalves Corrêa**  
**Vivek Nigam**

**Iguatemi Eduardo da Fonseca**  
Universidade Federal da Paraíba, Centro de  
Informática, João Pessoa-PB, Brasil

**RESUMO:** Este artigo propõe um módulo para defesa de ataques de negação de serviço na camada de aplicação. O módulo é executado em um servidor Web Apache e apresenta vantagens entre outros existentes na literatura, além de resultados similares a outra estratégia utilizada para o mesmo fim, mas que opera como um *proxy*. Nos experimentos realizados mostrou-se que o módulo proposto apresenta bons resultados em termos de disponibilidade, TTS (*time to service*), consumo de memória e CPU da máquina em que está sendo executado.  
**PALAVRAS-CHAVE:** Ataques de negação de serviço, Segurança na *Internet*, Redes de computadores.

**ABSTRACT:** This paper proposes a module that can be used as a defense against applications denial of service attacks. This module works in an Apache (*Web*) server and presents advantages when compared with others proposals in literature, as well as similar performance to another strategy that

runs as a *proxy*. The experimental results show that the proposed module presents appropriated values for availability, TTS (*time to service*), and memory and CPU consumption.  
**KEYWORDS:** Denial of service attacks, Internet security, Computer networks.

### 1 | INTRODUÇÃO

Ataques de Negação de Serviço (DoS – *Denial of Service*), são considerados uns dos mais perigosos e utilizados ataques contra redes e serviços (Gu and Liu 2007). Seu objetivo é causar indisponibilidade do serviço, website ou aplicação alvo para usuários honestos, consumindo todos os recursos disponibilizados de forma temporária ou indefinida. Seu poder é exponencialmente aumentado com a aplicação de Ataques de Negação de Serviço Distribuídos (DDoS – *Distributed DoS*) (Chang 2002). No DDoS, a fonte do ataque não é mais única, e sim dezenas, centenas e até milhares (Stephen and Lee 2004). Ataques DoS na camada de aplicação (ADoS – *Application Layer DoS*) são ainda mais difíceis de serem detectados, pois o alvo desses ataques são as vulnerabilidades de protocolos utilizados na camada de aplicação do modelo OSI (Xie and Yu 2009), como: HTTP, HTTPS, DNS, VoIP, FTP e SMTP. ADoS permitem aos atacantes a possibilidade



de focar seu ataque somente em uma aplicação ou serviço, deixando outros serviços disponíveis, e assim, dificultando a detecção do ataque (Xie and Yu 2009). Outro fator é que o tráfego gerado por ataques do tipo ADoS *LowRate* é similar ao de usuários honestos, dificultando sua detecção e mitigação (Dantas et al. 2014).

Uma defesa para mitigação de ataques ADoS, chamada SeVen (*Selective Verification in Application Layer*) foi desenvolvida com base em uma estratégia seletiva. O SeVen é uma estratégia nova na literatura, datada de 2014, que possui resultados bastante satisfatórios, mantendo servidores *Web*, quando sob ataque, com cerca de 95% de disponibilidade (Dantas et al. 2014). SeVen funciona como um *proxy*, fazendo a *interface* entre as requisições dos clientes com o servidor, aumentando a complexidade da ferramenta, pois deve se preocupar, além da execução da estratégia em si, com outros aspectos como: (1) qual tipo e versão de servidor estão sendo utilizados e suas peculiaridades; (2) qual(is) o(s) protocolo(s) esta(ão) sendo utilizado(s); (3) configuração (saber e replicar a atual configuração do servidor para a estratégia funcionar de acordo); (4) segurança (integridade, confidencialidade, disponibilidade e autenticidade da ferramenta em si); (5) robustez (garantir funcionamento ininterrupto em qualquer situação, para não prejudicar o servidor que está sendo protegido). Devido a esses fatores, faz-se necessário o uso de uma abordagem menos dependente e que possa ao mesmo tempo ser eficaz e fácil para o uso de administradores de redes. Para atingir esse objetivo pode-se aplicar o conceito de módulo, que funciona como um *mini-software* diretamente acoplado em um servidor (MódulosApache 2016). O Apache HTTP Server Project, mais conhecido por servidor Apache, é o servidor *Web* mais popular e utilizado atualmente. De acordo com pesquisas da W3Tech e a BuiltWith, servidores Apache são usados por 55,9% e 51% entre todos os sites na *Internet* em Dezembro de 2015 (W3tech 2016) (BuiltWith 2015). Por ser *OpenSource*, servidores Apache fornecem a liberdade e extensibilidade de seu funcionamento, a partir da implementação de módulos (MódulosApache 2016). Os objetivos e contribuições deste trabalho são, portanto, desenvolver um módulo Apache, chamado *mod\_seven*, para mitigar ataques ADoS. Os resultados dos experimentos mostraram que o *mod\_seven* obteve resultados equivalentes e alinhados com os obtidos na versão *proxy* do SeVen, em alguns casos com resultados até melhores. Além de replicar os experimentos realizados na versão *proxy*, também foram realizados testes baseados em situações reais de ataques, ou seja, ataque com durações maiores, além de ataques no protocolo HTTPS (não suportado pela versão *proxy*). Também comparamos o módulo proposto com outros encontrados na literatura, relatando suas vantagens sobre os mesmos. Outros módulos Apache existentes no mercado propõem-se a mitigar ataques ADoS (Monshouwer 2013) (Morimoto 2013) (Reqtimeout 2014). O *mod\_antiloris* tem como estratégia a contagem de conexões simultâneas abertas de um mesmo endereço IP. Quando essa contagem superar o valor configurado (o padrão é 10), o módulo rejeita todas as requisições provenientes daquele IP. O *mod\_pacify\_loris* (Morimoto 2013)



possui a mesma estratégia de defesa (mas utiliza 50 conexões por padrão), porém ainda implementa mais duas análises: contagem de GET HEADER enviados em uma mesma conexão e a taxa de requisições GET HEADER enviadas por segundo. Já o *mod\_reqtimeout* (Reqtimeout 2014), o mais atual e utilizado dentre eles, baseia-se em uma análise mais avançada das taxas de envio dos cabeçalhos e corpo das requisições. O módulo possui uma diretiva chamada *RequestReadTimeout* em que há dois parâmetros configuráveis para cada um dos campos (cabeçalho e corpo) de uma requisição, que são *timeout* e *minrate*. O seu diferencial quanto aos outros módulos é que ele avalia tais parâmetros em conjunto e a cada vez que pacotes de dados de uma requisição chegam ao servidor, o módulo renova suas janelas de *timeout*.

A Seção 2 apresenta a natureza, característica e funcionamento de ataques ADDoS. Na Seção 3 é apresentada a adaptação da estratégia SeVen como um módulo, bem como a arquitetura, funcionamento e interligação do módulo com o servidor. A Seção 4 descreve os experimentos realizados, discute e compara os resultados obtidos. Finalmente, na Seção 5 são apresentadas as conclusões e trabalhos futuros.

## 2 | ATAQUES DE NEGAÇÃO DE SERVIÇO NA CAMADA DE APLICAÇÃO

Dentre os ataques ADDoS destaca-se: o *Flooding* e o *LowRate*. No primeiro, tem-se a geração de um enorme fluxo de tráfego, para consumir todos os recursos da aplicação, até que ela fique incapaz de atender novos clientes. O segundo, consiste na geração de tráfego similar ao de clientes honestos, porém, utilizando-se de vulnerabilidades encontradas nos protocolos para manter requisições em atendimento por tempo indeterminado (Durcekova et al. 2012), assim não necessitando de grandes recursos para geração dos ataques (Dantas 2015). Ferramentas de geração desse tipo de ataque são facilmente encontradas na Internet e simples de usar (Loic 2013) (Slowloris 2013) (Rudy 2013) (Slowhttptest 2013). Por esses motivos, ataques *LowRate* são mais traiçoeiros e perigosos, dificultando detecção e proteção por parte da defesa. A seguir tem-se a descrição e funcionamento dos dois ataques do tipo *LowRate* mais utilizados:

- **Slowloris:** Consiste no envio de requisições HTTP GET HEADER incompletas (sem os campos *CR-Carriage Return, ASCII 13, /r,* e o *LF - Line Feed, ASCII 10, /n*) no final do pacote (Owasp 2009), em um certo intervalo de tempo. Fazendo com que o servidor nunca saiba quando requisições foram finalizadas, mantendo-as no *pool* de atendimento até o valor de *timeout* pré-configurado no servidor, a partir de um momento, todos os recursos estarão sendo consumidos pelas requisições maliciosas, indisponibilizando a aplicação;
- **HTTP POST:** Este ataque envia requisições HTTP GET HEADER com-

pletas, realizando o *Hand-shake* com o servidor, porém seus dados do campo BODY, enviados pelo método HTTP POST são enviados de maneira muito lenta, fazendo com que o servidor aguarde até o *timeout* configurado ou a quantidade máxima de dados no BODY de uma requisição seja atingida. Assim, essas requisições lentas tomam posse de todo o *pool* de atendimento e indisponibilizam a aplicação.

### 3 | O MÓDULO APACHE: *MOD\_SEVEN*

Em (Kew 2007) é descrito o *Apache HTTP Server* composto por um pequeno *core* e um conjunto de módulos. Dentre todos os módulos, existe um módulo “especial”, chamado MPM (Módulo de Processamento Múltiplo), que serve como um otimizador para a comunicação entre o sistema operacional e a APR (*Apache Portable Runtime*), conjunto de bibliotecas de suporte para o servidor Apache que fornece um conjunto de APIs para comunicação com o sistema operacional). Módulos são a chave para a extensibilidade provida pelo Apache, com eles tem-se a liberdade de customizar e criar novos processos. A execução de módulos é baseada em ganchos, que funcionam como métodos que trabalham diretamente com o *core* do Apache para melhor desempenho (MódulosApache 2016). Há vários tipos de ganchos já definidos no Apache, mas novos ganchos também podem ser criados por programadores (Kew 2007). O *mod\_seven* possui três ganchos: (1) o *ap\_hook\_post\_config* para recolher informações do servidor, como tamanho do *pool* de atendimento, criação e alocação de *threads*; (2) o *ap\_hook\_create\_request* para detecção de *dummy requests* (Hayden 2008), um *bug* encontrado no Apache durante a implementação desse trabalho, e registro de requisições no filtro utilizado pelo módulo; (3) o *mod\_seven\_input\_filter* que é onde ocorre praticamente toda a aplicação da estratégia. Quando uma requisição é registrada em um filtro, todo e qualquer dado enviado por aquela requisição será analisado e processado pelo filtro. Para simplificação e melhor entendimento, o funcionamento do *mod\_seven\_input\_filter* do módulo foi dividido em 4 fases distintas explicadas abaixo:

- **Fase de Reconhecimento:** Nessa fase o filtro extrai informações da requisição (endereço IP, porta e *socket* da conexão), e aloca uma variável-estrutura interna da APR do tipo *worker\_score* para representar uma requisição no *pool* de atendimento do Apache, chamado *scoreboard*;
- **Fase de Detecção:** Aqui acontece uma adaptação necessária necessária à estratégia para adequar-se ao comportamento e Módulos Apache. Como o Apache não permite a extração e manuseamento direto de requisições que se encontram no *scoreboard*, essa fase realiza uma varredura no *scoreboard* atual da aplicação verificando se a requisição atual (que está sendo processada pelo filtro) está com *flag* que foi selecionada para ser removida pela estratégia, caso positivo, a conexão daquela requisição será fechada imediatamente; caso negativo, o fluxo do módulo continua para a próxima fase, a Fase de Adição;

- **Fase de Adição:** É uma fase simples e direta, após escolher as informações da requisição e verificar que a mesma não encontra-se selecionada para deleção, o módulo conclui que ela está apta a ser atendida e processada, o *worker score* da mesma é adicionado ao *scoreboard* de acordo com a *thread* e número de processo alocados pelo servidor para atender aquela requisição;
- **Fase de Análise e Decisão:** É a fase mais completa e que aplica toda a lógica da estratégia. Uma contagem de requisições no *pool* de atendimento é realizada, similar ao da Fase de Detecção, para concluir se o servidor encontra-se sobrecarregado ou não, utilizando uma variável para comparar com o valor do parâmetro *server limit*, que representa o máximo de conexões simultâneas que o servidor pode atender. Neste ponto, uma função de probabilidade *FP1* decide a aceitação ou não da requisição. Caso a requisição seja rejeitada pela estratégia, sua conexão é automaticamente fechada usando *APR\_DECLINED*, *ap\_conn\_close* e *apr\_socket\_close*. Caso não, uma requisição é escolhida de acordo com uma função de distribuição uniforme *FP2*, para ser substituída do *pool* de atendimento (origem da característica seletiva da estratégia). Após a escolha, o módulo resgata informação do *worker\_score* escolhido, e adiciona a *flag* para deleção, assim, quando qualquer dado referente àquela requisição chegar ao servidor, a mesma será rejeitada automaticamente na Fase de Detecção.

É importante salientar que enquanto a aplicação não se encontra sobrecarregada, o módulo simplesmente executa as Fases de Reconhecimento e Adição. Na Figura 1 tem-se o fluxo de funcionamento e processos do *mod\_seven* realizados pelos seus ganchos e métodos internos. Na Figura 2 tem-se uma visão geral da divisão e execução de processos ocorridos nas distintas fases do *mod\_seven\_input\_filter*.

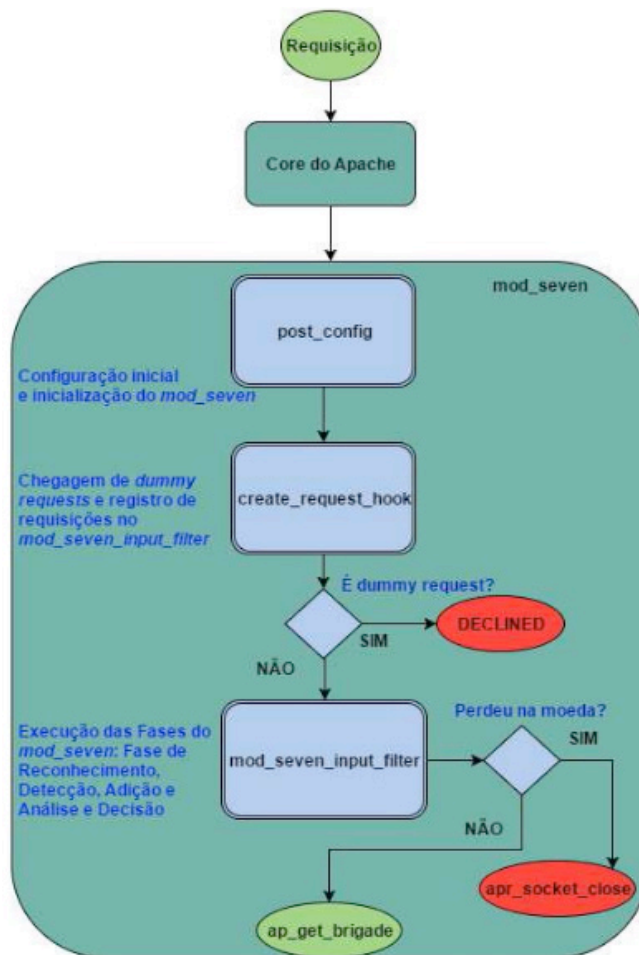


Figura 1. Organização e ordem de fluxo de execução do mod\_seven

## 4 | EXPERIMENTOS E RESULTADOS

### 4.1 Cenário e Configuração dos Experimentos

Os testes foram realizados usando três máquinas em dois diferentes *Campus* da Universidade Federal da Paraíba (UFPB). Duas máquinas situadas no *Campus V*, gerando tráfego de clientes legítimos e de atacantes (cada máquina gerando um tipo de tráfego). E uma outra máquina no *Campus I* hospedando a página Web padrão do Apache, utilizando o MPM PREFORK, que é o MPM padrão para sistemas operacionais Unix (ApacheMPM 2014). As máquinas para geração de tráfego possuem processador Intel i5-3470 de 3.20Ghz e 4GB de RAM e o servidor um Intel Xeon E5-2620 de 2.00GHz e 8GB de RAM. O objetivo dessa configuração é de simular, com um maior grau de realidade, um ataque DoS, em que o tráfego situa-se em redes distintas e separadas fisicamente. Para geração do tráfego cliente utilizamos a ferramenta Siege (ferramenta de benchmark para medir desempenho de aplicações web (Siege 2015)) e para o tráfego atacante usamos duas ferramentas: *Slowloris* (Slowloris 2013) e *Slowhttptest* (Slowhttptest 2013). À quesito de comparação de resultados, a configuração e cenários dos experimentos realizados nesse trabalho foram replicados de acordo com os realizados em (Dantas et al. 2014).

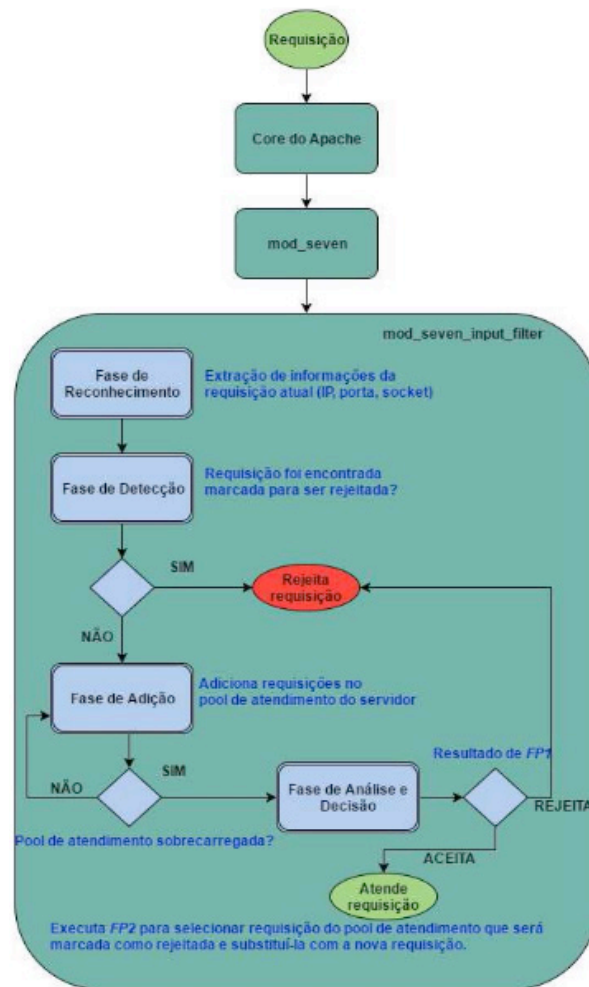


Figura 2. Ordem de fluxo e fases do mod\_seven\_input\_filter

As configurações dos testes levam ao cenário ideal para que o ataque obtenha sucesso e possa indisponibilizar a página Web alvo. O servidor foi configurado da seguinte maneira:

- **Timeout:** Tempo, em segundos, que o servidor irá aguardar para continuar a receber dados de requisições em uma mesma conexão. Configurado como 40 segundos;
- **MaxRequestWorkers:** É o número de requisições simultâneas que serão atendidas pelo servidor. Ou seja, o limite de atendimento da aplicação Web. Configurado com 200;

Os experimentos realizados no *mod\_seven*, nos outros módulos e com o Apache puro (sem nenhum módulo) tiveram as seguintes especificações

- **Quantidade de Atacantes:** 250 atacantes para cada tipo de ataque: *Slowloris* e *HTTP POST*;
- **Tráfego Atacante:** Conexões enviando requisições a cada 35 segundos. Aproximadamente 7,14 requisições por segundo;

- **Quantidade de Clientes Honestos:** 100 clientes gerados pelo *Siege*. Aproximadamente 10 requisições por segundo;
- **Tráfego de Clientes Honestos:** Requisições de cada cliente são enviadas em um intervalo de 0 a 3 segundos a fim de melhor simular um tráfego Web legítimo;
- **Protocolos:** HTTP (porta 80) e HTTPS (porta 443);
- **Duração:** Três repetições para os testes de 5 minutos e uma para os testes de 2 horas;
- **Tipo de Testes:** sem *mod\_seven* e sem ataque; sem *mod\_seven* e com ataque; com *mod\_seven* e sem ataque; com {*mod\_seven*, *mod\_antiloris*, *mod\_pacify\_loris*, *mod\_reqtimeout*} e com ataque.

Como métricas de desempenho, utilizou-se os seguintes parâmetros: (i) **Disponibilidade:** Porcentagem dos clientes atendidos com sucesso; (ii) **TTS:** Tempo médio de resposta para cada requisição; (iii) **Consumo de memória:** Porcentagem do consumo médio de memória durante o teste; (iv) **Consumo de CPU:** Porcentagem do consumo médio de CPU durante o teste.

## 4.2 Resultados e Discussão

	Sem <i>mod_seven</i>				Com <i>mod_seven</i>			
	Disponibilidade	TTS	Memória	CPU	Disponibilidade	TTS	Memória	CPU
<i>Sem Ataque</i>	100%	0,01s	0,9%	7,8%	100%	0,03s	0,9%	8,7%
<i>Slowloris</i>	0,0%	∞	17,5%	0,0%	98,7%	0,07s	18,2%	2,6%
<i>HTTP POST</i>	0,0%	∞	16,6%	0,0%	96,6%	0,03s	13,5%	1,8%

TABELA I: Disponibilidade, TTS, consumo de memória e CPU dos testes realizados

Pela Tabela I percebe-se que o *mod\_seven* não influencia na disponibilidade da aplicação em situações normais (sem ataque), percebe-se somente um aumento insignificante do consumo de CPU (de 7,8% para 8,7%) e de 0,02 segundos no TTS, devido aos processamentos adicionais realizados pelo módulo.

Analisando os cenários com ataques, percebe-se a eficiência do *mod seven*, que manteve a aplicação com uma disponibilidade de 98,7% e 96,6% nos ataques *Slowloris* e *HTTP POST*, respectivamente, e com valores de TTS baixos.

Pela Figura 4, nota-se que o consumo de memória, quando utilizado o módulo não é elevado e no caso do ataque *HTTP POST* o consumo foi reduzido. Isso é explicado devido à natureza do ataque, que não envia tantas requisições como o *Slowloris*. Outro fator interessante é o consumo de CPU ser nulo (teve pico de 0,3%) quando sob ataque, pois, com a aplicação indisponibilizada, a mesma não processa mais nenhuma requisição, conseqüentemente não consumindo mais recursos da CPU.



Porém, ainda há de memória, uma vez que as requisições continuam chegando e sendo armazenadas, esperando atendimento (ver Figura 3). No cenário com *mod\_seven* (ver Figura 4) percebe-se o consumo de recursos intercalados, mostrando que o servidor encontra-se ativo e em funcionamento.

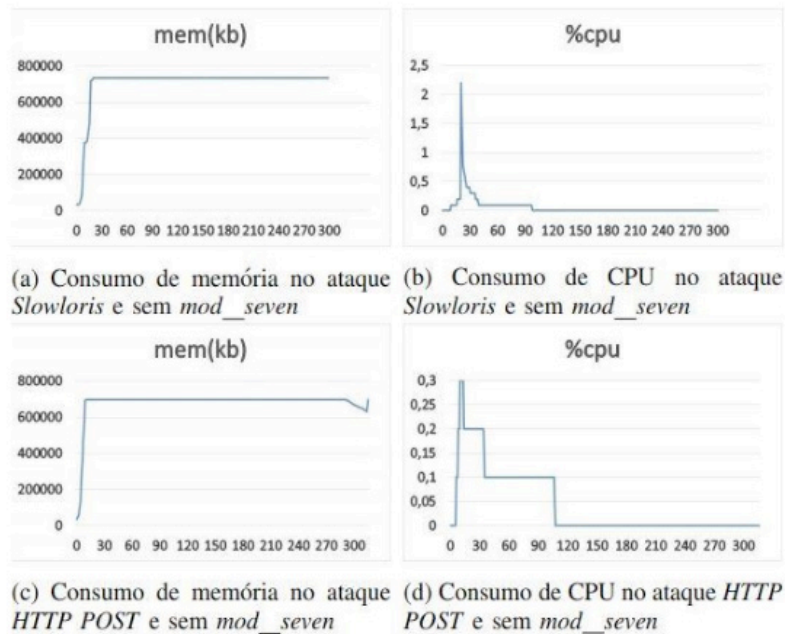
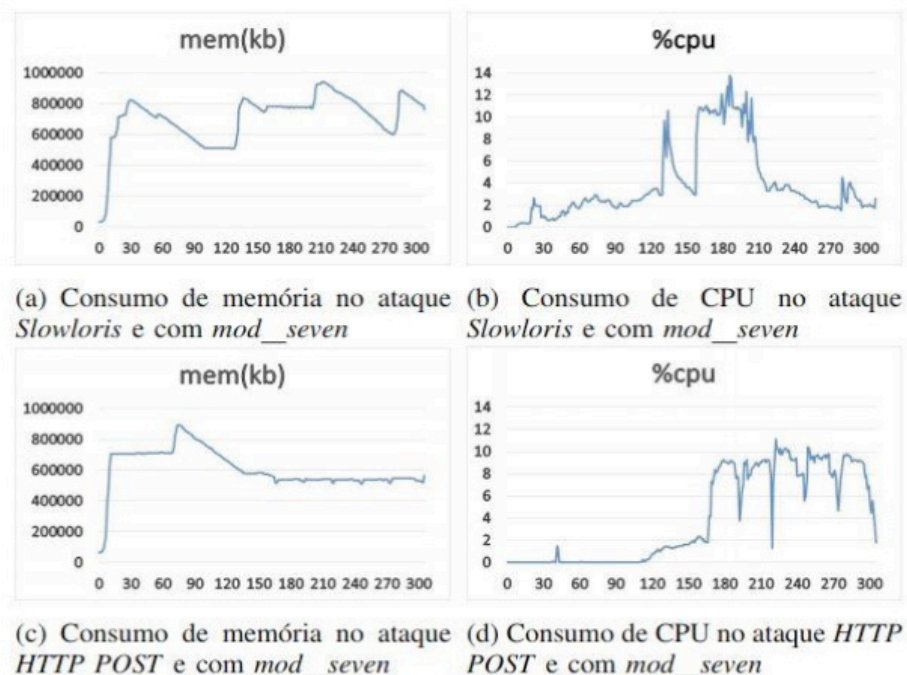


Figura 3. Consumo de Memória e CPU com ataque e sem *mod\_seven*

Nos testes de 2 horas de duração (protocolos HTTP e HTTPS), confirmou-se o bom desempenho do *mod\_seven*. Verifica-se uma pequena queda na disponibilidade para o ataque HTTP POST (ver Tabela II), uma vez que a taxa de requisições por segundo aumentou, transformando o ataque em um ataque *mini-flooding*, dando indícios que o *mod\_seven* possa obter bons resultados contra ataques do tipo *Flooding*.





No protocolo HTTPS, verificou-se que o ataque *slowloris.pl* mostrou-se incapaz de realizar o ataque nesse protocolo. Devido a esse fato, utilizamos a ferramenta *Slowhttptest*. Houve uma redução da disponibilidade e um pequeno aumento no TTS, sobretudo porque o protocolo HTTPS aplica mais métodos de segurança e criptografia (contudo, uma maior e mais específica análise deve ser realizada em relação ao *overhead* causado pelo HTTPS, em um trabalho futuro). Outra importante conclusão é da robustez e bom gerenciamento de *threads* e processos do módulo, uma vez que suportou uma carga elevada (recebeu cerca de 79.945 pacotes atacantes e 422.511 requisições clientes nos testes de 2 horas).

Ataque	Protocolo	Disponibilidade	TTS
<i>Slowloris</i>	HTTP	97,5%	0,14s
<i>Slowloris</i>	HTTPS	91,6%	0,09s
<i>HTTP POST</i>	HTTP	91,2%	0,05s
<i>HTTP POST</i>	HTTP	92,4%	0,07s

TABELA II: Resultados dos experimentos de 2 horas do mod\_seven

Em relação ao *mod\_antiloris* e *mod\_pacify\_loris*, o *mod\_seven* se mostra uma defesa agnóstica, pois não discrimina as requisições recebidas, ou seja, todas possuem as mesmas chances de serem atendidas. Esses módulos utilizam uma estratégia discriminatória, trazendo algumas desvantagens. Além disso, esses módulos podem prejudicar usuários honestos que estejam na mesma família de endereçamento IP ou em uma mesma rede pública que um atacante (o que ocorre em redes de grandes organizações, onde um único IP público pode representar mais de uma máquina na sua rede interna). Uma vez que a defesa bloqueia requisições pela informação do IP público, usuários honestos podem ser prejudicados caso um atacante esteja na sua mesma sub-rede. Pela tática de taxa de contagem de cabeçalhos por segundo implementada pelo *mod\_pacify\_loris*, ele pode erroneamente rejeitar requisições de clientes honestos que possuem conexões lenta.

Por sua vez, o *mod\_reqtimeout*, apesar de possuir uma estratégia mais inteligente, baseada em *timeouts* e *minrate* renováveis, possui vulnerabilidade. Uma vez que atacantes podem utilizar estratégias de detecção desses valores a partir de experimentos prévios ao ataque. Por exemplo, enviando requisições em diferentes intervalos de tempo, e verificando o comportamento e as respostas do servidor, é possível encontrar um valor aproximado do tempo que suas conexões mantêm-se em atendimento até que comecem a ser rejeitadas. Esse valor de tempo, é justamente o valor configurado para a diretiva *timeout* configurado na defesa. Com essa informação, realiza-se o ataque utilizando o *timeout* de suas conexões com um valor próximo ao

detectado, assim, renovando suas conexões antes de serem removidas, mantendo-as indefinidamente em atendimento.

Comparando com o *SeVen proxy*, os resultados do módulo mostram uma pequena melhora da disponibilidade da aplicação (aumento médio de 1%) nos experimentos com os mesmos cenários usados em (Dantas et al. 2014), além de possibilitar a aplicação da estratégia no protocolo HTTPS e em qualquer outro, desde que suportado pelo Apache. Outra vantagem é que a estratégia pode ser melhor difundida dado a grande utilização de servidores Apache pela comunidade, além de sua facilidade de utilização.

## 5 | CONCLUSÃO E TRABALHOS FUTUROS

Esse artigo apresentou uma solução para uns dos maiores problemas na Internet na atualidade, ataques DDoS. A defesa proposta é construída como um Módulo Apache, que é o servidor Web mais utilizado nos dias atuais. O *mod\_seven* obteve resultados consistentes em vários cenários de testes e também quando comparado com o *SeVen proxy* e com outros módulos Apache existentes na literatura. Além disso, o *mod seven* trouxe melhorias e vantagens em relação ao *SeVen proxy*, conseguindo: mitigar ataques DoS no protocolo HTTPS; melhora de performance; modesto consumo de memória e CPU; utilização e implantação mais fácil e maior robustez da defesa. Como trabalho futuro, objetiva-se testar o *mod seven* na mitigação de ataques DoS em outros protocolos suportados pelo Apache; testá-lo contra um ataque do tipo *LowRate* mais recente, chamado *SlowRead* (Park et al. 2015), bem como contra ataques do tipo *Flooding*.

## 6 | AGRADECIMENTOS

A CAPES, CNPq e RNP pelo apoio no desenvolvimento deste trabalho.

## REFERÊNCIAS

[ApacheMPM 2014] ApacheMPM (2014). **Multi-Processing Modules (MPMs)**. <http://docs.apache.org/docs/2.2/en/mpm.html>. Acessado em: 18 de Agosto de 2015.

[BuiltWith 2015] BuiltWith (2015). **Web server usage statistics – Statistics for websites using web server technologies**. <http://trends.builtwith.com/web-server>. Acessado em: 18 de Dezembro de 2015.

[Chang 2002] Chang, RKC. (2002). **Defending against flooding-based distributed denial-of-service attacks: a tutorial**. *Communication Magazine, IEEE*, 40(10), 42-51.

[Dantas et al. 2014] Dantas, Y. G., Nigam, V., and Fonseca, I. E. (2014). **A selective defense for application layer ddos attacks**. In *Intelligence and Security Informatics Conference (JISIC)*, 2014 IEEE Joint, pages 75–82. IEEE.

[Dantas 2015] Dantas, Y. G. (2015). **Estratégias para tratamento de ataques de negação de serviço na camada de aplicação em redes ip**. Master Thesis in Portuguese.

- [Durcekova et al. 2012] Durcekova, V., Schwartz, L., and Shahmehri, N. (2012). **Sophisticated denial of service attacks aimed at application layer**. In ELEKTRO, 2012, pages 55–60. IEEE.
- [Gu and Liu 2007] Gu, Q. and Liu, P. (2007). **Denial of service attacks**. Handbook of Computer Networks: Distributed Networks, Network Planning, Control, Management, and New Trends and Applications, 3:454–468.
- [Hayden 2008] Hayden, M. (2008). **Apache 2.2: internal dummy connection**. <https://major.io/2008/09/23/apache-22-internal-dummy-connection/>. Acessado em: 25 de Julho de 2015.
- [Loic 2013] Loic (2013). **A network stress testing application**. <https://github.com/NewEraCracker/LOIC/downloads>. Acessado em: 28 de Janeiro de 2015.
- [Kew 2007] Kew, N. (2007). **The Apache modules book: application development with Apache**. Prentice Hall Professional.
- [Monshouwer 2013] Monshouwer, K. (2013). **mod antiloris**. <https://sourceforge.net/projects/mod-antiloris/>. Acessado em: 10 de Agosto de 2015.
- [Morimoto 2013] Morimoto, S. (2013). **mod pacify loris**. [http://mod-pacify-slowloris.googlecode.com/svn/trunk/mod\\_pacify\\_loris.c](http://mod-pacify-slowloris.googlecode.com/svn/trunk/mod_pacify_loris.c). Acessado em: 10 de Agosto de 2015.
- [MódulosApache 2016] MódulosApache (2016). **Developing modules for the Apache HTTP Server 2.4**. <http://httpd.apache.org/docs/2.4/developer/modguide.html>. Acessado em: 07 de Outubro de 2016.
- [Owasp 2009] Owasp (2009). **CRLF Injection**. <https://www.owasp.org/index.php/CRLFInjection>. Acessado em: 27 de Março de 2015.
- [Park et al. 2015] Park, J., Iwai, K., Tanaka, H., and Kurokawa, T. (2015). **Analysis of slow read dos attack and countermeasures on web servers**. International Journal of Cyber-Security and Digital Forensics (IJCSDF), 4(2):339–353.
- [Reqtimeout 2014] Reqtimeout (2014). **mod reqtimeout**. [https://httpd.apache.org/docs/2.4/mod/mod\\_reqtimeout.html](https://httpd.apache.org/docs/2.4/mod/mod_reqtimeout.html). Acessado em: 11 de Agosto de 2015.
- [Rudy 2013] Rudy (2013). **R.U.D.Y – Are you dead yet**. <https://code.google.com/p/r-u-dead-yet/>. Acessado em: 07 de Fevereiro de 2015.
- [Slowhttpptest 2013] Slowhttpptest (2013). **Slowhttpptest tool**. <https://code.google.com/p/slowhttpptest/>. Acessado em: 02 de Fevereiro de 2015.
- [Slowloris 2013] Slowloris (2013). **Slowloris tool**. <http://ha.ckers.org/slowloris/>. Acessado em: 02 de Fevereiro de 2015.
- [Siege 2015] Siege (2015). **Linux man page: siege - An HTTP/HTTPS stress tester**. <http://linux.die.net/man/1/siege>. Acessado em: 18 de Dezembro de 2015.
- [Stephen and Lee 2004] Stephen, S. and Lee, Ruby B. (2004). **Distributed Denial of Service: Taxonomies of Attacks, Tools, and Countermeasures**. ISCA, PDCS, 42-51.
- [Xie and Yu 2009] Xie, Y. and Yu, S.-Z. (2009). **Monitoring the application-layer ddos attacks for popular websites**. Networking, IEEE/AcM Transactions on, 17(1):15–25.
- [W3tech 2016] W3tech (2016). **Usage of web servers for website**. [http://w3techs.com/technologies/overview/web\\_server/all](http://w3techs.com/technologies/overview/web_server/all). Acessado em: 18 de Agosto de 2016.

## AN EMPIRICAL RATE BALANCED ALIEN XTALK MITIGATION METHOD FOR G.FAST SYSTEMS

### Diego de Azevedo Gomes

Universidade Federal do Sul e Sudeste do Pará,  
Faculdade de Computação e Engenharia Elétrica,  
Marabá – Pará

### Cláudio de Castro Coutinho Filho

Universidade Federal do Sul e Sudeste do Pará,  
Faculdade de Computação e Engenharia Elétrica,  
Marabá – Pará

### João Victor Costa Carmona

Universidade Federal do Sul e Sudeste do Pará,  
Faculdade de Computação e Engenharia Elétrica,  
Marabá – Pará

### Evaldo Gonçalves Pelaes

Universidade Federal do Pará, Instituto de  
Tecnologia, Belém – Pará

**ABSTRACT:** The G.fast is the most recently DSL standard released by the ITU. The G.fast deployment environment is very noisy, in which several external interferences may impair the transmission. One of the most problematic interference source is the alien xtalk. Despite the difficult to mitigate this type of interference, there are some mitigation methods that reduce the impact of the alien xtalk over DSL systems. However, these methods present a setback that is to unbalance the transmission rates achieved by the coordinated lines. In this paper we propose an empirical method to achieve rate balancing simultaneously with

alien xtalk mitigation. Additionally, we present a set of simulation results performed with cable measurements, that show the effectiveness of our method.

**KEYWORDS:** G.fast, Alien Crosstalk, Rate Balancing.

### 1 | INTRODUÇÃO

The demand for higher data rates has taken the transmission rates of the access technologies to the order of the Gbps, such as the G.fast (ITU-T G.9700, 2014). This technology was standardized by the International Telecommunication Union, which aims to achieve 1 Gbps over short copper loops, with a bandwidth of 106 or 212 MHz. Due to this large bandwidth, the G.fast is susceptible to a variety of interferences, both from the coordinated pairs (in-domain crosstalk) as from external sources, in which alien crosstalk (AXT) is one of the most dangerous. The issue of combating the in-domain crosstalk is addressed by the utilization of Vectoring techniques, i.e., a set of techniques capable of effectively canceling crosstalk, in order to provide better service quality and data transmission (R. ZIDANE et.al, 2013).

The term “Vectoring” is used to describe the definitions of the ITU G.Vector standard (ITU-T Rec. G.993.5, 2010). The Vectoring

techniques have started to be implemented in the commercial networks in some parts of the globe and it will possibly be in focus in the next years.

In the Vectoring field, the Precoder Matrix is one of the most important structures, due to the possibility of surpassing the Far-End-Crosstalk (FEXT), i.e., the crosstalk caused by each line on every other at the far end of the cable. However, as the frequency increases (G.fast works with frequencies up to 200 MHz), the Linear Precoder, or Diagonalizing Precoder (DP), becomes unstable, with the crosstalk channel being even higher than the direct channel (BRITISH TELECOM, 2013). Given the fact that the crosstalk increases with frequency, a simple normalization method is presented in (R. CENDRILLON et.al, 2007). The method, however, causes decrease in the system's performance, therefore impacting the final user rates. Similarly, a heuristic approach was proposed in (ALCATEL-LUCENT, 2013). The presented method was able of reducing the normalization penalty among lines. The so-called rate balancing techniques can also be found in other scenarios (C. HELLINGS e W. UTSCHICK, 2013; A. TARIGHATI e J. JALDEN, 2016; N. LI et. al, 2014) given the impact of this situation.

A similar behavior is observed in the AXT mitigation methods, in which the rates of some lines achieve greater rates than others (GEORGE GINIS e CHIA-NING PENG, 2006; A. AWASTHI et. al, 2012; AHMAD et. al, 2014). In (GEORGE GINIS e CHIA-NING PENG, 2006) and (A. AWASTHI et. al, 2012), the lines are benefited distinctly due to the decision feedback structure, which provides more prediction taps for the last user, thus increasing the AXT mitigation, and, consequently, providing more rates for these lines. Conversely, in (AHMAD et. al, 2014) the first lines appreciate the higher rates, because in this method the gains of each line are the diagonal elements of the upper-triangular matrix of a QR decomposition (GILBERT STRANG, 1986).

For that matter, this work introduces the rate-balancing methods, or techniques, to reduce the rate unbalancing in the AXT mitigation methods. As their name already attest, these techniques try to balance the bitrates among users (or lines) in the same binder at their Customer Premises Equipment (CPEs), by altering the order in which they are processed at the Central Office (CO). This sorting is able to achieve a good approximation among lines. In this context, an empirical low computational cost method is proposed to define a line order, that minimizes the total bitrate variance of all lines and maximizing the total bitrate sum. This can be a convenient tool for a proper comparison between each method and the optimal achieved results for testing their performance.

This chapter is organized as follows. Section II describes the system model adopted in this document. The rate balancing method is presented in Section III. Section IV shows some simulation results of the application of the proposed method. Finally, Section V concludes the chapter.

## 2 | 2. AXT SYSTEM MODEL

In the current DSL infrastructure, it is possible to find groups of pairs, which are coordinated in at least one of the ends of the cables. This arrangement allows to employ MIMO techniques to combat in-domain crosstalk (crosstalk among the pairs of the same coordinated group), such as the standardized vectoring techniques described in (ITU-T Rec. G.993.5, 2010). However, this vectored group can share the same environment with pairs of another company. In this scenario, the vectored lines (VLs) are susceptible to suffer the impairment called alien crosstalk (or out-of-domain crosstalk), which is the crosstalk caused by lines that are not in the vectored group. Consequently, the total interference observed in the VLs is composed by the AXT and by the background noise (here we do not consider the interference caused by other types of sources, such as RFI, impulsive noise etc.).

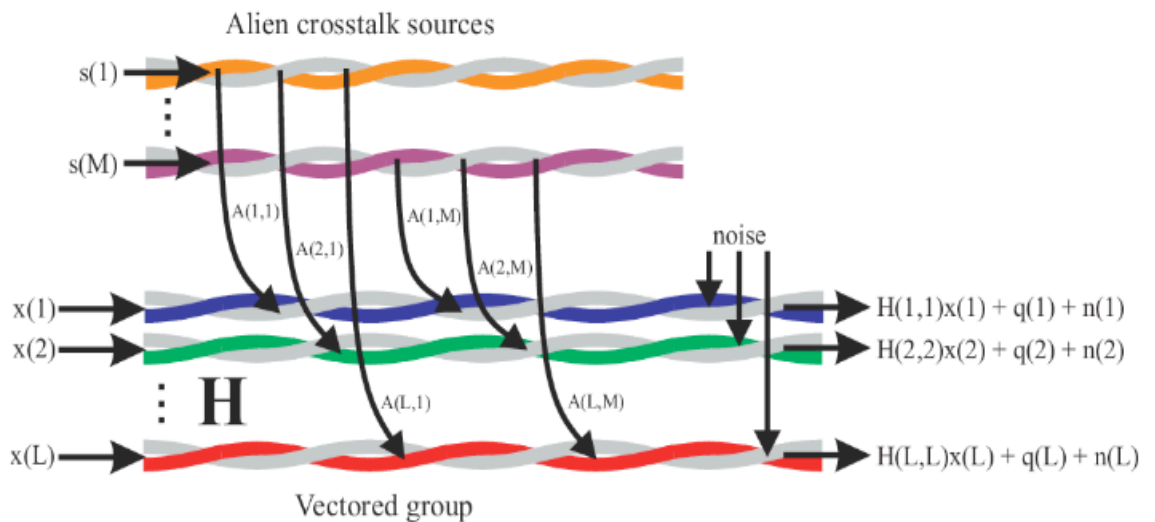


Figure 1: Complete system with vectored lines and AXT.

Figure 1 depicts this scenario, in which a group of  $L$  vectored lines is impaired by the crosstalk caused by  $M$  alien lines (ALs). Assuming that both VLs and ALs use DMT modulation, and that all lines are synchronized (there is no Inter-Carrier Interference - ICI), the received symbols,  $y_k = [y_{1,k}, \dots, y_{L,k}]^T$ , in each independent  $k$  tone can be modelled by:

$$y_k = H_k x_k + z_k, k = 1, \dots, N, \quad (1)$$

where  $N$  is the DFT length, i.e., the number of tones,  $H_k$  is a complex  $L \times L$  matrix composed by the direct and in-domain crosstalk channel frequency values at tone  $k$ ,  $x_k = [x_{1,k}, \dots, x_{L,k}]^T$  is the  $L \times 1$  vector containing the QAM encoded transmit symbols at the  $k$ -th tone and  $z_k$  is an  $L \times 1$  vector corresponding to both additive white Gaussian (AWG) noise and AXT at the  $k$ -th tone, that is:



$$z_k = q_k + n_k, \quad (2)$$

where  $q_k = [q_{1,k}, \dots, q_{L,k}]^T$  and  $n_k = [n_{1,k}, \dots, n_{L,k}]^T$  are the vectors corresponding to AXT and background noise, respectively. In Eq. (2),  $q_k$  is obtained by the multiplication  $A_k s_k$ , where  $A_k$  is a  $L \times M$  matrix with the coupling channels from the alien lines to the vectored lines, and  $s_k$  is a  $L \times 1$  vector with the signals transmitted by the alien lines. In order to simplify the notation, the underscript  $k$  will be omitted in the following equations.

### 3 I RATE BALANCING FOR AXT MITIGATION

The majority of the multiline AXT mitigation methods are based on steps in which the decoding is carried out iteratively, in which the next lines (here we assume that the lines are organized in some order) need the information of previous lines to be decoded (GEORGE GINIS e CHIA-NING PENG, 2006; A. AWASTHI et. al, 2012; AHMAD et. al, 2014). This structure allows the lines to share information and then the process is boosted. Nevertheless, this processing form causes the last lines to be more benefited from the mitigation. A consequence of this strategy is that the last lines achieve the greater rates, and the system, in an overall view, becomes unbalanced. Thus, we present here a strategy to make the AXT mitigation method to yield similar rates across the coordinated lines. We will take as basis the work presented in (AHMAD et. al, 2014) (which we will call DFC), because it needs processing only at the receiver, and then it can be applied in standard DSL applications. The DFC works as follows. In the training phase (ITU-T Rec. G. 9700, 2014) the spatial correlation matrix of the total interference is computed as:

$$R_{zz} = E[zz^H], \quad (3)$$

where the operator  $E [.]$  denotes statistical expectation, and the overscript  $H$  indicates transposing and conjugation. The spatial correlation matrix provides information to calculate the decorrelation matrix (or whitening matrix), which will remove the colored part of the interference:

$$W = R_{zz}^{-1/2}. \quad (4)$$

In the showtime transmission phase, the whitening matrix is used for multiplying the received symbols, as follows:

$$\begin{aligned} \tilde{y} &= Wy \\ &= WHx + \tilde{z}, \end{aligned} \quad (5)$$



where  $\tilde{z} = Wz$ . This operation turns the correlation matrix of the interference into an identity matrix  $I$ . However, the channel matrix has lost the column-wise diagonal dominance (CWDD). To overcome this problem, we compute the QR decomposition of the resulting matrix  $WH$  as

$$WH = QR, \quad (6)$$

where  $Q$  is a  $L \times L$  unitary matrix, and  $R$  is a  $L \times L$  upper-triangular matrix. Then, multiplying Eq. (5) by  $Q^H$ , we get

$$\begin{aligned} \hat{y} &= Q^H \tilde{y} \\ &= Rx + Q^H \tilde{z}. \end{aligned} \quad (7)$$

This operation allows the received signal to be decoded in a successive interference cancellation fashion (AHMAD et. al, 2014). Additionally, it does not change the power of the remnant interference since  $Q^H$  is unitary.

This method unbalances the rates across the lines, because the gain of each line becomes the diagonal of the  $R$  matrix, which contains decreasing values. Then, we propose the following strategy to balance the rates, which we call ordered line permutation (OLP). In the OLP, the decoding order of the lines is changed across the tones, in order to allow each line to get the advantage of being the first line (and having the greatest value of the diagonal of the  $R$  matrix as its channel gain). For example, let us assume a G.fast system with 3 lines  $a$ ,  $b$  and  $c$ . At tone  $k$  we assume the order  $a$ ,  $b$  and  $c$ , in which line  $a$  will get the greatest AXT cancellation; at tone  $k + 1$  we change the line order to  $b$ ,  $c$  and  $a$ , in which line  $b$  will appreciate the best AXT cancelling; at tone  $k + 2$  the order will be  $c$ ,  $a$  and  $b$ , and so on. This strategy is possible due to the transmission (and processing) independence among tones, and because the physical way in which the twisted pairs are plugged in the DSLAMs of the Central Office do not define the way in which the data streams will be processed.

## 4 | RESULTS

In this section, we present the results of the proposed method in comparison to the traditional use of the AXT mitigation method. In simulations, we used a program developed in MATLAB, which takes the impulse response of the channels to perform time domain simulations. Table I shows the parameters used in the simulations.

Parameter	Value
Bandwidth	106 MHz
Numberof Tones	4096
Tx PSD	-76 dBm/Hz
Background Noise PSD	-150 dBm/Hz
SINR Gap	9.75 dB
NoiseMargin	6 dB

Table I - Simulation parameters.

The data relative to the channels were obtained from measurements of 2 standard cables found in DSL deployments, and with a CAT5 cable impaired by other CAT5. One cable has 16 twisted pairs (C16) and was 200 m long. The other standard cable has 8 twisted pairs (C8), 100 m long. The CAT5 cables were 50 m long, representing a specific deployment, developed to achieve high bitrates.

In simulations we assumed very noisy scenarios, in which the number of alien lines was equal to the number of the vectored lines, to represent a situation in which the rates tend to be very unbalanced [14]. Thus, in simulations with the cable with 16 twisted pairs, 8 twisted pairs were used as the vectored lines and the other 8 were assumed to be the AXT sources. Similarly, when the simulations were fed by the cable with 8 twisted pairs, 4 lines were the vectored group, and the remaining ones were used as AXT lines. In simulations with CAT5, all the 4 pairs of a cable were assumed to be the vectored group, and the pairs of the other cable were the alien lines.

The evaluation was defined in a way in which we can observe the gain of the use of AXT mitigation, against a system that employs only in-domain crosstalk mitigation (vectoring). Additionally, the evaluation allows to compare the OLP with the traditional AXT mitigation method (AHMAD et. al, 2014). Table II summarizes the simulations results. Note that clearly the use of AXT mitigation method improves the overall performance of the system, observed by the increase in the aggregate rate (sum of the rates across the vectored lines) when AXT mitigation is active. For example, in simulations with the cable with 16 twisted pairs, the AXT mitigation has increased the aggregate rate in approximately 40%. In the same table, one can also observe the high range of values in the column referent to traditional AXT mitigation, in which in simulations with C16 we have observed rate variations from 97.91 Mbps up to 145.81 Mbps, from 209.3 up to 354.11 in the simulations with C8, and from 317.08 up to 673.36 in the results for CAT5. In the former, the standard deviation is approximately 21 Mbps, in the second the standard deviation is around 65 Mbps, and in the CAT5, 152.43 Mbps. Nevertheless, when the proposed method was employed, this standard deviation have passed to 1.96 Mbps, 2.17 Mbps, and 54.98, respectively, confirming the efficiency of the proposed method on balancing the rates. In spite of the method having achieved a considerable reduction in the bitrate's variance, the standard deviation in the CAT5 kept a great value, due to the great transmission rate achieved by the line

c when compared to the other ones. This fact is supported because, in general, the AXT channel to the line is weaker than in the other one. Additionally, we can observe an increase in the rates across the scenarios, which is attributed, mainly, to the cable length reduction, that allows higher transmission rates.

Line	Vectoring	AXT mitigation	OLP
<b>Cable with 16 twisted pairs - 200 m</b>			
a	91.36	91.38	121.44
b	93.32	106.36	122,22
c	111.81	145.81	124,10
d	92.91	118.81	118,96
e	97.58	142.74	121.49
f	83.23	142.54	125.33
g	57.51	97.91	123.63
h	74.97	131.28	121,90
aggregate rate	702.67	976.81	979.06
standard dev.	16.23	21.64	1.96
<b>Cable with 8 twisted pairs - 100 m</b>			
a	248.6	248.275	260.42
b	297.68	354.11	260.795
c	177.265	223.51	257.055
d	166.075	209.03	256.68
aggregate rate	889.62	1034.925	1034.95
standard dev.	62.08	65.61	2.17
<b>CAT5 cable - 100 m</b>			
a	537.03	673.36	475.81
b	241.19	317.08	484.57
c	420.23	436.79	591.39
d	548.23	548.48	484.82
aggregate rate	1746.68	1975.71	2036.59
standard dev.	142.59	152.43	54.98

Table II - Rate achieved (in mbps) in the uplink direction by each vectored line in 2 distinct cables in scenarios with AXT. Vectoring denotes the situation in which only vectoring is employed; AXT mitigation indicates that traditional AXT mitigation was used; and AXT rate balancing refers to AXT mitigation with rate balancing.

We can also note in Table II that the rate of some lines was increased considerably with the OLP method when compared to the standard AXT mitigation method. For example, in simulations with the C16 the rate of line was increased approximately 22%, and line of the CAT5 simulations has experienced a rate improvement of around 35%. In order to have a better view of the balance improvement provided by the proposed method, we have plotted the rates of lines from C16 and C8 in Figure 2 and Figure 3, respectively.

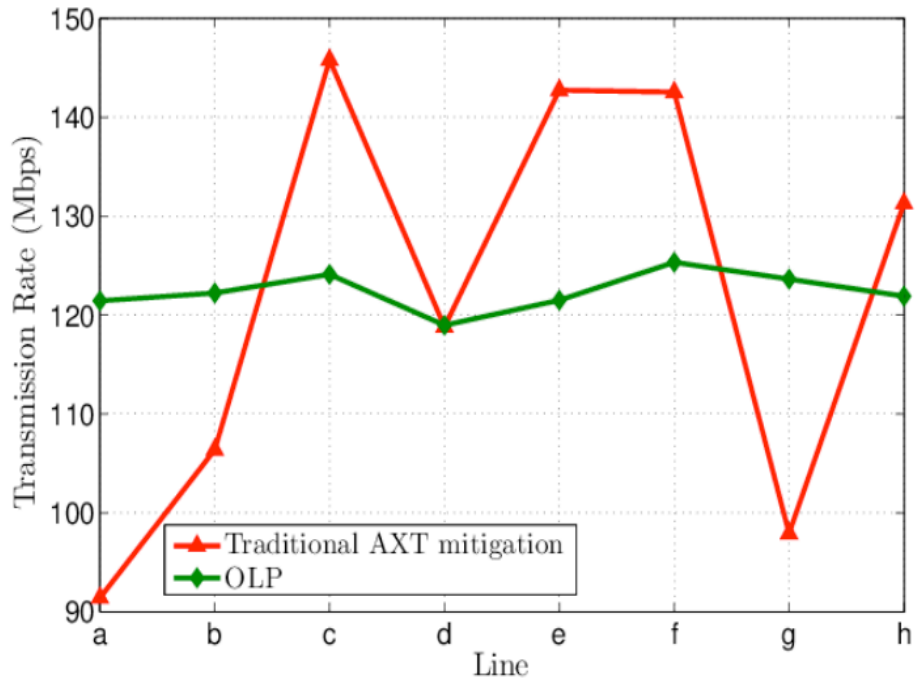


Figure 2: Comparison of transmission rate achieved by each line with traditional AXT mitigation method and with OLP in the scenario with C16 cable

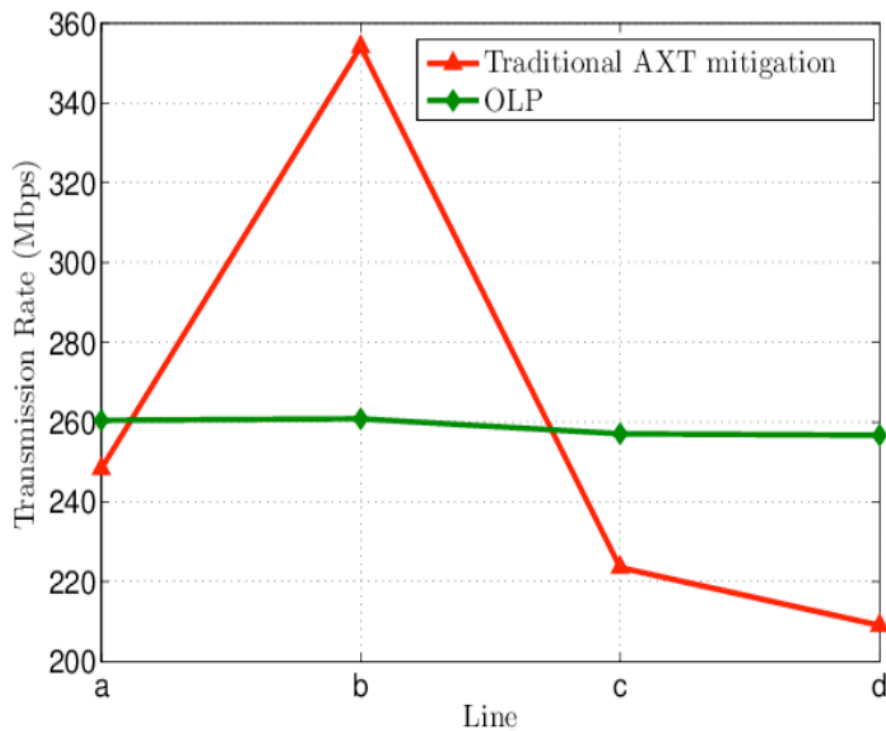


Figure 3: Comparison of transmission rate achieved by each line with traditional AXT mitigation method and with OLP in the scenario with C8 cable.

The transmission rates showed in Table II seem to be higher for the current consumption demand, and one could argue that the rate gained by the rate balancing is not significant due to the high value of the original transmission rate. However, this technique can, for example, supply a Wi-Fi Network of a cafe, a hotel or a restaurant (Z. SHAH et. al, 2015). In these environments, the transmission rate of the Internet access must be suitable to accommodate the many clients using online multimedia

services, such as video conference. In the other hand, there are studies (ERICSSON, 2015) that expect a fast growing in the number of equipment connected to the Internet (Internet of Things (L. TAN e N. WANG, 2010), for example), in which a balanced access network can provide a similar experience for all the subscribers. Additionally, this rate equilibrium allows the telecommunications companies to deliver similar service packages for the consumers, and to provide the rates necessary for the current and future online services.

## 5 | CONCLUSIONS

This document has presented an empirical method to achieve rate balancing when alien xtalk mitigation is applied in G.fast systems (OLP). The OLP is quite simple and does not have an exhaustive phase for the search of the optimal users order. The OLP is applied in a tone perspective, in which the position of the lines in the decoding is regularly altered. The simulations results have shown good standard deviation reduction in the transmission rates through the coordinated lines, in which measurements of the standard DSL cables were used. For example, in a scenario with 4 lines, the standard deviation was reduced from 65 Mbps to 2.1 Mbps. Future works include the comparison of the OLP with the state of the art rate-balancing algorithms and an evaluation of the effect of the rate balancing in the QoE of the users.

## REFERENCES

- A, TARIGHATI; J, JALDEN. **Optimality of rate balancing in wireless sensor networks**. IEEE Transactions on Signal Processing, vol. PP, no. 99, pp. 1–1, 2016.
- A, AWASTHI; N, AL-DHAHIR; O. E. ELIEZER; PORAS T. BALSARA. **Alien crosstalk mitigation in vectored DSL systems for backhaul applications**. IEEE International Conference on Communications, pp. 3852 – 3856, 2012.
- AHMAD A. AMAYREH; JRME LE MASSON; MARY LINEHLARD; MERYEMOUZZIF. **Alien crosstalk elimination in digital subscriber line systems**. IET Communications, pp. 1714 – 1723, 2014.
- ALCATEL-LUCENT. **G.fast: Comparison of linear and non-linear precoding**. Geneva, Switzerland: Contribution ITU-T 2013-01-Q4-046, 2013.
- BRITISH TELECOM. **G.fast: Release of BT cable measurements for use in simulations**. Geneva, Switzerland: Contribution ITU-T 2013-01-Q4-066, 2013.
- C. HELLINGS; W. UTSCHICK. **Energy-efficient rate balancing in vector broadcast channels with linear transceivers**. In 2012 International Symposium on Wireless Communication Systems (ISWCS), Aug 2012, pp. 1044–1048.
- D. A. GOMES; I. FREIRE; A. KLAUTAU; E. PELAES. **Feasibility of alien crosstalk mitigation with receiver-side MIMO processing on G.fast systems**. In 2015 International Workshop on Telecommunications, Jun. 2015.
- ERICSSON. **ERICSSON MOBILITY REPORT**. Tech. Rep. Disponível em < <https://www.ericsson.com/en/mobility-report>>, Acessado em: 05 maio, 2015.

G. GINIS; C. N. PENG. **Alien crosstalk cancellation for multipair digital subscriber line systems**. Applied Signal Processing, EURASIP Journal on, vol. 2006, pp. 1–12, 2006.

GILBERT STRANG. **Introduction to Applied Mathematics**. Wellesley Cambridge Press, 1 edition, 1986.

ITU-T. **Self-FEXT cancellation (vectoring) for use with VDSL2 transceivers**. 2010. Rec. G.993.5-2010, International Telecommunication Union. Disponível em <<https://www.itu.int/rec/T-REC-G.993.5>> acessado em: 03 maio. 2010

ITU-T. **Fast access to subscriber terminals (G.fast) - physical layer specification**. Rec. G. 9700-2014, International Telecommunication Union. Disponível em <[https://www.itu.int/rec/dologin\\_pub.asp?lang=e&id=T-REC-G.9700-201404-I!!PDF-E&type=items](https://www.itu.int/rec/dologin_pub.asp?lang=e&id=T-REC-G.9700-201404-I!!PDF-E&type=items)> acessado em: 02 maio. 2014.

L. TAN; N. WANG. **Future internet: The internet of things**. 3rd International Conference on Advanced Computer Theory and Engineering (ICACTE), pp. 376–380, Aug. 2010.

N. Li; X. Tao; Q. Cui; J. Bai; Y. Hu. **Secrecy rate balancing for the downlink multiuser mimo system with independent confidential messages**. In 2014 IEEE Wireless Communications and Networking Conference (WCNC), April 2014, pp. 2173–2177.

R. CENDRILLON; G. GINIS; E. V. D. BOGAERT; M. MOONEN. **A near-optimal linear crosstalk precoder for downstream VDSL**. Communications, IEEE Transactions on, vol. 55, no. 5, pp. 860–863, 2007.

R. ZIDANE; S. HUBERMAN; C. LEUNG; THOLENGOC. **Vectored DSL: Benefits and Challenges for Service Providers**. IEEE Communications Magazine, Feb. 2013.

Z. SHAH; S. RAU; ADEELBAIG. **Throughput comparison of ieee 802.11ac and ieee 802.11n in an indoor environment with interference**. International Telecommunication Networks and Applications Conference, pp. 196–201, Nov. 2015.

## REPRESENTAÇÃO ESPARSA UTILIZANDO WAVELETS E VARIAÇÃO TOTAL APLICADOS AO PROCESSAMENTO DE SINAIS DE DESCARGAS PARCIAIS

**Paulo Vitor do Carmo Batista**

Instituto Federal de Educação, Ciência e  
Tecnologia do Norte de Minas Gerais  
Januária – MG

**Hilton de Oliveira Mota**

Universidade Federal de Minas Gerais,  
Departamento de Engenharia Elétrica  
Belo Horizonte – MG

**RESUMO:** Técnicas de processamento de sinais são desenvolvidas com o intuito de auxiliar no processo de manutenções em equipamentos, evitando falhas e paradas indesejadas. Uma dessas falhas é conhecida como descarga parcial. O objetivo deste trabalho é identificar as descargas parciais de um sinal, aplicando um método conhecido como Variação Total com *Wavelets*. São utilizados sinais sintéticos e medidos em laboratório contendo ruídos gaussianos, harmônicos e impulsivos. Os resultados obtidos são promissores e mostram que o método reconstrói bem o sinal, mantendo as descargas.

**PALAVRAS-CHAVE:** Processamento de Sinais, *Wavelets*, Variação Total, Descargas Parciais.

**ABSTRACT:** Techniques of signal processing are developed with the purpose of assisting in the process maintenance in equipment, avoiding failures and unwanted stops. One

such fault is known as partial discharge. The objective of this work is to identify the partial discharges of a signal, applying a method known as Total Variation with Wavelets. Synthetic and laboratory signals containing gaussian, harmonic and impulsives noises are used. The results obtained are promising and show that the method reconstructs well the signal, maintaining the discharges.

**KEYWORDS:** Signal Processing, Wavelets, Total Variation, Partial Discharge.

### 1 | INTRODUÇÃO

Descargas parciais (DPs) são descargas elétricas que acontecem parcialmente no interior de sistemas de isolamento que, em sua maioria, são compostos por resinas, plástico ou borracha. Normalmente são formadas cavidades durante o processo de construção desses sistemas, seja por defeito de fabricação seja em função das características construtivas destes sistemas de isolamento.

São nessas cavidades que ocorrem as descargas, que são ditas parciais por não ocasionar uma ruptura total do sistema de isolamento, ou seja, não une as duas extremidades dessa região. Em Boggs (1990), o autor define que uma DP ocorre quando alguma forma de atividade elétrica dentro do sistema



resulta em uma mudança rápida da configuração de campo elétrico que faz com que uma corrente flua em um condutor conectado ao mundo externo.

As DPs ocasionam vários fenômenos físicos e químicos entre os quais pode-se destacar:

- Fenômenos elétricos;
- Emissão acústica;
- Transformações químicas;
- Geração de luz ultravioleta;
- Radiação eletromagnética.

Entre estes, a detecção de DPs por meios elétricos é geralmente a mais utilizada por fornecer informações mais detalhadas (MOTA et al, 2016).

Como mostrado na Figura 1, as descargas apresentam como características localidade temporal e baixa energia, aparecendo em forma de pulsos elétricos de curta duração. Sua reincidência pode ocasionar defeitos nos sistemas de isolamento devido às alterações que causam nas estruturas físicas e químicas dos materiais (MOTA et al, 2016). Em casos mais extremos, podem ocorrer rupturas no isolamento causando graves falhas de operação, como por exemplo em sistemas de alta tensão.

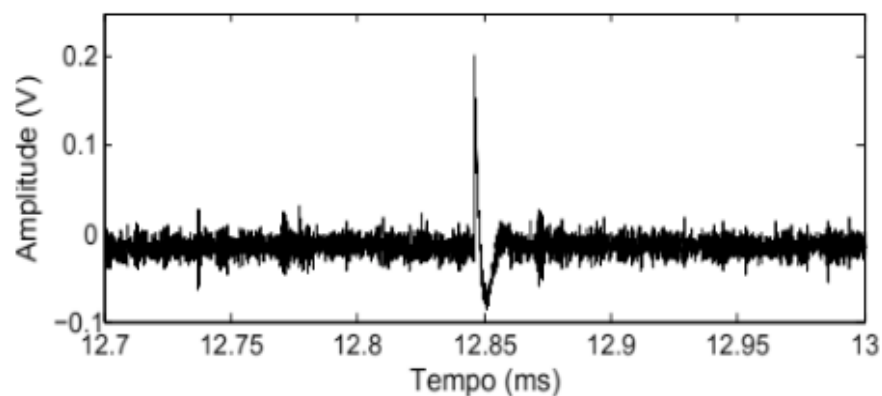


Figura 1 - Descarga Parcial

Tais falhas resultam em altos prejuízos financeiros, em manutenção ou substituição de equipamentos. Desta forma faz-se necessária a utilização de métodos preditivos com o intuito de verificar a existência de descargas parciais.

## 2 | PROCESSAMENTO DE SINAIS DE DESCARGAS PARCIAIS

Sistemas de monitoramento, processamento e diagnóstico do fenômeno de DPs fornecem informações que permitem identificar e prever falhas que podem ocorrer

em equipamentos. Assim, são desenvolvidas e utilizadas técnicas de processamento de sinais que visam possibilitar intervenções e manutenções programadas, evitando falhas e paradas indesejadas.

Técnicas como a Transformada de *Fourier* (TF) e filtros FIR (*Finite Impulse Response*) são muito utilizadas em filtragem de sinais. No entanto, por constituir-se de uma soma ponderada de senos e cossenos que são sinais periódicos, a representação das DPs por Transformada de *Fourier* não é bem feita por se tratar de pulsos de características transitórias e não-estacionárias.

Outra técnica existente, a Transformada de *Wavelets* (TW), tem sido cada vez mais utilizada, pois entre outras características, permite a decomposição do sinal em componentes localizados no tempo (translação do sinal) e na escala (dilatação/ contração do sinal). Isso favorece a representação de sinais especificamente localizados, como é o caso de sinais de DPs.

Em uma variação da TW, em Coifman e Donoho (1995) os autores propuseram aplicar os métodos de eliminação de ruídos não apenas ao sinal original, mas a versões deslocadas circularmente de uma unidade de tempo, obtendo como sinal reconstruído a média dos sinais obtidos para cada deslocamento. Esse algoritmo é conhecido como Transformada de *Wavelets* Não-Dizimada (TWND) ou invariante ao deslocamento.

Um método frequentemente utilizado para eliminação de ruídos em imagens digitais, como visto em Condat (2013), Figueiredo et al (2006) e Strong e Chan (2003), é o método da Variação Total (VT) que, proposto em Rudin et al (1992), é baseado em um tipo de otimização restrita. O VT baseia-se no princípio de que os sinais com oscilações excessivas, possivelmente associadas a ruídos, e possuem alta variação total, ou seja, a integral do gradiente absoluto do sinal é alta. Busca-se então minimizar essa variação total sujeito a restrições impostas usando multiplicadores de Lagrange..

Assim, neste capítulo é apresentado os resultados obtidos aplicando o método proposto em Ding e Selesnick (2015) em sinais de descargas parciais, que ao realizar análise visual e estatística, percebe-se a eficiência do método, que resulta excelentes níveis de atenuação de diversos tipos de ruído.

### 3 | TRABALHOS RELACIONADOS

Sabe-se que a Transformada de *Wavelets* Dizimada (TWD) fornece um conjunto de coeficientes que, associados a respectivas funções base, permitem reconstruir qualquer sinal de energia finita. A partir dessa ideia é possível reconstruir parte do sinal selecionando os coeficientes de interesse. Assim, em Donoho (1995) o objetivo é reconstruir um sinal baseando-se em um limiar previamente estabelecido. Todo coeficiente menor que este valor foi então considerado ruído e, conseqüentemente, zerado; os coeficientes resultantes foram então utilizados na reconstrução do sinal. Esta metodologia não apresenta bons resultados nos sinais utilizados neste trabalho, visto que os ruídos impulsivos possuem características semelhantes às descargas

parciais. Sendo assim, as descargas podem ser eliminadas ou os ruídos impulsivos, mantidos.

Já o trabalho apresentado em Ma et al (2000) envolveu TW e descargas parciais, propondo um método de filtragem desses sinais utilizando DSP (*Digital Signal Processor*). O trabalho obteve expressivos resultados em desempenho e custo computacional para atenuação de ruídos senoidais de alta frequência e ruídos impulsivos.

Em outra pesquisa envolvendo DPs, os autores em Cuenca et al (2004) utilizaram a análise de componentes principais (*Principal Component Analysis* - PCA) para reduzir a dimensão do espaço original de entrada, gerando subespaços de menor dimensão produzidos pelas componentes mais relevantes, para auxiliar na identificação de tais descargas em transformadores de potência. Foram descartadas as projeções que apresentavam baixa energia na representação e manteve-se as demais. Após aplicar o resultado em uma rede neural, obteve-se 93% de acurácia. Diferentemente do trabalho aqui apresentado, a análise de componentes principais e aplicação em uma rede neural eleva o custo computacional do processo. Sendo assim, optou-se por utilizar otimização convexa.

Já no trabalho de de Oliveira Mota et al (2011), os autores baseados em um procedimento de seleção de coeficientes adaptativa espacialmente, propuseram um método que visasse explorar a capacidade de processamento localizado da TW. Assim, obtiveram informações adicionais que aprimorou a separação entre os coeficientes relacionados ao sinal e os coeficientes relacionados ao ruído, obtendo excelentes resultados.

Por outro lado, utilizando a técnica de otimização *Basis Pursuit* através do algoritmo SALSA, em de Azevedo Silva et al (2017) é apresentado um método de filtragem de sinais de DPs utilizando dicionários de *wavelets* sobrecompletos. O método nomeado *Overcomplete Dictionary Basis Pursuit Denoising* (ODBPD) consiste na união de diferentes famílias *wavelets* com o objetivo de obter representações mais próximas das características do sinal. Assim como o *Wavelet Total Variation* (WATV), esse método utiliza otimização convexa e dicionários sobrecompletos em sinais de DPs. No entanto, ao aplicá-lo em sinais com ruídos impulsivos, o método não apresentou bons resultados. Tanto as DPs quanto os ruídos impulsivos mantiveram-se no sinal, ou eram eliminados juntos.

Já em Mitiche et al (2017) os autores desenvolveram uma nova abordagem ao método *Adaptive Local Iterative Filtering* (ALIF) proposto por Cicone et al (2016). Inspirado no método WATV e denominado ALIF-TV, o método busca reconstruir o sinal sem ruído. Seus testes foram realizados em sinais de DPs medidos com ruídos Gaussianos, obtendo bons resultados. No entanto, além de apresentar elevado custo computacional, assim como em Donoho (1995) o método funciona bem para sinais com ruídos Gaussianos apenas.

Percebendo que até então não há uma maneira convencional para realizar

*denoising* em sinais de descargas parciais, buscou-se então uma alternativa aos métodos já existentes.

#### 4 | MÉTODO UTILIZADO

A VT foi proposta em Rudin et al (1992) como um critério de regularização para eliminação de ruídos, e consiste em medir o quanto seus valores mudam entre si. Em um sinal  $x(n)$  de  $N$  pontos, a VT é definida como:

$$VT(x) = \sum_{n=2}^N |x_n - x_{n-1}|$$

que também pode ser escrita em forma matricial como

$$VT(x) = \|Dx\|_1 \quad (2)$$

onde  $\|x\|_1$  é a norma  $\ell_1$  de  $x$  e  $D$  é a matriz de diferenças de primeira ordem de tamanho  $(N-1) \times N$ .

Sendo assim, a abordagem denominada eliminação de ruídos por *total variation* tem como objetivo encontrar uma aproximação que obtenha a menor variação total cuja reconstrução esteja mais próxima ao sinal original  $x(n)$ . Entre algumas medidas de proximidade possíveis, a soma dos erros quadrados é eficiente e é dada pela Equação (3):

$$E(x, y) = \frac{1}{2} \sum_{n=1}^N (x_n - y_n)^2 \quad (3)$$

onde  $y_n$  é o sinal reconstruído.

A eliminação de ruído por variação total consiste em minimizar a função objetivo dada pela Equação (4):

$$F = \min_x \|x - y\|_2^2 + \lambda \|Dx\|_1 \quad (4)$$

onde  $\lambda$  é um parâmetro de regularização.

Baseado na Equação (4), no algoritmo SALSA proposto em Afonso et al [2010] e utilizado neste trabalho, em Ding e Selesnick (2015) os autores propuseram a técnica conhecida como WATV, que consiste no emprego de penalidades não convexas com o objetivo de promover a esparsidade no domínio de *wavelets* Nikolova et al (2010). Ao utilizar penalidades não convexas, afeta-se a convexidade da função objetivo. No

entanto, em seu método, os autores restringem tal penalidade de modo a garantir a convexidade daquela função.

Sendo a transformada de *wavelets* representada por  $W$  e  $w$  os coeficientes que satisfazem  $w=Wx$ , o método utilizado neste trabalho, consiste em encontrar os coeficientes *wavelets* que resolvem o problema de otimização da Equação (5):

$$w^* = \underset{w}{\operatorname{argmin}} (F(w) = \frac{1}{2} \|Wy - w\|_2^2 + \sum_{j,k} \lambda_j \phi(w_{j,k}; a_j) + \beta \|DW^T w\|_1) \quad (5)$$

em que  $\lambda_j$  são os multiplicadores de Lagrange, e  $a_j$  e  $\beta$  são parâmetros de regularização.

Tais parâmetros são calculados baseados nas Equações (6), (7) e (8).

$$\lambda_j = \frac{2.5\eta\sigma}{2^{\frac{j}{2}}} \quad (6)$$

$$\beta = \frac{(1 - \eta)\sqrt{N}\sigma}{4} \quad (7)$$

$$a_j = \frac{1}{\lambda_j} \quad (8)$$

onde  $0 < \eta < 1$  controla o peso da *wavelet* e a regularização da VT, e  $\sigma$  é a variância do ruído.  $\sigma$  e  $\eta$  são definidos empiricamente neste trabalho, respectivamente como 0.15 e 0.99 para os sinais sintéticos. Para os sinais medidos, o parâmetro  $\sigma$  variou para cada sinal, pois os melhores resultados foram atingidos usando os valores apresentados na Tabela 1. O parâmetro  $a$ , neste caso, garante a convexidade da função objetivo por ser inversamente proporcional a  $\lambda_j$  na Equação (9):

$$f(x) = \frac{1}{2} (y - x)^2 + \lambda\phi(x; a) \quad (9)$$

O critério para a escolha da família a ser utilizada no processamento do sinal foi a relação entre a curva e a descarga parcial. Essa medida é dada pelo coeficiente de correlação, que como visto em Ma et al (2002), possui maior valor para a família *Daubechies 2*.

Ainda em Ma et al (2002), entre as *wavelets* disponíveis, a *Daubechies 2* possui quase todas as propriedades necessárias, como compactação, duração limitada, ortogonalidade e assimetria para análise de pulsos irregulares e transitórios rápidos.

Por fim, empiricamente adotou-se 4 níveis de decomposição para os sinais

sintéticos. Já para os sinais medidos, a quantidade de níveis é apresentada na Tabela 1.

Índice do Sinal	Tipo do Ruído	$\sigma$	$\eta$	Níveis
1	Harmônico	0.01	0.99	5
2	Harmônico	0.007	0.99	5
3	Impulsivo	0.02	0.99	10
4	Impulsivo	0.09	0.99	3
5	Harmônico	0.09	0.99	5

Tabela 1- Características dos Sinais Medidos

A função de penalidade  $\phi$  utilizada foi a arc-tangente definida na Equação (10) que como pode ser visto na Figura 2 é uma função de limiar não-linear contínua, que apresenta valores sempre entre os valores das funções usuais *hard* e *soft-thresholding*.

$$\phi(x; a) = \begin{cases} \frac{2}{a\sqrt{3}} \left( \tan^{-1} \left( \frac{1+2a|x|}{\sqrt{3}} \right) - \frac{\pi}{6} \right), & \text{se } a > 0, \\ |x|, & \text{se } a = 0. \end{cases} \quad (10)$$

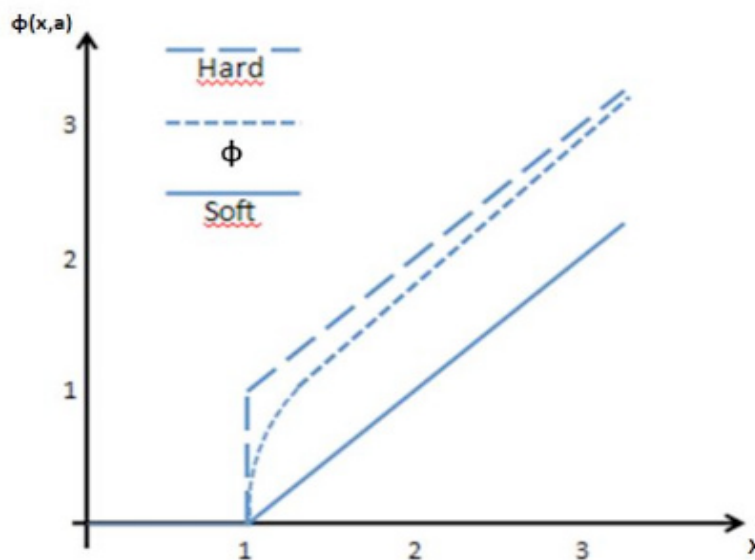


Figura 2- Função Threshold

## 5 | SINAIS UTILIZADOS

Com o objetivo de avaliar o método de filtragem, é necessário aplicá-lo a sinais simulados. Conhecer o sinal antes da aplicação do método torna possível avaliá-lo quantitativamente utilizando-se métricas que comparam o sinal original e o sinal obtido após aplicação do método.

Assim, foram gerados 20 sinais de DPs sintéticos e em cada um destes sinais foram adicionados ruídos do tipo:

- ruído branco (Gaussiano);
- ruído harmônico (tipo AM – Amplitude Modulada);
- ruído do tipo impulsivo;

que também foram gerados de maneira sintética.

Além desses sinais, aplicou-se o método à 5 sinais medidos no laboratório de medição e instrumentação computacional da Universidade Federal de Minas Gerais. Todos eles fazem parte de uma base de dados de sinais de DPs oriundos de pesquisas contidos em Mota et al (2001).

## 6 | RESULTADOS

Dados os sinais apresentados na Seção IV, objetiva-se com o método, filtrar o sinal preservando as descargas parciais (pulsos) nele contidas.

A Figura 3 apresenta um sinal simulado com descargas parciais regularmente espaçadas corrompido por ruído gaussiano, e o mesmo sinal após aplicação do método de filtragem WATV. Como pode ser visto, o método apresentou excelentes níveis de atenuação do ruído gaussiano, eliminando-o quase totalmente. Além disso, manteve as descargas parciais apresentando baixo nível de atenuação da amplitude das mesmas.

A Figura 4 corresponde a um sinal simulado com descargas parciais regularmente espaçadas corrompido por ruídos impulsivos. Como pode ser visto, os ruídos impulsivos foram praticamente eliminados e as descargas parciais foram levemente atenuadas.

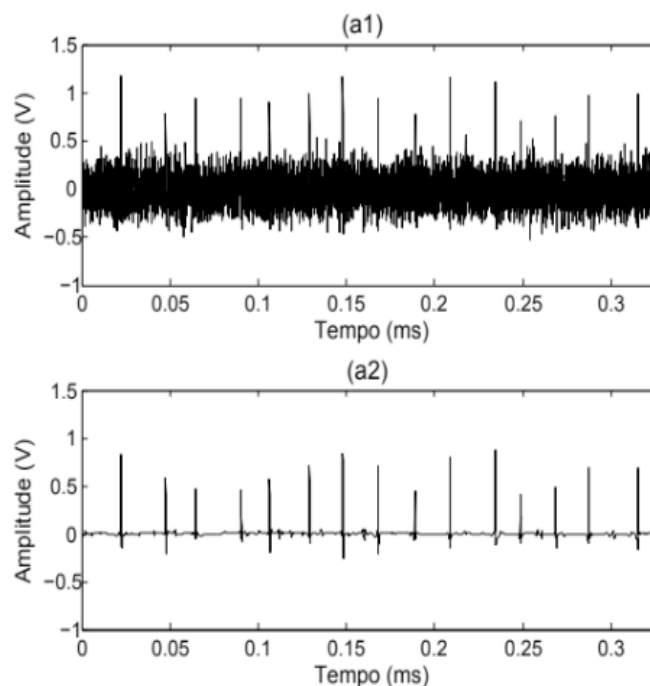


Figura 3- Sinal Simulado com Ruído Gaussiano



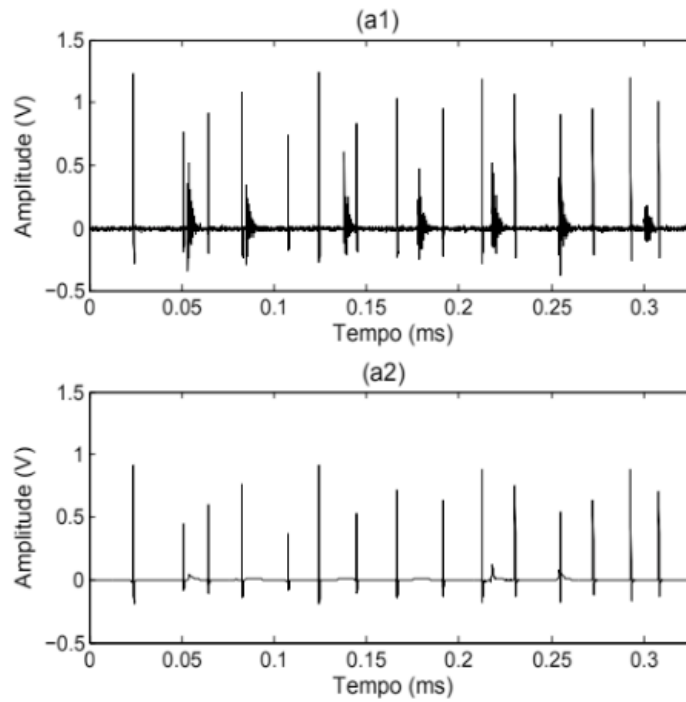


Figura 4 - Sinal Simulado com Ruído Impulsivo

A Figura 5 representa um sinal simulado com descargas parciais regularmente distribuídas corrompido por ruído harmônico do tipo AM. A mesma figura apresenta também o sinal após aplicar o método WATV. Nota-se que obteve-se excelente atenuação do ruído, sendo este praticamente eliminado, além de manter as descargas parciais que foram levemente atenuadas.

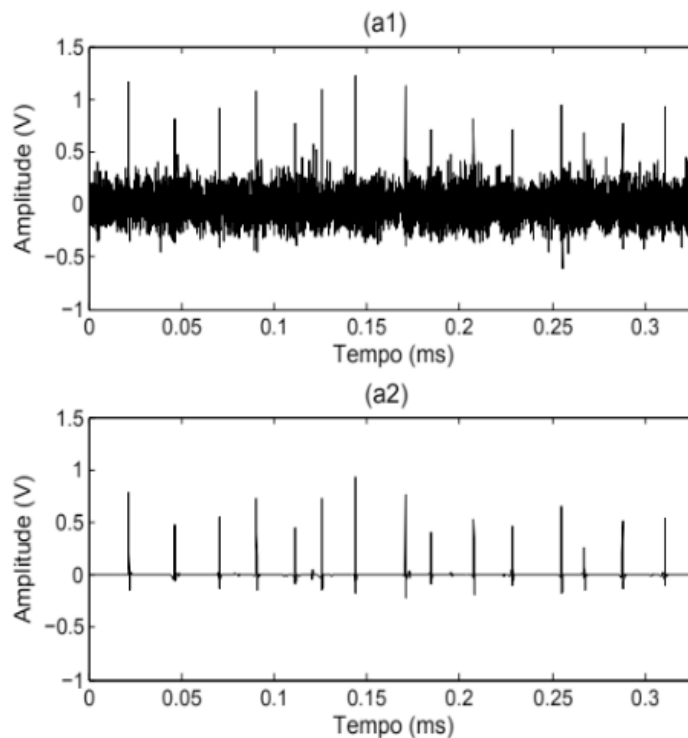


Figura 5- Sinal Simulado com Ruído Harmônico

Após empregar o método em sinais simulados e verificar sua eficiência, aplicou-se o mesmo a sinais medidos em laboratório.

A Figura 6 corresponde ao sinal 1 da Tabela 1, e possui ruído predominantemente harmônico e diversas descargas parciais concentradas entre os instantes 12 ms e 15 ms. Ao aplicar o método, percebe-se no sinal filtrado a eliminação quase total do ruído harmônico e a manutenção das descargas parciais, que foram suavemente atenuadas.

Assim como o sinal 1 mostrado na Figura 6, o sinal 5, mostrado na Figura 7, possui ruído predominantemente harmônico, com alguns ruídos impulsivos distribuídos. Diferentemente do sinal anterior, ele possui apenas três descargas parciais entre os instantes 9 ms e 14 ms, que após aplicação do método foram mantidas com suave atenuação. Como pode ser visto, o ruído harmônico bem como os ruídos impulsivos foram removidos, mostrando que para este tipo de sinal o método apresentou-se muito eficiente.

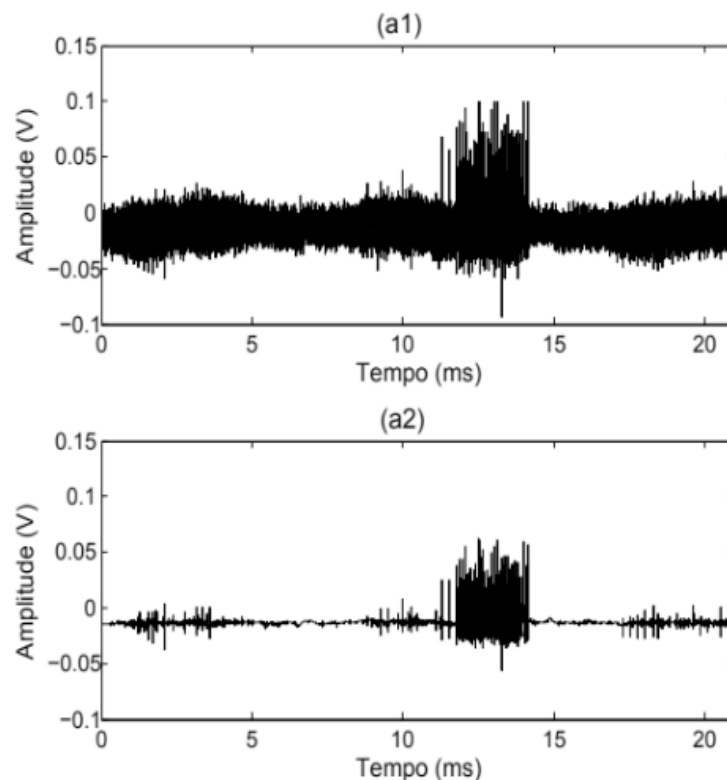


Figura 6- Sinal Medido 1 com Ruído Harmônico

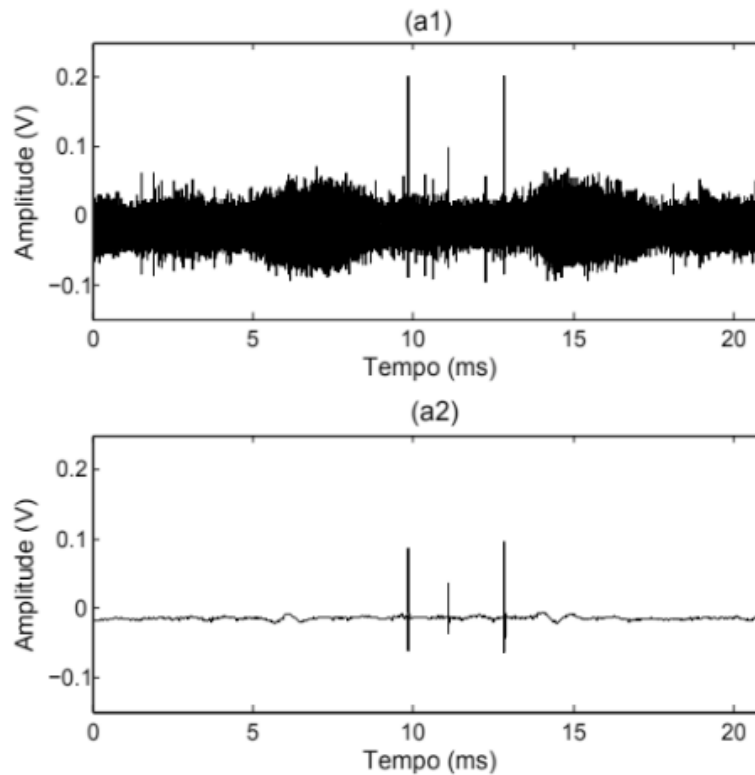


Figura 7- Sinal Medido 5 com Ruído Harmônico

Como pode ser visto na Figura 8, o sinal 2 também é composto por ruído harmônico do tipo AM. No entanto, este sinal possui diversas descargas parciais entre os instantes 1 ms e 4 ms com amplitudes que variam entre 0.02 V e 0.1 V. Assim como no resultado do método aplicado ao sinal 1, o ruído contido neste sinal não foi totalmente eliminado, mas houve atenuação do mesmo e as descargas parciais foram mantidas.

A Figura 9 corresponde ao sinal 3 que contém, ao longo de sua duração, diversos ruídos impulsivos. No entanto, como mostrado em a2 e a3 na Figura 10, nos instantes 10.28 ms e 11.76 ms acontecem duas descargas parciais com amplitudes aproximadas de 0.2 V e 0.1 V respectivamente. Além disso, ressalta-se um ruído impulsivo de amplitude aproximada de 0.2 V que ocorre aos 16.14 ms mostrado em a4, que foi totalmente eliminado.

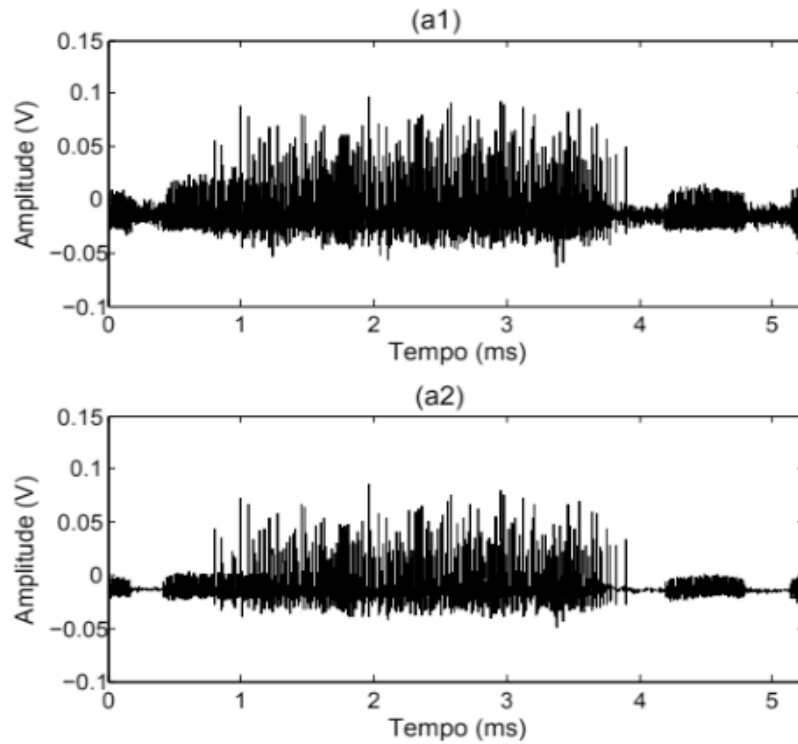


Figura 8- Sinal Medido 2 com Ruído Harmônico

Por fim, o sinal 4 mostrado na Figura 11 possui predominantemente ruído harmônico, duas descargas parciais nos instantes 10.2 ms e 11.6 ms e, no instante 14.98 ms apresenta um ruído impulsivo com amplitudes positiva e negativa de aproximadamente 0.5 V, como mostrados respectivamente em a2, a3 e a4 na Figura 12. Após aplicar o método, as descargas foram mantidas e os ruídos harmônico e impulsivo foram totalmente eliminados.

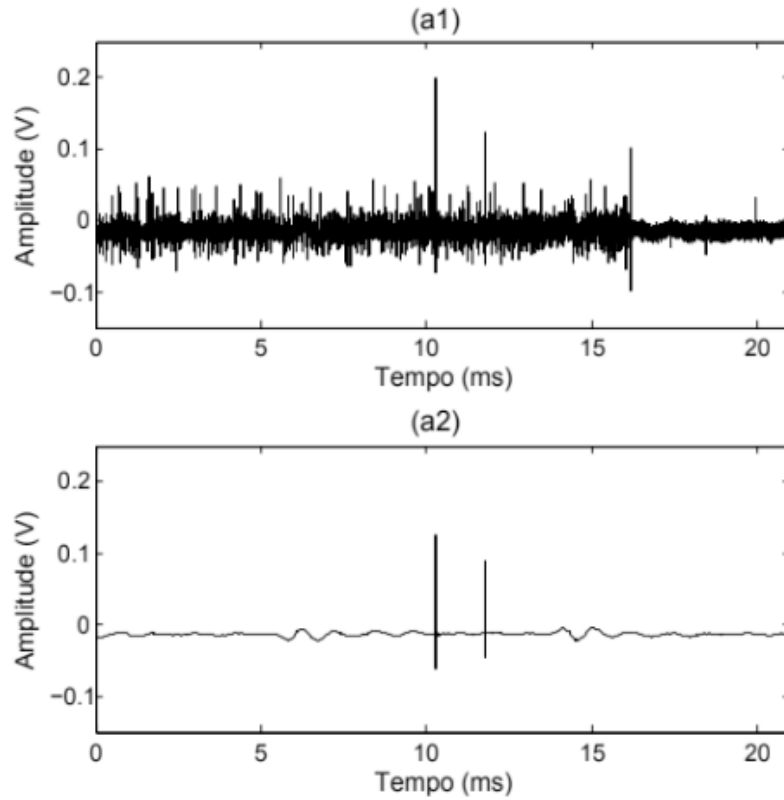


Figura 9- Sinal Medido 3 com Ruído Impulsivo

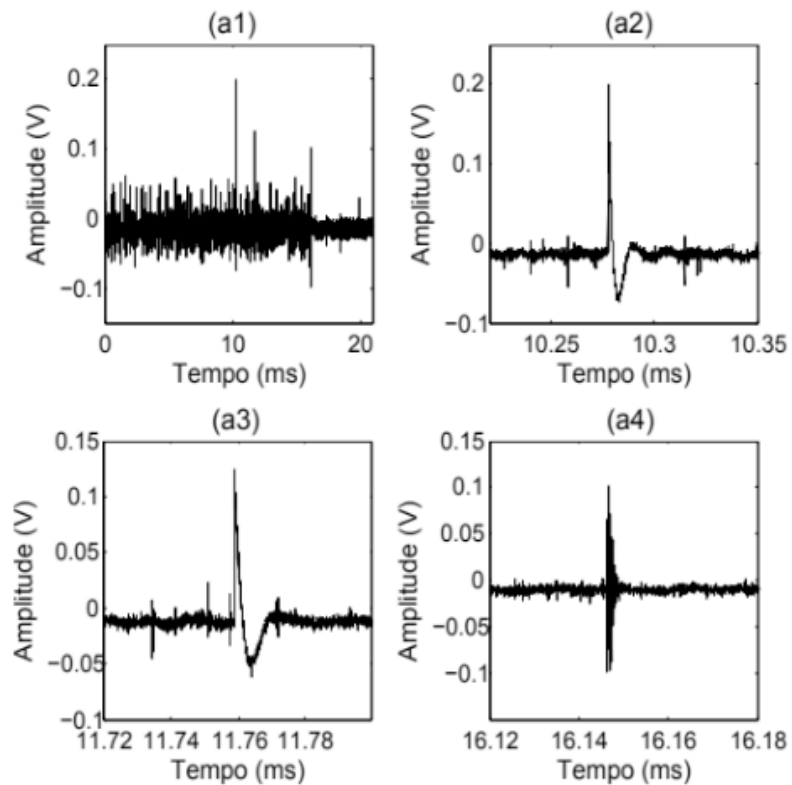


Figura 10- Detalhes do Sinal Medido 3

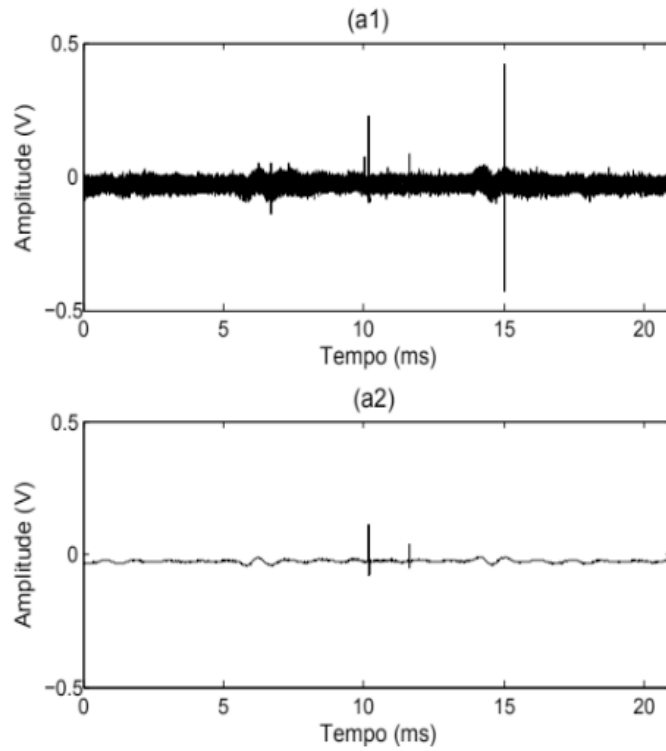


Figura 11- Sinal Medido 4 com Ruído Impulsivo

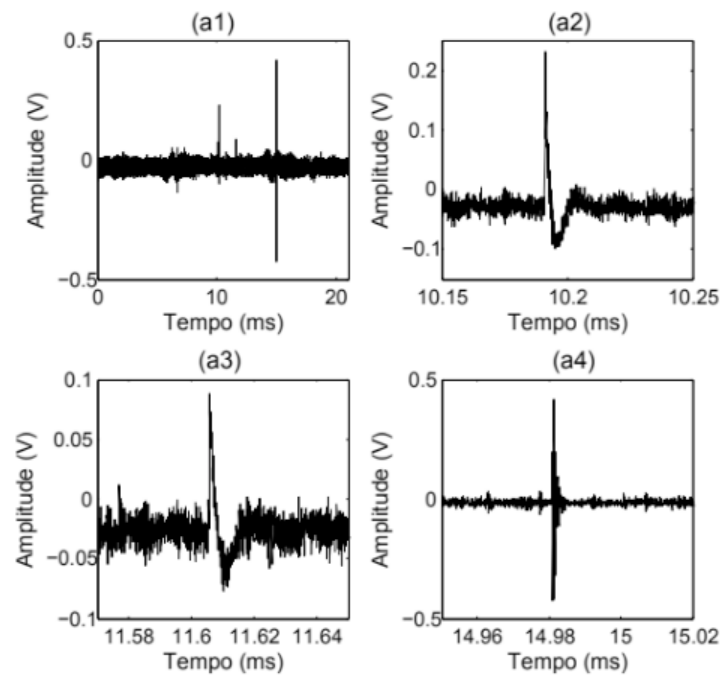


Figura 12- Detalhes do Sinal Medido 4

Após analisar os resultados do método em sinais sintéticos e medidos, é feita uma análise estatística de sua eficiência. Uma boa medida para atestar sua precisão é o RMSE (*Root Mean Square Error*) (11), que calcula a raiz do erro quadrático médio entre o sinal desejado e o sinal obtido.

$$RMSE = \sqrt{\frac{1}{2} \sum_{n=1}^N (x_n - x_n^*)^2} \quad (11)$$

Assim, na Tabela 2 são apresentados os valores de RMSEs para os métodos WATV e ODBPD nos 60 sinais simulados. Percebe-se então que o método WATV apresentou maior RMSE apenas para o sinal 12 corrompido por ruído gaussiano.

Sinal	Gaussiano		Impulsivo		Harmônico	
	ODBPD	WATV	ODBPD	WATV	ODBPD	WATV
1	0.0417	<b>0.0362</b>	0.0390	<b>0.0222</b>	0.0322	<b>0.0231</b>
2	0.0406	<b>0.0350</b>	0.0362	<b>0.0231</b>	0.0293	<b>0.0222</b>
3	0.0396	<b>0.0388</b>	0.0270	<b>0.0192</b>	0.0305	<b>0.0219</b>
4	0.0404	<b>0.0339</b>	0.0370	<b>0.0219</b>	0.0304	<b>0.0231</b>
5	0.0409	<b>0.0374</b>	0.0381	<b>0.0224</b>	0.0322	<b>0.0249</b>
6	0.0395	<b>0.0359</b>	0.0380	<b>0.0246</b>	0.0322	<b>0.0218</b>
7	0.0440	<b>0.0370</b>	0.0294	<b>0.0227</b>	0.0301	<b>0.0228</b>
8	0.0392	<b>0.0357</b>	0.0369	<b>0.0219</b>	0.0306	<b>0.0236</b>
9	0.0435	<b>0.0345</b>	0.0315	<b>0.0205</b>	0.0301	<b>0.0219</b>
10	0.0406	<b>0.0349</b>	0.0226	<b>0.0182</b>	0.0336	<b>0.0256</b>
11	0.0406	<b>0.0376</b>	0.0324	<b>0.0206</b>	0.0307	<b>0.0219</b>
12	<b>0.0365</b>	0.0371	0.0383	<b>0.0211</b>	0.0329	<b>0.0224</b>
13	0.0395	<b>0.0382</b>	0.0375	<b>0.0215</b>	0.0293	<b>0.0236</b>
14	0.0414	<b>0.0394</b>	0.0266	<b>0.0227</b>	0.0319	<b>0.0229</b>
15	0.0371	<b>0.0359</b>	0.0412	<b>0.0220</b>	0.0288	<b>0.0217</b>
16	0.0410	<b>0.0368</b>	0.0374	<b>0.0212</b>	0.0285	<b>0.0214</b>
17	0.0394	<b>0.0354</b>	0.0268	<b>0.0193</b>	0.0290	<b>0.0213</b>
18	0.0378	<b>0.0377</b>	0.0336	<b>0.0219</b>	0.0302	<b>0.0242</b>
19	0.0365	<b>0.0392</b>	0.0272	<b>0.0199</b>	0.0305	<b>0.0216</b>
20	0.0450	<b>0.0415</b>	0.0252	<b>0.0184</b>	0.0331	<b>0.0233</b>

Tabela 2- RMSEs do Sinais Simulados

A partir dessa análise, realizou-se um teste estatístico para verificar se o método WATV é melhor que o método ODBPD para os sinais utilizados. Primeiramente verificou-se a premissa de normalidade dos resíduos realizando o teste estatístico de Shapiro-Wilk. Como mostrado na Tabela 3, todos os p-valor são superiores a 0,05 e portanto, a premissa de normalidade dos resíduos é atendida em todos os casos.

Tipos de Ruído	TESTE DE SHAPIRO-WILK		
	Gaussiano	Impulsivo	Harmônico
p-valor	0.6392	0.2203	0.4989

Tabela 3- Teste Shapiro-Wilk

Visto que os dados apresentam uma distribuição normal, aplicou-se então o Teste



T pareado unilateral, para verificar se o método WATV é estatisticamente equivalente ou melhor que o método ODBPD. Como apresentado na Tabela 4, os p-valor de 4.284 e-05, 7.437e-11 e 2.305e-09 comprovam que para todos os sinais testados o método WATV apresenta-se mais eficiente que o método ODBPD.

Tipos de Ruído	TESTE T PAREADO		
	Gaussiano	Impulsivo	Harmônico
p-valor	4.284 e-05	7.437 e-11	2.305 e-09

Tabela 4- Teste T Pareado

## 7 | CONCLUSÕES

Este trabalho apresentou a aplicação do método de Variação Total com *Wavelets* utilizando o algoritmo SALSA para eliminação de ruídos gaussianos, harmônicos e impulsivos em sinais de descargas parciais sintéticos e gerados em laboratório.

Entre os ruídos apresentados, os impulsivos são mais complexos de serem extraídos devido a sua similaridade com as próprias descargas parciais. Com os resultados apresentados, e após comparação estatística com método da literatura, conclui-se que o método WATV obteve expressivos níveis de eliminação de ruídos, sendo mais eficiente que o método ODBPD.

Em trabalhos futuros pretende-se realizar outras métricas de avaliação de filtragem, comparar o WATV a outros métodos da literatura e utilizar outros sinais sintéticos e medidos.

## REFERÊNCIAS

AFONSO, Many V.; BIOUCAS-DIAS, José M.; FIGUEIREDO, Mário AT. Fast image recovery using variable splitting and constrained optimization. **IEEE transactions on image processing**, v. 19, n. 9, p. 2345-2356, 2010.

BOGGS, Steven A. Partial discharge: overview and signal generation. **IEEE Electrical Insulation Magazine**, v. 6, n. 4, p. 33-39, 1990.

CICONE, Antonio; LIU, Jingfang; ZHOU, Haomin. Adaptive local iterative filtering for signal decomposition and instantaneous frequency analysis. **Applied and Computational Harmonic Analysis**, v. 41, n. 2, p. 384-411, 2016.

COIFMAN, Ronald R.; DONOHO, David L. Translation-invariant de-noising. In: **Wavelets and statistics**. Springer, New York, NY, 1995. p. 125-150.

CONDAT, Laurent. A direct algorithm for 1-D total variation denoising. **IEEE Signal Processing Letters**, v. 20, n. 11, p. 1054-1057, 2013.

CUENCA, W. M.; SEIXAS, J. M.; LEVY, A. F. Análise de Componentes Principais para Identificar Descargas Parciais em Transformadores de Potência. In: **Brazilian Symposium on Neural Nets, Rio de Janeiro, Brazil**. 2004.

DE AZEVEDO SILVA, Fernando Thomé; DE OLIVEIRA MOTA, Hilton. Partial discharge signal

- processing using overcomplete dictionaries and sparse representations. In: **2017 IEEE Electrical Insulation Conference (EIC)**. IEEE, 2017. p. 388-391.
- DE OLIVEIRA MOTA, Hilton et al. Partial discharge signal denoising with spatially adaptive wavelet thresholding and support vector machines. **Electric Power Systems Research**, v. 81, n. 2, p. 644-659, 2011.
- DING, Yin; SELESNICK, Ivan W. Artifact-free wavelet denoising: non-convex sparse regularization, convex optimization. **IEEE signal processing letters**, v. 22, n. 9, p. 1364-1368, 2015.
- DONOHO, David L. De-noising by soft-thresholding. **IEEE transactions on information theory**, v. 41, n. 3, p. 613-627, 1995.
- FIGUEIREDO, Mario AT et al. On total variation denoising: A new majorization-minimization algorithm and an experimental comparison with wavelet denoising. In: **2006 International Conference on Image Processing**. IEEE, 2006. p. 2633-2636.
- MA, X. D.; ZHOU, C.; KEMP, I. J. DSP based partial discharge characterisation by wavelet analysis. In: **Proceedings ISDEIV. 19th International Symposium on Discharges and Electrical Insulation in Vacuum (Cat. No. 00CH37041)**. IEEE, 2000. p. 780-783.
- MA, X.; ZHOU, C.; KEMP, I. J. Interpretation of wavelet analysis and its application in partial discharge detection. **IEEE Transactions on Dielectrics and Electrical Insulation**, v. 9, n. 3, p. 446-457, 2002.
- MITICHE, I. et al. An evaluation of total variation signal denoising methods for partial discharge signals. In: **2017 INSUCON-13th International Electrical Insulation Conference (INSUCON)**. IEEE, 2017. p. 1-5.
- MOTA, Hilton de O.; VASCONCELOS, Flávio H. A partial discharge data acquisition system based on programmable digital oscilloscopes. In: **IMTC 2001. Proceedings of the 18th IEEE Instrumentation and Measurement Technology Conference. Rediscovering Measurement in the Age of Informatics (Cat. No. 01CH 37188)**. IEEE, 2001. p. 994-999.
- MOTA, Hilton de O.; VASCONCELOS, Flávio H.; DE CASTRO, Cristiano L. A comparison of cycle spinning versus stationary wavelet transform for the extraction of features of partial discharge signals. **IEEE Transactions on Dielectrics and Electrical Insulation**, v. 23, n. 2, p. 1106-1118, 2016.
- NIKOLOVA, Mila; NG, Michael K.; TAM, Chi-Pan. Fast nonconvex nonsmooth minimization methods for image restoration and reconstruction. **IEEE Transactions on Image Processing**, v. 19, n. 12, p. 3073-3088, 2010.
- RUDIN, Leonid I.; OSHER, Stanley; FATEMI, Emad. Nonlinear total variation based noise removal algorithms. **Physica D: nonlinear phenomena**, v. 60, n. 1-4, p. 259-268, 1992.
- STRONG, David; CHAN, Tony. Edge-preserving and scale-dependent properties of total variation regularization. **Inverse problems**, v. 19, n. 6, p. S165, 2003.

## REDUÇÃO DE DIMENSÕES USANDO TRANSFORMADA DE KARHUNEN–LOÈVE EM SISTEMAS MIMO MASSIVO DISTRIBUÍDO COM FRONTHAUL LIMITADO

**Ricardo de Souza Cerqueira**

Departamento de Engenharia Elétrica,  
Universidade de Brasília  
Brasília–DF

**André Noll Barreto**

Departamento de Engenharia Elétrica,  
Universidade de Brasília  
Brasília–DF

**RESUMO:** MIMO massivo distribuído é amplamente considerado como uma das principais tecnologias para sistemas de comunicação sem fio 5G. No entanto, um dos aspectos limitantes desta técnica é a capacidade dos enlaces de *fronthaul*, que transmitem a informação de cada arranjo de antenas a uma unidade de processamento centralizado. Este artigo propõe o uso da transformada Karhunen–Loève para reduzir o número de dimensões do sinal recebido no *uplink* com múltiplas antenas, e assim, reduzir os requisitos de taxa de dados do *fronthaul* com pequena perda no desempenho da transmissão.

**PALAVRAS-CHAVE:** Transformada de Karhunen–Loève, KLT, MIMO massivo distribuído, *fronthaul* limitado.

**ABSTRACT:** Distributed Massive MIMO is widely regarded as one of the key technologies

for 5G wireless communications systems. However, one of the limiting aspects of this technique is the capacity of the fronthaul links that convey the information from each individual antenna array to a centralized processing unit. In this paper, we propose the use of the Karhunen–Loève transform to reduce the number of dimensions of the received multi-antenna uplink signal, and, thus, reduce the fronthaul data-rate requirements, with little loss in the performance. **KEYWORDS:** Karhunen–Loève Transform, KLT, Distributed Massive MIMO, Limited Fronthaul.

### 1 | INTRODUÇÃO

Dentre as várias tecnologias atualmente consideradas para redes 5G, o *massive multiple-input multiple-output* (MIMO), ou MIMO massivo, é amplamente considerado como um dos mais promissores, permitindo aumentos substanciais de alcance e de taxa de dados (BOCCARDI *et al.*, 2014).

Sistemas *massive* MIMO são arranjos com grande número de antenas que têm, pelo menos, várias dezenas ou centenas de antenas. Estas antenas podem também ser distribuídas em diferentes painéis em locais diferentes, chamados de *remote radio heads* (RRH), com cada painel ligado a uma unidade central de processamento, ou *base band unit*

(BBU) através de um enlace de *fronthaul*, que pode ser implementado tanto com fibras ópticas, como com enlaces sem fio, seja por radiofrequência (RF) ou *free-space optics* (FSO), conforme ilustrado na Figura 1.

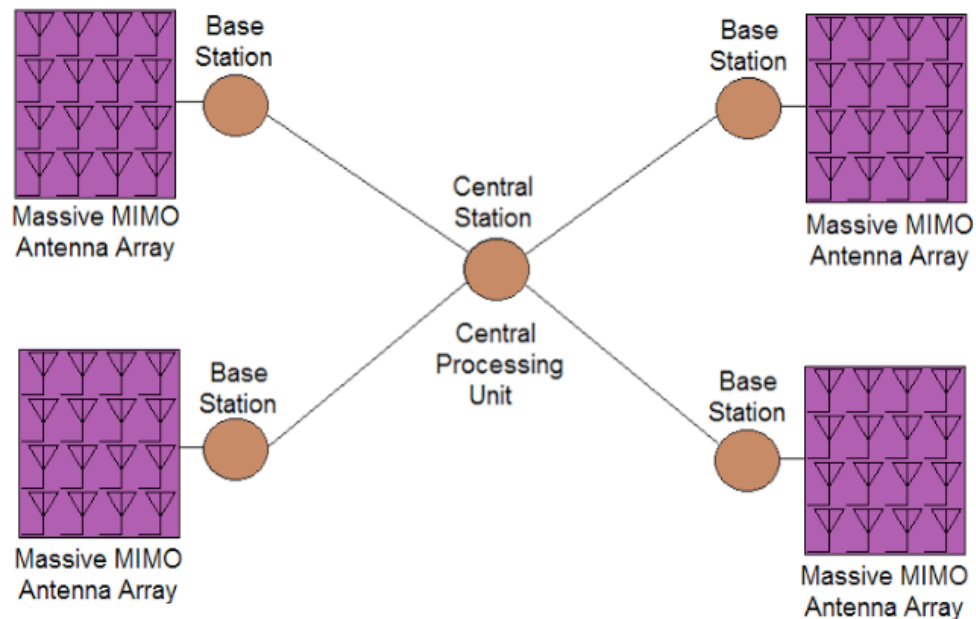


Figura 1: Topologia de um sistema MIMO massivo distribuído.

Tais sistemas são conhecidos como sistemas de antenas distribuídas (*Distributed Antenna Systems – DAS*) (BOCCARDI *et al.*, 2014), (ANDREWS *et al.*, 2014), e proporcionam não só multiplexação espacial e micro-diversidade, mas também ganhos de macro-diversidade, já que as respostas de canal de diferentes antenas têm diferentes fatores de somreamento e são tipicamente descorrelacionadas.

No entanto, a maioria dos artigos sobre MIMO distribuído considera um *fronthaul* ideal, sem limitações de capacidade (WANG *et al.*, 2013), mas o *fronthaul* com limitações é um problema fundamental nas redes sem fio futuras (MARSCH; FETTWEIS, 2007), particularmente se o número de antenas em cada painel for grande.

Nós investigamos o *uplink* de um DAS para um usuário MIMO, em que um equipamento de usuário (*user equipment – UE*), transmite feixes espaciais para um conjunto de painéis de antenas, que por sua vez são conectados a uma unidade central de processamento por enlaces *fronthaul*. Devido às limitações de capacidade, esses enlaces podem não ser capazes de transmitir todas as informações de cada antena para a unidade de processamento, e esta informação deve ser comprimida em cada painel antes de ser enviada. Neste contexto, alguns autores propuseram diferentes técnicas de compressão, por exemplo, o uso da compressão distribuída com perdas Wyner–Ziv (DEL COSO; SIMOENS, 2009), (ZHOU; YU, 2014) para lidar com o *fronthaul* com restrições de capacidade.

Nesta contribuição, propomos a utilização de uma técnica linear, a transformada Karhunen–Loève (KLT), que realiza uma fatorização da matriz para reduzir a quantidade

de informação a ser transmitida nos enlaces de *fronthaul*. A KLT baseia-se na matriz de covariância da antena de longo prazo, na qual separa o sinal em componentes descorrelacionados.

O uso da KLT em sistemas de múltiplas antenas foi sugerido em (JELITTO; FETTWEIS, 2002), na qual propõe que alguns dos componentes de maior energia sejam selecionados após a KLT para redução de dimensões do sinal. Na literatura, a KLT também foi empregada em um sistema MIMO (SIM *et al.*, 2014), mas para reduzir a informação de estado de canal (*channel state information* – CSI), no canal de retorno e não nos dados em si.

Sugerimos a extensão do proposto em (JELITTO; FETTWEIS, 2002) a um sistema com antenas distribuídas, em que uma KLT é realizada em cada arranjo e somente os componentes de maior potência são enviados ao nó central, pois consideramos uma restrição na capacidade do enlace do *fronthaul*. Isto é semelhante à KLT distribuída proposto em (GASTPAR; DRAGOTTI; VETTERLI, 2006), mas que foi aplicada a outras tarefas de processamento de sinal.

## 2 | MODELO DO SISTEMA

Neste trabalho consideramos o *uplink* de um sistema sem fio em um cenário de propagação urbano com multipercursos na qual os UEs transmitem por meio de painéis com grande número de antenas. Consideramos um cenário de usuário único, em que o UE está localizado no ponto médio entre dois painéis de antenas, distante 100 m de cada painel. Também foi considerado que cada UE possui 4 antenas de transmissão, e que cada painel de antenas tem 25 antenas de recepção, em um arranjo linear de antenas. Estas configurações podem ser facilmente modificadas para outros cenários com diferentes números de painéis de antenas e usuários, ou para outros modelos de propagação, distâncias, configurações de antena e formatos de painel.

Conforme mostrado na Figura 2 cada painel realiza uma compressão com perdas no sinal multidimensional recebido e transmite esta informação reduzida a um servidor central, que tenta detectar o sinal transmitido baseado na combinação da informação, com perdas, de todos os painéis.

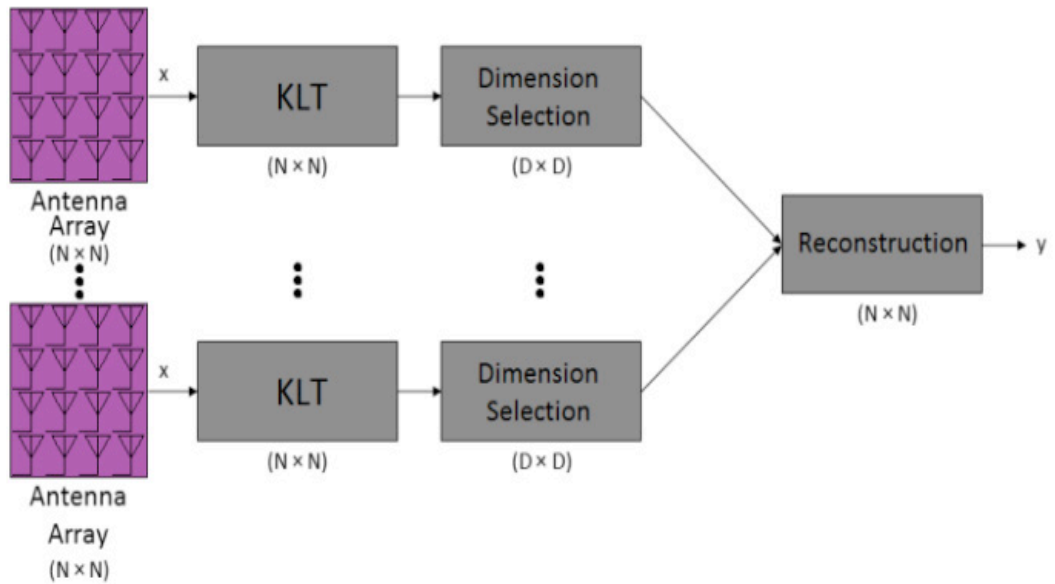


Figura 2: Redução de dimensões no *massive* MIMO distribuído.

Como as simulações são relacionadas com o desempenho na camada de enlace, o modelo de perda por percurso não é crítico, e foram consideradas somente as perdas de percurso por espaço livre (*free-space path loss* – FSPL), baseada na fórmula de Friis (FRIIS, 1946)

$$FSPL = \left(\frac{4\pi d}{\lambda}\right)^2 \quad (1)$$

Onde  $d$  é a distância entre o transmissor e o receptor, e  $\lambda$  é o comprimento de onda. Foi considerada a frequência portadora de 60 GHz.

A cada intervalo de símbolo foi considerada a transmissão de um vetor  $\mathbf{x}$ , que consiste em  $M$  diferentes símbolos QAM/PSK, mapeados em  $M$  diferentes antenas de transmissão e feixes espaciais.

No lado do receptor, tem-se  $k$  sinais recebidos, um para cada um dos  $K$  painéis de antenas, que são definidos como:

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x} + \mathbf{n}_k \quad (2)$$

Onde  $\mathbf{y}_k$  é o vetor do sinal recebido pelo painel  $k$ ,  $\mathbf{H}_k$  é a matriz de canal com dimensões  $N \times M$ , sendo  $N$  o número de antenas de recepção em cada painel,  $M$  o número de antenas de transmissão,  $\mathbf{x}$  é o vetor do sinal transmitido, e  $\mathbf{n}_k$  é o vetor de ruído gaussiano branco aditivo (AWGN).

Os sinais dos  $K$  painéis são enviados ao servidor central, onde os sinais recebidos por cada painel são concatenados em um vetor global:

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_0 \\ \mathbf{y}_1 \\ \vdots \\ \mathbf{y}_{k-1} \end{bmatrix} = \mathbf{H}\mathbf{x} + \mathbf{n} = \begin{bmatrix} \mathbf{H}_0 \\ \mathbf{H}_1 \\ \vdots \\ \mathbf{H}_{k-1} \end{bmatrix} \mathbf{x} + \begin{bmatrix} \mathbf{n}_0 \\ \mathbf{n}_1 \\ \vdots \\ \mathbf{n}_{k-1} \end{bmatrix} \quad (3)$$

Foi considerado inicialmente um cenário sem linha de visada (*non-line-of-sight* – NLOS), em que a matriz do canal é modelada como variáveis aleatórias de desvanecimento Rayleigh. Também consideramos que as matrizes de canal em diferentes painéis de antenas são descorrelacionadas. Porém, especialmente considerando o grande número de antenas em um painel de tamanho limitado, é muito provável que o sinal em diferentes elementos de antena sejam correlacionados.

Em cenários NLOS típicos e para tamanhos de matriz moderados, a matriz de covariância do canal  $\mathbf{R}_H$  pode ser bem aproximada pelo produto de Kronecker das matrizes de covariância no lado do transmissor e do receptor, respectivamente (YU, 2005):

$$\mathbf{R}_H = \mathbf{R}_{Tx} \otimes \mathbf{R}_{Rx} \quad (4)$$

Onde  $\otimes$  denota o produto de Kronecker,  $\mathbf{R}_{Tx}$  é a matriz de covariância do canal lado do transmissor, com dimensões  $M \times M$ , e  $\mathbf{R}_{Rx}$  é a matriz de covariância do canal do lado receptor, com dimensões  $N \times N$ .

Considerando as matrizes de correlação de antena dadas, a correlação de antena da equação (4) pode ser modelada na matriz de canal de cada painel, fazendo (YU *et al.*, 2002):

$$\mathbf{H}_k = \mathbf{R}_{Rx}^{1/2} \mathbf{H}_{w_k} \left( \mathbf{R}_{Tx}^{1/2} \right)^T \quad (5)$$

Onde  $\mathbf{H}_{w_k}$  é uma matriz de canal com elementos independentes de desvanecimento Rayleigh, com dimensões  $N \times M$ , e  $(\cdot)^T$  é a matriz transposta.

## 2.1 MATRIZ DE COVARIÂNCIA DA ANTENA

Considerando que as antenas vizinhas provavelmente têm uma grande correlação, tanto as matrizes de covariância da antena de recepção e de transmissão,  $\mathbf{R}_{Rx}$  e  $\mathbf{R}_{Tx}$ , foram implementadas fazendo-se  $R_{ij} = \alpha^{li-jl}$ , onde  $i$  e  $j$  são os índices da linha e da coluna, respectivamente, e conseqüentemente,  $li-jl$  representa a distância da antena, ou seja, a correlação da antena diminui com a distância.

Esta abordagem também foi feita em (GASTPAR; DRAGOTTI; VETTERLI, 2006), resultando na seguinte matriz covariância circulante de recepção:



$$\mathbf{R}_{\text{Rx}} = \begin{bmatrix} \alpha^0 & \alpha^1 & \alpha^2 & \dots & \alpha^{N-1} \\ \alpha^1 & \alpha^0 & \alpha^1 & \alpha^2 & \vdots \\ \alpha^2 & \alpha^1 & \alpha^0 & \alpha^1 & \alpha^2 \\ \vdots & \alpha^2 & \alpha^1 & \ddots & \alpha^1 \\ \alpha^{N-1} & \dots & \alpha^2 & \alpha^1 & \alpha^0 \end{bmatrix} \quad (6)$$

O nível de correlação depende do valor do parâmetro  $\alpha$ . Por exemplo, temos correlação total entre os sinais recebidos quando  $\alpha = 1$ , e sinais completamente descorrelacionados quando  $\alpha = 0$ .

A matriz de correlação de transmissão  $\mathbf{R}_{\text{Tx}}$  é criada da mesma forma que do lado de recepção, porém com dimensões  $M \times M$ .

### 3 | REDUÇÃO DE DIMENSÕES

Considerando o grande número de antenas receptoras e limitações de espaço nos painéis é muito provável que os sinais recebidos sejam altamente correlacionadas entre antenas diferentes. Isto é o que permite o uso de técnicas de redução de dimensões com pouca perda de informação. Neste artigo, foi empregada a transformada Karhunen–Loève para comprimir o sinal a partir da redução do número de dimensões que contém informações relevantes, assumindo o perfeito conhecimento da matriz de covariância da antena de recepção. Como se trata de uma métrica estatística de longo prazo aplicado a um receptor fixo, pode-se considerar que esta seja uma suposição razoável.

Adotamos o termo transformada Karhunen–Loève em vez de decomposição em valores singulares (*single value decomposition* – SVD). Ambos são intercambiáveis e usados na literatura. Em (GERBRANDS, 1981) tanto a KLT e a análise de componentes principais, ou *principal component analysis* (PCA), são descritos e pode-se ver que, no cenário de desvanecimento Rayleigh considerado neste trabalho, ambas as abordagens são equivalentes já que os ganhos de canal têm uma média nula.

Após a execução da KLT é realizada uma seleção dos  $D$  componentes mais fortes, cuja quantidade de componentes selecionados depende da capacidade dos enlaces de *fronthaul*. Isto equivale a uma compressão com perdas, já que é descartada informação dos componentes mais fracos, mas, por outro lado, também é descartado o ruído contido nestes componentes (JELITTO; FETTWEIS, 2002).

Como foi considerado um sistema com painéis de antenas distribuídas, foi utilizada a KLT distribuída (GASTPAR; DRAGOTTI; VETTERLI, 2006). Ela difere da KLT pois os dados são divididos em subgrupos. A KLT é realizada em cada subconjunto individual, e, em seguida, o sinal é reconstruído em uma unidade de processamento central pela união de cada subconjunto processado. Neste trabalho, os subconjuntos representam os diferentes painéis de antenas, tal como mostrado na Figura 2.

A KLT é realizada com base na decomposição em valores singulares da matriz

de covariância, cuja saída são valores singulares em ordem decrescente de energia e seus correspondentes vetores singulares. Se a SVD for aplicada em uma matriz de covariância  $\mathbf{R}$ , os vetores singulares esquerdos e direitos serão os mesmos. No caso deste trabalho, a matriz covariância das antenas de recepção foi decomposta, tal que:

$$\mathbf{R}_{\text{Rx}} = \mathbf{U} \mathbf{\Lambda} \mathbf{U}^{\text{H}} \quad (7)$$

Onde  $\mathbf{U}$  é uma matriz unitária com dimensões  $N \times N$ , cujas colunas representam os vetores singulares esquerdos,  $\mathbf{\Lambda}$  é uma matriz diagonal com valores singulares  $\lambda_1, \lambda_2, \dots, \lambda_N$  em ordem decrescente, e  $(\cdot)^{\text{H}}$  é transposta Hermitiana.

Com o conhecimento da matriz de covariância, a KLT efetua uma transformação linear do sinal de entrada, de tal modo que as saídas são descorrelacionadas. Isto é feito pela aplicação da seguinte transformação em cada painel antena:

$$\mathbf{z}_k = \mathbf{U}^{\text{H}} \mathbf{y}_k \quad (8)$$

Resultando em um vetor  $\mathbf{z}_k$  de  $N$  sinais descorrelacionados, com valores em ordem decrescente de potência e correspondentes aos valores singulares  $\lambda_j$ .

Dependendo do nível de correlação, a potência estará concentrada em alguns poucos componentes e pode-se realizar a seleção de dimensões pela remoção de alguns destes componentes que não têm contribuição relevante para a reconstrução do sinal. Enquanto o número de dimensões for maior ou igual ao número de feixes espaciais, a detecção dos sinais espacialmente multiplexados ainda será possível.

Na etapa de seleção de dimensões o tamanho do sinal é reduzido de  $N$  para  $D$  dimensões, onde  $D \leq N$ , conforme mostrado na Figura 2.

Para representar essa seleção de dimensões é necessário criar uma matriz de transformação de redução:

$$\mathbf{T} = [\mathbf{I}_D | \mathbf{0}_{[D \times (N-D)]}]_{[D \times N]} \quad (9)$$

Onde  $\mathbf{I}_D$  é uma matriz identidade com dimensões  $D \times D$  e  $\mathbf{0}$  é uma matriz nula com dimensões  $D \times (N-D)$ .

A seleção de dimensões pode ser representada como:

$$\mathbf{y}_{\text{KLT},k} = \mathbf{T} \mathbf{z}_k \quad (10)$$

Esta redução de dimensões permite a diminuição da capacidade dos enlaces de *fronthaul* entre os painéis de antenas e a unidade de processamento central. Ela também ajuda a reduzir o processamento na estação central, ou *central station* (CS), porque o número de dimensões é significativamente menor do que o sinal original,

porém, ao custo da execução da KLT nos painéis de antenas.

Pela substituição da equação (8) na equação (10), temos:

$$\mathbf{y}_{\text{KLT},k} = \mathbf{T}\mathbf{U}^H\mathbf{y}_k \quad (11)$$

E, substituindo a equação (2) na equação (11), temos:

$$\mathbf{y}_{\text{KLT},k} = \mathbf{T}\mathbf{U}^H\mathbf{H}_k\mathbf{x} + \mathbf{T}\mathbf{U}^H\mathbf{n}_k = \mathbf{H}_{\text{KLT},k}\mathbf{x} + \mathbf{n}_{\text{KLT},k} \quad (12)$$

Onde  $\mathbf{H}_{\text{KLT},k}$  é a matriz de canal com dimensões reduzidas pela KLT e  $\mathbf{n}_{\text{KLT},k}$  é o vetor de ruído após a transformação, que, em razão de  $\mathbf{U}$  ser unitário, tem amostras de ruído com a mesma potência do vetor de ruído original  $\mathbf{n}$ .

#### 4 | DETECÇÃO DE SINAL

Após reduzirmos as dimensões, todos os sinais com dimensões reduzidas são concatenados na unidade central de processamento, tal que, em vez da equação (3), tem-se o sinal composto que representa um canal MIMO para D antenas receptoras decorrelacionadas.

$$\begin{aligned} \mathbf{y}_{\text{KLT}} &= \begin{bmatrix} \mathbf{y}_{\text{KLT},0} \\ \mathbf{y}_{\text{KLT},1} \\ \vdots \\ \mathbf{y}_{\text{KLT},k-1} \end{bmatrix} = \begin{bmatrix} \mathbf{H}_{\text{KLT},0} \\ \mathbf{H}_{\text{KLT},1} \\ \vdots \\ \mathbf{H}_{\text{KLT},k-1} \end{bmatrix} \mathbf{x} + \begin{bmatrix} \mathbf{n}_{\text{KLT},0} \\ \mathbf{n}_{\text{KLT},1} \\ \vdots \\ \mathbf{n}_{\text{KLT},k-1} \end{bmatrix} = \\ &= \mathbf{H}_{\text{KLT}}\mathbf{x} + \mathbf{n}_{\text{KLT}} \end{aligned} \quad (13)$$

O sinal transmitido pode então ser detectado usando-se esquemas padrão de detecção de multiplexação espacial. Nesta contribuição foi considerada apenas a detecção linear, isto é, o sinal de transmissão pode ser estimado como:

$$\hat{\mathbf{x}} = \mathbf{A}\mathbf{y}_{\text{KLT}} \quad (14)$$

Neste estudo, foi investigado o equalizador Zero Forcing (ZF), tal que:

$$\hat{\mathbf{x}}_{\text{ZF}} = \mathbf{H}_{\text{KLT}}^\dagger \mathbf{z}_{\text{KLT}} \quad (15)$$

Onde  $\hat{\mathbf{x}}_{\text{ZF}}$  é o sinal estimado pelo método ZF e  $(\cdot)^\dagger$  é a pseudo-inversa de Moore–Penrose.

Também foi considerado o método minimum mean square error (MMSE) que

pode ser calculado como:

$$\hat{\mathbf{x}}_{\text{MMSE}} = \left[ \left( \mathbf{H}_{\text{KLT}}^{-1} \mathbf{H}_{\text{KLT}} + \frac{1}{E_s/N_0} \right)^{-1} \mathbf{H}_{\text{KLT}}^{-1} \right] \mathbf{z}_{\text{KLT}} \quad (16)$$

Onde  $\hat{\mathbf{x}}_{\text{MMSE}}$  é o sinal estimado pelo método MMSE,  $E_s/N_0$  é a razão entre a energia de símbolo e a densidade espectral da potência do ruído, e  $(\cdot)^{-1}$  é a matriz inversa.

## 5 | RESULTADOS DAS SIMULAÇÕES

A fim de avaliar o desempenho da redução de dimensões de sistemas MIMO distribuído foram realizadas simulações tanto sem compressão, quanto com redução de dimensões baseadas na KLT. Simulações também foram realizadas com painéis lineares de antenas e com painéis de antenas quadrados, em um arranjo 5×5, na recepção.

Nos resultados apresentados neste trabalho, foi considerado um *layout* com um UE contendo 4 antenas de transmissão, e dois painéis de antenas igualmente distantes do UE, cada um com 25 antenas receptoras. Isto representa um sistema equivalente com 4 antenas de transmissão e um total de 50 antenas de recepção.

Nas simulações foi considerado um cenário NLOS com desvanecimento Rayleigh e AWGN, multiplexação espacial e modulação BPSK, para os métodos de estimação ZF e MMSE.

Foram realizadas diversas simulações utilizando diferentes parâmetros covariância  $\alpha$ , com valores de  $\alpha = 1/4$ ,  $\alpha = 1/2$  e  $\alpha = 3/4$ , tanto para painéis lineares quanto para painéis quadrados. Nas Figuras 3 a 5 são apresentados os resultados de simulações de um cenário com um UE transmissor e dois painéis lineares de antenas.

As curvas de taxa de erro de bit, ou *bit error rate* (BER), para um canal com o parâmetro covariância  $\alpha = 1/4$  são mostradas na Figura 3 e, como esperado, o desempenho se deteriora quando se reduz o número de dimensões a ser transmitido pelos enlaces de *fronthaul*. No entanto, a deterioração do desempenho é aceitável, mesmo com uma redução razoavelmente grande, como no caso de 50 para 20 antenas, e que constitui em uma enorme redução comparada ao *fronthaul* sem compressão. No caso da redução de 50 para 30 dimensões, o desempenho é praticamente idêntico ao sistema sem redução.

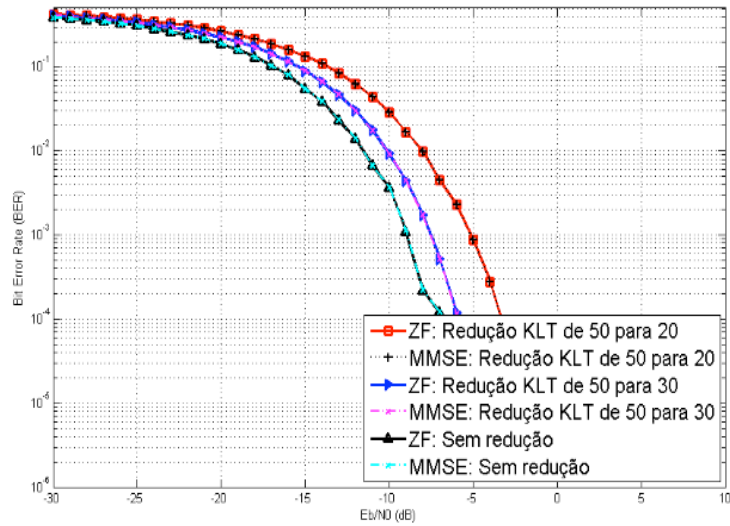


Figura 3: BER para arranjo linear com  $\alpha=1/4$ .

Nas Figuras 4 e 5 foram realizadas as mesmas simulações, para um arranjo linear de antenas, mas considerando um sistema com maior correlação entre as antenas, isto é,  $\alpha = 1/2$  e  $\alpha = 3/4$ .

As simulações, conforme apresentado nas Figuras 3 a 5, comprovam que, quanto menor o fator de correlação das antenas, melhor o desempenho, pois significa que há uma maior diversidade. Também pode ser observado que, como esperado, a redução de dimensões apresenta menores perdas de desempenho quando a correlação das antenas é alta, pois neste caso, há uma maior redundância, que resulta numa compressão mais eficaz.

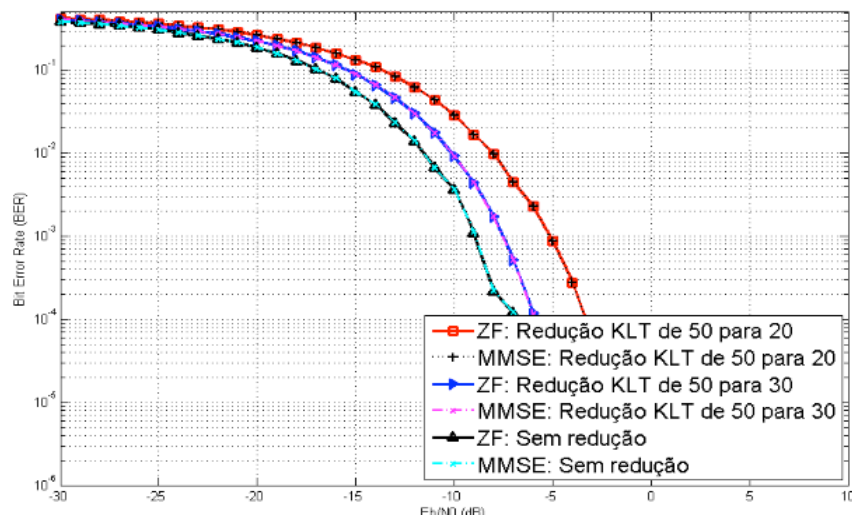


Figura 4: BER para arranjo linear com  $\alpha=1/2$ .

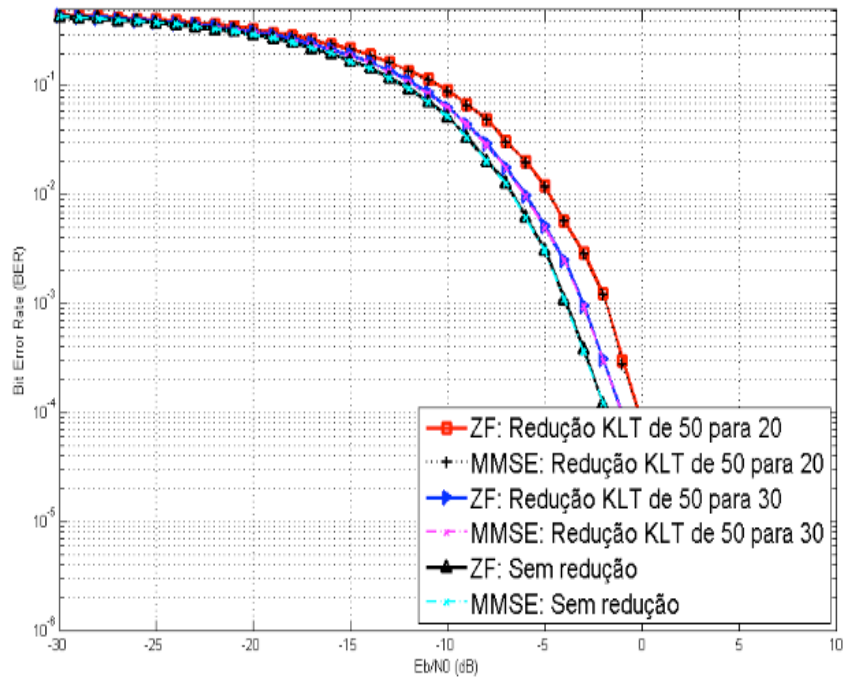


Figura 5: BER para arranjo linear com  $\alpha=3/4$ .

Da mesma forma que as simulações para um sistema com arranjo linear de antenas e diferentes covariâncias, as simulações para um arranjo quadrado de antenas, isto é dois painéis de 25 antenas de recepção em um arranjo  $5 \times 5$ , onde a matriz de correlação é modificada, pode-se verificar que o comportamento é semelhante, porém o painel quadrado de antenas é levemente melhor que o arranjo linear, conforme mostrado nas Figuras 6 a 8.

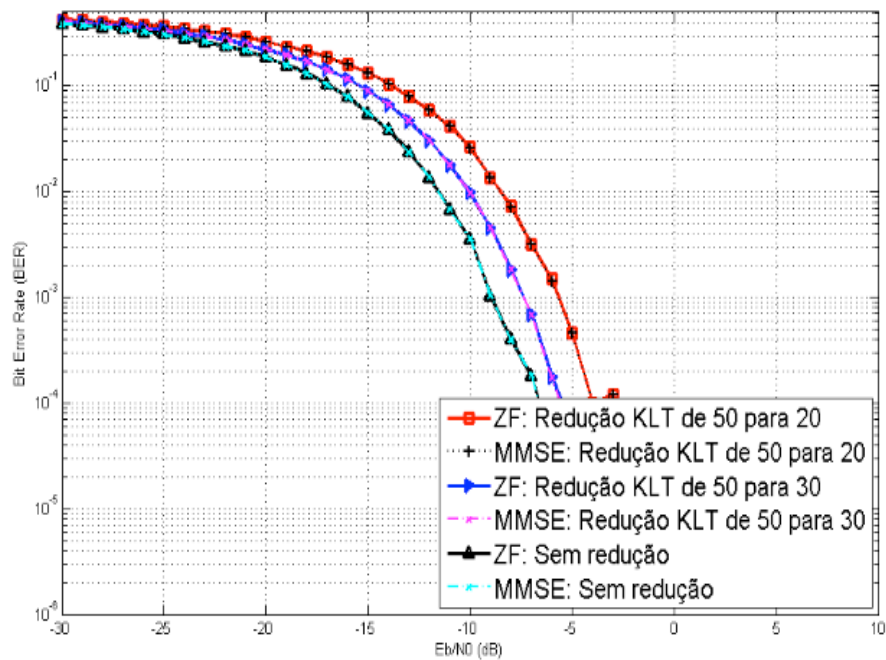


Figura 6: BER para painel quadrado com  $\alpha=1/4$ .

Assim como o sistema com arranjo linear, o arranjo com painel quadrado de antenas apresenta um desempenho um pouco melhor quando o fator de covariância



é menor.

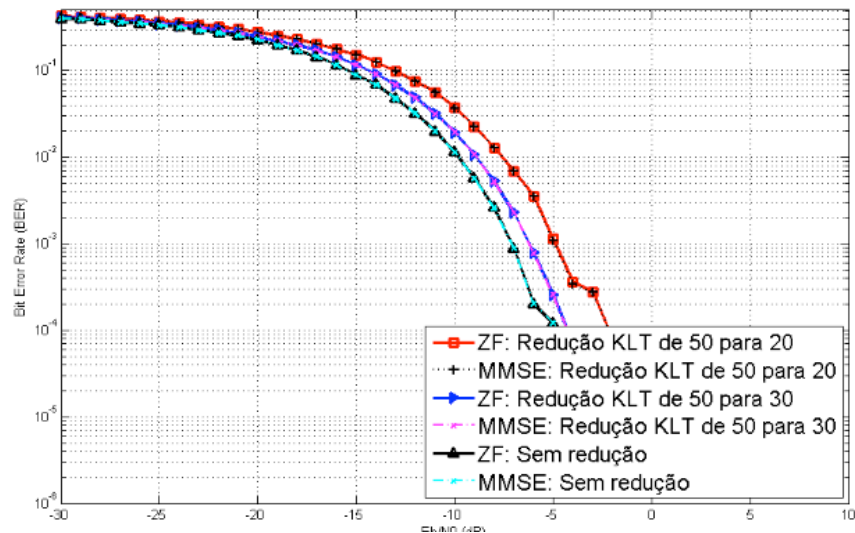


Figura 7: BER para painel quadrado com  $\alpha=1/2$ .

Quando a correlação das antenas é alta, assim como no arranjo linear, o desempenho do sistema é pior que um sistema com menor correlação.

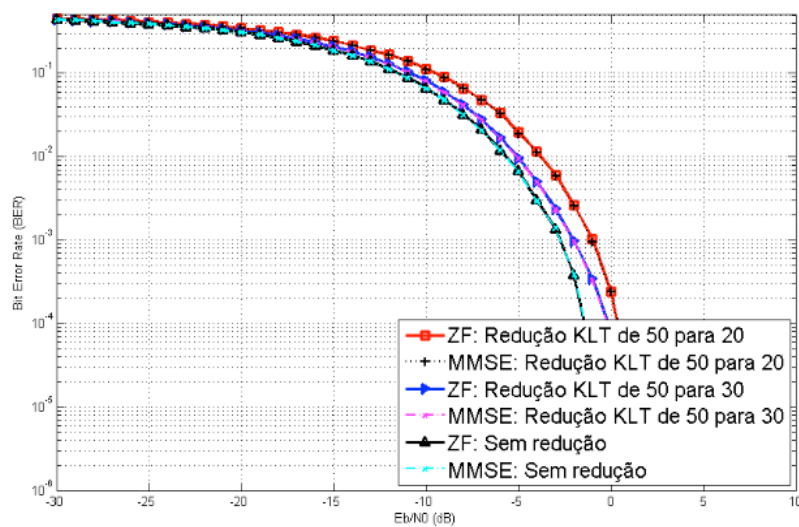


Figura 8: BER para painel quadrado com  $\alpha=3/4$ .

Ao se comparar três cenários com a mesma covariância ( $\alpha$ ), ou seja, sem redução, com redução de 50 para 30 dimensões e com redução de 50 para 20, pode-se verificar que o sistema de antenas distribuídas e compressão por KLT permite a obtenção de desempenho semelhante ao sistema sem compressão.

Como esperado, com mais antenas de recepção, a redução de dimensões pode selecionar os sinais mais fortes que chegam a cada painel e reduzir significativamente a quantidade de sinais requeridos, resultando na necessidade de um enlace de menor capacidade.

O número de dimensões que pode ser reduzido sem grande perda de desempenho dependerá da matriz de covariância de canal e da relação sinal-ruído, ou *signal-to-*



noise ratio (SNR), do sistema.

Na Figura 9 pode-se verificar o desempenho de cada arranjo linear com 25 antenas, pois mostra a BER para diferentes fatores de redução de dimensões para valores fixos, tanto de SNR, quanto de covariância. Neste cenário selecionado pode-se observar que o desempenho se deteriora quando se reduz o número de dimensões, e também que a partir de um número de dimensões, o desempenho é bruscamente melhorado. Este ponto é facilmente visualizado para SNRs altas e, para os cenários simulados, foi de 10 dimensões.

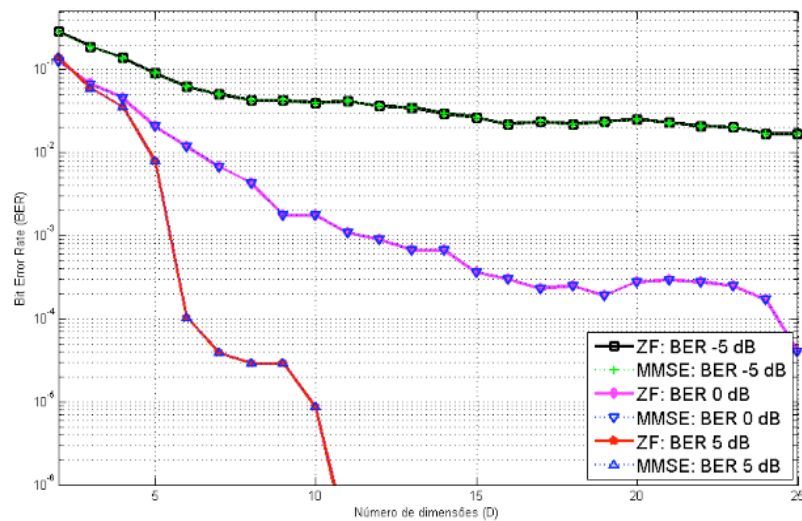


Figura 9: BER para arranjo linear de antenas com  $\alpha=1/2$  e diversos fatores de redução, para  $E_b/N_0$  fixos.

A Figura 10 apresenta o mesmo que a Figura 9, mas para um painel quadrado de antenas com arranjo  $5 \times 5$ . É possível verificar que a partir de 10 dimensões, assim como no arranjo linear, o desempenho é repentinamente melhorado.

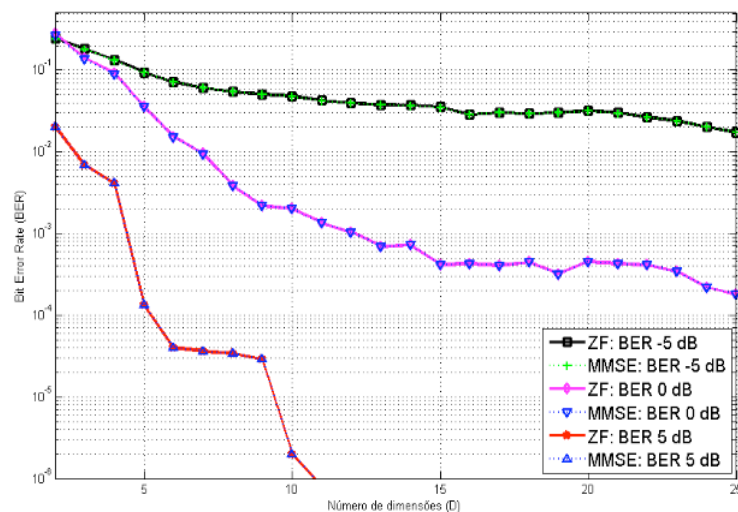


Figura 10: BER para painel de antenas com  $\alpha=1/2$  e diversos fatores de redução, para  $E_b/N_0$  fixos.

Caso se tenha um enlace de *fronthaul* com capacidade limitada, podemos

questionar se devemos simplesmente usar o número exato de antenas suportadas ou se devemos acrescentar um grande número de antenas e comprimir esse sinal. Tal situação é abordada nos resultados de simulação mostrados nas Figuras 9 e 10. Pode ser observado que a partir destes resultados uma redução de dimensões baseada na KLT permite um desempenho BER similar ao de um sistema sem compressão.

Em outras palavras, nestas simulações tem-se que a utilização da redução de dimensões baseada na KLT requer cerca de metade da taxa necessária para o mesmo enlace de *fronthaul* sem compressão.

A partir das simulações realizadas, pode-se questionar caso se tenha um enlace de *fronthaul* com restrição de capacidade, qual é o fator de redução de dimensão que se tem a melhor relação custo-benefício. Para estas simulações o fator de compressão se aproxima a metade da taxa original.

## 6 | CONCLUSÕES E ESTUDOS FUTUROS

Nesta contribuição verificou-se a viabilidade de um esquema de redução de dimensões em sistemas MIMO massivo distribuído por meio de uma transformação linear, a KLT, do sinal multidimensional recebido. Também foi observado que, dependendo da correlação entre as antenas, a compressão resulta em pouca perda de desempenho dos enlaces de *fronthaul*. No cenário simulado foi possível reduzir a cerca de metade da taxa original.

Esta abordagem pode ajudar na transmissão de dados relacionada a operação de *fronthaul* com limitação de capacidade, pois representa economia nas despesas operacionais das operadoras de telecomunicações.

Os resultados obtidos são promissores, mas entre diferentes questões a serem consideradas em trabalhos futuros, será necessário considerar um modelo de propagação mais realista para ondas milimétricas (mmWave) e uma matriz de antena covariância realista, de acordo com o ambiente considerado. Um modelo de propagação para linha de visada, ou *line-of-sight* (LOS), e NLOS, incluindo cenários *indoor* e *outdoor*, em mmWave e outras frequências, permitirá encontrar uma compreensão mais adequada sobre estas questões.

A estimativa da matriz de covariância também é uma questão importante a ser investigada, pois neste trabalho foi pressuposto o seu perfeito conhecimento. Será necessário executar medições e estabelecer as matrizes de covariância mais viáveis.

Este trabalho avança em uma técnica promissora que pode ter um papel fundamental no futuro da 5G, mas mais estudos e investigações ainda são necessários.

## REFERÊNCIAS

ANDREWS, J. *et al.*, **What will 5G be?**, IEEE J. Select. Areas Commun., vol. 32, n. 6, p. 1065-1082, Jun. 2014.

BOCCARDI, F. *et al.*, **Five disruptive technology directions for 5G**, IEEE Commun. Mag., vol. 52, p. 74-80, Fev. 2014.

DEL COSO, A.; SIMOENS, S., **Distributed compression for MIMO coordinated networks with a backhaul constraint**, IEEE Trans. on Wireless Commun., vol. 8, n. 9, p. 4698-4709, Set. 2009.

FRIIS, H. T., **A Note on a Simple Transmission Formula**, Proc. IRE, vol. 34, n. 5, p. 254-256, Maio 1946.

GASTPAR, M.; DRAGOTTI, P. L.; VETTERLI, M., **The distributed Karhunen–Loève transform**, IEEE Trans. on Inform. Theory, p. 5177-5196, Dez. 2006.

GERBRANDS, J. J., **On the relationships between SVD, KLT and PCA**, Pattern Recognition, vol. 14, n. 1-6, p. 375-381, 1981.

JELITTO, J.; FETTWEIS, G., **Reduced dimension space-time processing for multi-antenna wireless systems**, IEEE Wireless Commun., vol. 9, n. 6, p. 18-25, Dez. 2002.

MARSCH, P.; FETTWEIS, G., **A framework for optimizing the uplink performance of distributed antenna systems under a constrained backhaul**, Proc. IEEE Int. Conf. on Commun., p. 975-979, Jun. 2007.

SIM, M. S. *et al.*, **Compressed channel feedback for correlated massive MIMO systems**, IEEE Globecom Workshops, p. 327-332, Dez. 2014.

WANG, D. *et al.*, **Spectral Efficiency of Distributed MIMO Systems**, IEEE J. Select. Areas Commun., vol. 31, n. 10, p. 2112-2127, Out. 2013.

YU, K., **Multiple-input multiple-output radio propagation channels characteristics and models**, Ph.D. dissertation, Dept. of signals, sensors and systems, Royal Institute of Technology, Stockholm, Sweden, 2005.

YU, K. *et al.*, **A wideband statistical model for NLOS indoor MIMO channels**, IEEE Veh. Technol. Conf., p. 370-374, Maio 2002.

ZHOU, Y.; YU, W., **Optimized backhaul compression for uplink cloud radio access network**, IEEE J. Select. Areas Commun., vol. 32, n. 6, p. 1295-1307, Jun. 2014.

## WSN COVERAGE IMPROVEMENT WITH ROF IN BUS TOPOLOGY FOR SMART CITIES

### Raphael Montali da Assumpção

University of Campinas (UNICAMP), DSIF,  
Campinas-SP, Brazil

### Indayara Bertoldi Martins

Pontifícia Universidade Católica de Campinas,  
CEATEC, Campinas-SP, Brazil

### Frank Herman Behrens

Pontifícia Universidade Católica de Campinas,  
CEATEC, Campinas-SP, Brazil

### Omar Carvalho Branquinho

Pontifícia Universidade Católica de Campinas,  
CEATEC, Campinas-SP, Brazil

### Fabiano Fruett

University of Campinas (UNICAMP),  
DSIF, Campinas-SP, Brazil

**ABSTRACT:** This work proposes an experimental coverage evaluation of a Wireless Sensor Network (WSN) connected with Radio-over-Fiber (RoF) network operating as system backhaul that can be implemented in WSN under the Internet of Things (IoT) context. The RoF is assembled in a bus topology and has a radio repeater for signal distribution to the WSN. Impairments present in RoF systems that degrade the RF signal are analyzed. The main parameter used to evaluate the WSN coverage of this system is distance by Packet Error Rate analysis. The results show that the radio repeater device brings benefits to signal quality,

improving the area coverage of a WSN. The bus topology used for RoF reduces equipment cost.

**KEYWORDS:** WSN, RoF, Bus Topology, Smart Cities, IoT.

### 1 | INTRODUCTION

The Internet of Things (IoT) is a recent communication paradigm aimed at connecting objects of everyday life equipped with sensors, microcontrollers and transceivers for digital communication that make them able to communicate with one another bringing new data and functionalities to the users through the Internet (Zanella *et al.*, 2014).

The application of the IoT paradigm to an urban context is of particular interest nowadays, as it permits the adoption of the information and communications technology (ICT) solutions to be used in the management of public affairs, thus realizing the concept known as Smart City (Hall *et al.*, 2013) environmental and economic benefit of the e-City dweller as did the introduction of utility and transport network infrastructures. Yet without action, the quantity of greenhouse gas emissions attributable to the increasing energy consumption of access networks will become a serious threat to the environment. This paper introduces the vision of a 'sustainable Digital City' and then considers strategies to overcome

economic and technical hurdles faced by engineers responsible for developing the information and communications technology (ICT).

In 2008, over 50% of the world's population was already living in cities, and this number is expected to rise to 80% by 2020. Cities have become important economic units in the global economy, where innovation takes place. The structure of cities has been radically transformed by the recent introduction of ICT network infrastructure, an important component to leverage cities into the digital age and e-economy. Ambient intelligence is embedded in the fabric of the city, wherein various sensors are widely and densely deployed to gather all sorts of information (traffic, pollution, temperature, etc.) which is then transported by wireless and wired access networks to data centers for analysis in order to optimize the use of resources (Zanella *et al.*, 2014).

Wireless sensor network (WSN) is a great area of applications currently under development. The monitoring and control of process in the context of Internet of Things (IoT) – where everything is connected – is becoming a great business for wireless communication. A WSN can be described as a network composed of sensor nodes that are highly distributed, small, and lightweight, that wirelessly monitors the environment or system through physical measurements. The topology of a WSN system design depends significantly on the application (Assumpção e Branquinho, 2012).

A known problem with WSN is its short-range coverage under low-power consumption restriction. In some applications, the sensor nodes need to operate in low power mode to preserve battery charge and extend its operating lifetime, thus limiting the WSN coverage by its restricted transmission power when sending monitored data. However, this limitation can be mitigated by using an optical fiber communication network as system backhaul (Andrade, de *et al.*, 2012; Assumpção e Branquinho, 2012). With the Radio over Fiber (RoF) backhaul, it is possible to increase the distance between a WSN cluster, comprised of several sensor nodes and a local coordinator node, and the Management Center by several kilometers if needed (Hossen e Jang, 2009) if only wireless ad-hoc network is used in large sensor network such as u-City, it requires too many hops causing problems like address assignment and time delay. Above problems are solved when the u-City is divided into multiple clusters, which are then connected with the personal area network coordinator (PANC. It is noteworthy that the path loss in free space is orders of magnitude higher than in the optical domain on a RoF based backhaul system, that is what makes it viable to transmit packets between the Management Center and WSN clusters as a single wireless hop, instead of multiple hops wireless links, which are lossy and have high latency.

A WSN topology for smart cities is shown in Figure 1.

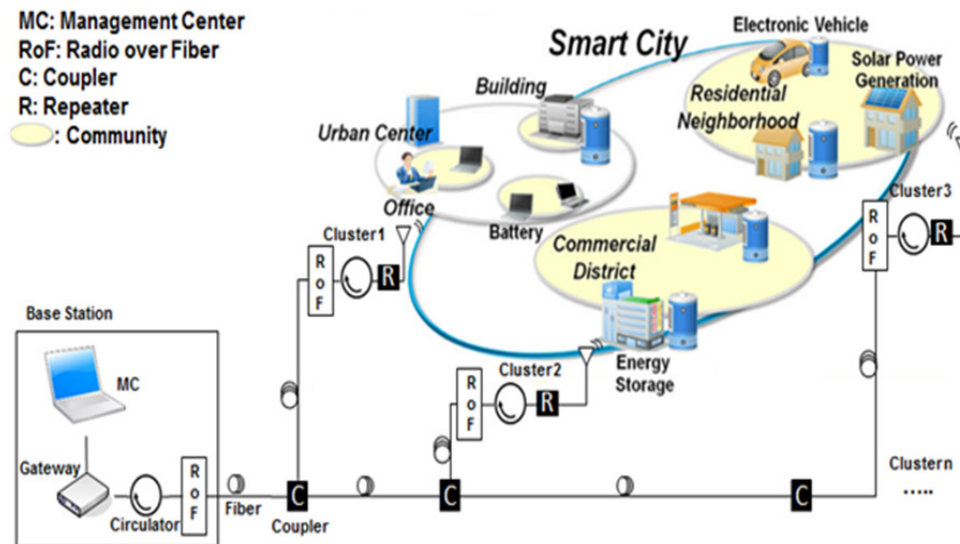


Fig. 1. Wireless Sensor Network with RoF in bus topology for Smart Cities

Fig. 1. Wireless Sensor Network with RoF in bus topology for Smart Cities

An important fact to be noted is that fiber is abundantly available because of great implantation of fiber in the last 30 years and the development of the WDM technique that uses only one pair of fibers to transport many streams of data (Ramaswami e Sivarajan, 2002).

Considering there is a fiber surplus, it can be used to create a robust WSN with RoF network with low delay and high reliability (A. WIBERG, 2008).

Although advantageous, the utilization of RoF as a backhaul for WSN implies that additional impairments must be considered. The conversion from radio to light and then from light to radio inserts a great amount of noise and also has its gain highly associated to laser modulation, demodulation, and optical attenuation.

RoF links are often used in a point-to-point topology (Fernando, 2001; Hall *et al.*, 2013; Hossen e Jang, 2009; Wiberg, 2008) environmental and economic benefit of the e-City dweller as did the introduction of utility and transport network infrastructures. Yet without action, the quantity of greenhouse gas emissions attributable to the increasing energy consumption of access networks will become a serious threat to the environment. This paper introduces the vision of a 'sustainable Digital City' and then considers strategies to overcome economic and technical hurdles faced by engineers responsible for developing the information and communications technology (ICT, which would lead to a central site with several RoF equipment connected to the clusters equipment. In order to reduce equipment cost it is proposed using the system in a bus topology, as in (Andrade, de *et al.*, 2012; Assumpção e Branquinho, 2012), this way connecting all the distant clusters RoF under only one RoF equipment on the central site.

This study extends the results obtained on another works (Assumpção e Branquinho, 2012), analyzing the coverage improvement on WSN brought by the use



of RoF bus network backhaul connected back-to-back to the radio repeaters. In the proposed testbed, the sensors nodes operate at 915 MHz, using FSK modulation, and are connected to the RoF bus network through a radio repeater. The packet loss and received signal strength were measured at the sensor. The results show that it is possible to use RoF in a WSN implementation. The impact of the penalty introduced by the RoF on the cluster coverage distance was mitigated with the use of the radio repeater, achieving a good performance for the RoF architecture proposed.

The chapter is organized as follows. Section 2 describes some basic concepts about RoF and WSN. In section 3 the proposed set up model is described, as well as its bus topology configuration and radio repeater. In Section 4 the results are presented and discussed. In Section 5 we show our conclusions and a perspective of future works.

## 2 | ROF AND WSN ARCHITECTURE

This section focuses on discussing aspects of the RoF and WSN architecture and their requirements.

### A. Radio over Fiber

The main elements of RoF systems are shown in Figure 2.

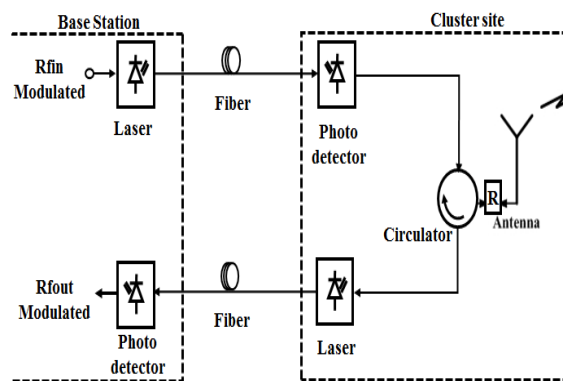


Fig. 2. RoF system architecture

The RoF transmit the RF signal by modulating a laser analogically as a mean to convert the electrical RF signal to an optical signal. The laser can be directly or externally modulated (Hossen e Jang, 2009) if only wireless ad-hoc network is used in large sensor network such as u-City, it requires too many hops causing problems like address assignment and time delay. Above problems are solved when the u-City is divided into multiple clusters, which are then connected with the personal area network coordinator (PANC). The RoF applied in this work uses direct modulation, and it's gain can be calculated by (Wiberg, 2008):

$$G = S_l(f)^2 \cdot \alpha^2 \cdot S_d(f)^2 \cdot \frac{R_d}{R_l} \quad (1)$$



Where  $G$  is the gain,  $S_l$  the laser slope,  $\alpha$  the attenuation,  $S_d$  the detector slope, and  $R_l$  and  $R_d$  the laser resistance and detector load resistance, respectively. RoF systems' noise figure comes from both generation and detection of the optical signal. Considering signal and link analysis, the most relevant noises in a direct modulation scheme are the laser relative intensity noise (RIN), the shot noise, and the thermal noise. Most direct modulated systems are RIN limited. The noise figure from a direct modulation RoF link can be calculated by Eq. (2) (Ackerman e Cox, 1998; Fernando e Anpalagan, 2004).

$$NF = 10 \cdot \log \left[ 1 + \frac{I_D^2 \cdot RIN \cdot R_d}{k \cdot T \cdot G} + \frac{2q \cdot I_D \cdot R_d}{k \cdot T \cdot G} + \frac{1}{G} \right] \quad (2)$$

Where  $I_D$  is the current at the photodetector output,  $R_d$  the detector load resistance,  $k$  the Boltzmann constant,  $T$  the temperature in kelvin,  $G$  the RoF gain, and  $q$  the charge of an electron.

Generally, commercial systems have their noise figure contribution presented in the form of equivalent input noise (EIN). The EIN represents the sum of all noise contributions and is given by Eq. (3) (Fernando, 2001).

$$EIN = N_{RIN} + N_{SHOT} + N_T \quad (3)$$

Where  $N_{RIN}$  means the noise detected at the receiver,  $N_{SHOT}$  the noise influenced by the photoreceptor current, and  $N_T$  the thermal noise, which depends on the temperature. EIN is usually presented in dBm/Hz units. Equation (3) is valid when the RoF link is composed of only one RoF as head end unit and another RoF as remote antenna unit (RAU). In this case, this system is considered to have a single cluster. For multi cluster systems, where more than one RAU is present, the EIN should consider the noise components from all the clusters. Thus, the EIN for a multi cluster system is given by Eq. (4), where the noise from all the clusters is considered.

$$EIN = \sum N_{RIN} + N_{SHOT} + N_T \quad (4)$$

It is important noticing that the RIN perceived and shot noise, both, are influenced by all the clusters in the system, while the thermal noise is dependent only on the two RoFs included in the link that is being analyzed.

From this, it is shown that noise figure analysis is case specific and must be done for each link.

## B. Wireless Sensor Network

The WSN consists of sensor nodes that are connected to each other through a

wireless network.

For the analysis of the wireless signal propagation, an additive white Gaussian noise channel may be considered. The signal-to-noise ratio (SNR) vs bit error rate (BER) curve for the modulation used can be calculated, aiming determine the necessary SNR required for any defined packet error rate (PER). The error probability for FSK-NC, considering Noncoherent (NC) frequency shift keying (FSK) modulation is given by Eq. (5) (Rappaport, 2002):

$$P_{e,FSK} = \frac{1}{2} e^{\left(-\frac{E_b}{N_0}\right)} \quad (5)$$

Where  $P_{e,FSK}$  is the error probability for FSK-NC,  $E_b$  the bit energy, and  $N_0$  the noise power spectral density. As the data rate used and the bandwidth necessary for the transmission are known, the SNR is given by Eq. (6):

$$\frac{S}{N} = \frac{E_b}{N_0} \cdot \frac{R}{B} \quad (6)$$

Where  $R$  is the data rate and  $B$  the bandwidth. With the BER and the transmission packet size, it is possible to obtain the PER value equivalent to the BER as shown in Eq. (7) (Texas Instruments, 2013).

$$PER = 1 - (1 - BER)^n \quad (7)$$

Where  $n$  is the packet size. The PER versus SNR for FSKNC modulation with 38.4 kbps bit rate, 406.25 kHz bandwidth and 496 bits packet, was considered.

The WSN's link budget is calculated seeing the RoF link as a gain, or attenuation, and a noise source. In this way the link budget equation considering the RoF gain is shown in Eq. (8).

$$P_{Rx} = P_{Tx} + G_{Tx} - L_{PL} + G_{Rx} + G_{RoF} \quad (8)$$

Where  $P_{Rx}$  is the reception power,  $P_{Tx}$  the transmission power,  $G_{Tx}$  the transmission antenna gain,  $L_{PL}$  the path loss,  $G_{Rx}$  and  $G_{RoF}$  the gains from reception antenna, and RoF, respectively. The path loss can be calculated by various methods, but for further calculations, the log-distance path loss model was used in Eq. (9) (Rappaport, 2002).

$$L_{PL} = 20 \cdot \log_{10} \left( \frac{4 \cdot \pi \cdot d_0}{\lambda} \right) + 10 \cdot \beta \cdot \log_{10} \left( \frac{d}{d_0} \right) \quad (9)$$

Where  $L_{PL}$  is the path loss,  $d_0$  the reference distance,  $\lambda$  the RF wavelength,  $\beta$  the path loss exponent and  $d$  the total distance.

The path loss exponent is dependent on the environment where the signal is propagated and, in this experiment, it was defined as 3.41, which is the exponent for an open area for WSN in 915 MHz (TIWARI e THAKUR, 2011)grocery stores, hard office, soft office, Textile Company, chemical company, metal company etc..

### C. WSN + RoF

The union of WSN and RoF result in some impairments for the radio communication. Since RoF system introduce a large amount of noise, it makes necessary to increase the signal power on the receiver to guarantee the same error rate.

The reception power can be defined as a function of the noise power, the SNR defined by the modulation, and the PER as given in Eq. (10).

$$P_{Rx} = E_{IN_{dBm/Hz}} + 10 \cdot \log_{10}(B) + SNR_{dB} \quad (10)$$

Where  $P_{Rx}$  is the reception power necessary for a given PER,  $E_{IN_{dBm/Hz}}$  the EIN as shown in Eq. (4) in dBm/Hz, B the signal bandwidth, and  $SNR_{dB}$  the SNR that is defined by a chosen PER given the modulation scheme.

Motivated by the huge degradation on distance coverage encountered in previous works (Assumpção e Branquinho, 2012), we decided to experiment using a multi-hop like scheme to mitigate the problem with the noise and coverage distance. For this, it was proposed a repeater formed by 2 sensor nodes, connected through a serial bus and containing a firmware which retransmitted in one radio what was received in the other. This repeater was then connected in one side directly to the RoF backbone and on the other side feeding the antenna system.

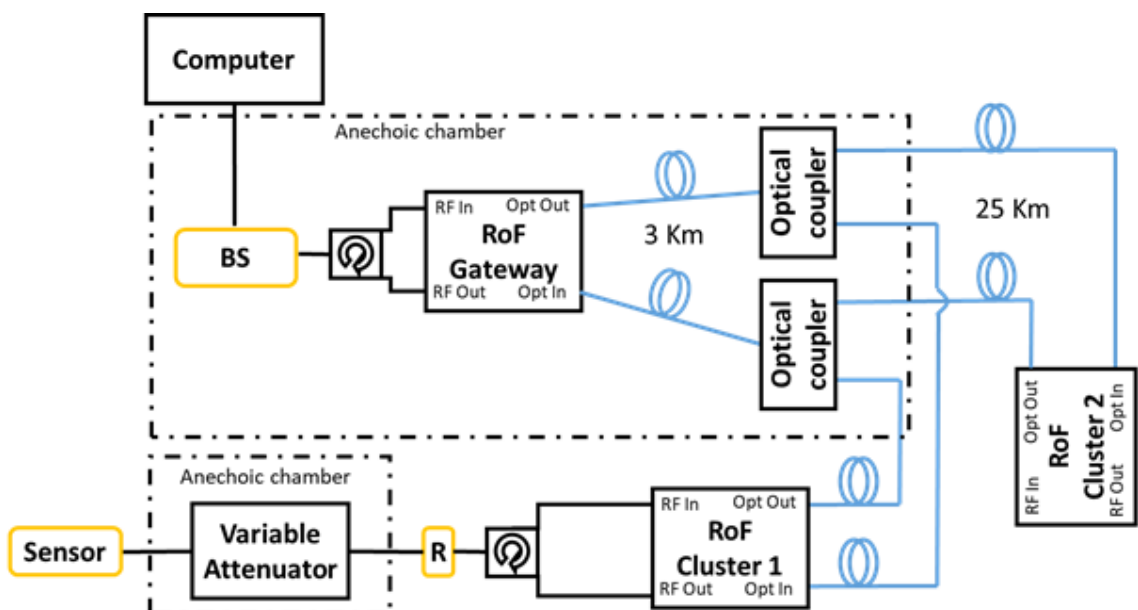


Fig. 3. Experimental Setup

### 3 | EXPERIMENTAL SETUP

The experimental setup was designed for testing the radio coverage adding a repeater to decouple the cluster coverage from the RoF backhaul impairments as shown in Figure 3.

The WSN with RoF testbed layout is modular and can be reorganized to match the proposed tests scenarios. The testbed is composed of a sensor node acting as the remote node, a variable attenuator to emulate the radio channel, a radio repeater, antenna duplexers, three commercial RoF equipment, two optical couplers, a sensor node working as base station, and a computer for data processing and logging.

To minimize uncertainties on radio propagation, such as multipath propagation, the testbed was cabled, and a variable attenuator was used to emulate the sensor being carried away from the RAU toward the cluster boundaries.

The Radiuino platform (Radiuino, [s.d.]) was used to configure the radio module parameters for the WSN's nodes such as channel frequencies, modulation, transmission rate, transmission power, serial baud rate and net address.

Each experiment round consisted of a set of data transmissions. While the attenuation was slowly raised, PER and RSSI data were collected. For each point of data 10,000 packets were transmitted from the sensor to the base station.

From the curves of PER vs RSSI obtained we collected the RSSI where 5% PER is observed and used it to calculate the theoretical coverage distance.

The modulation used was FSK with data rate of 38.4 kbps. The filter bandwidth was 406.25 kHz. The packets used in the transmission were 62 bytes long. The packet architecture used is similar to the one shown in (Assumpção e Branquinho, 2012), with a payload 43 bytes longer. Although CRC was included in the transmitted packet, correction code was not used. The RoF system used is a commercial unit (Zonu, [s.d.]). It works in 1310 nm and has an optical output of 3 dBm and a typical RF gain of 0 dB without optical attenuation (back to back). The EIN presented in the datasheet is 134 dBm/Hz and the RIN given by the manufacturer is 140 dB/Hz. The testbed has an optical attenuation of 6.5 dB between RoF clusters and RoF base station (BS). This was reflected as a measured RF gain of 15 dB in the performed tests.

The tests of this work were realized in three rounds as follow:

A) In the first experiment round, the BS was connected back-to-back to the repeater without any RoF. The repeater was connected to variable attenuator (radio channel emulator). The results from this experiment were used as a reference for the other tests.

B) In the second experiment round, the RoF from the BS and the first cluster were introduced in the test.

C) In the third experiment round, a second RoF (cluster 2) was connected to the testbed. As we just wanted to measure the degradation caused by adding the cluster, no sensor was connected to it (i.e., the RoF equipment on cluster 2 emitted a

continuous-wave signal).

Using the repeater connected directly with the RoF makes it possible to guarantee that the necessary power is received at the radio receiver. Since the packet is received and retransmitted on the repeater the cluster is thus decoupled from the RoF backbone transmission.

These tests were proposed to verify the feasibility of such arrangement for future use on a real case scenario.

#### 4 | RESULTS AND COMMENTS

The results of the experiments (A, B and C) described in the last section are presented in Figure 4.

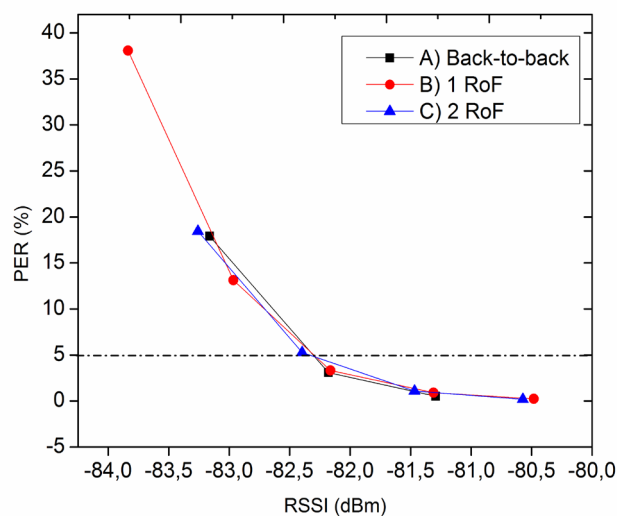


Fig. 4. PER versus RSSI

These results show that in all tests the WSN network performed the same in PER x RSSI, with the 3 curves overlapping each other.

From these results, taking the practical RSSI values for the links to perform with a PER of 5%, the coverage distances radius were calculated using Eq. (7).

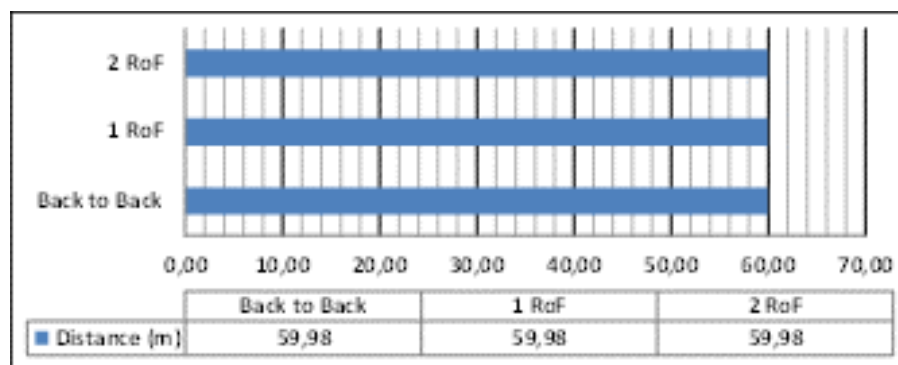


Fig. 5. Calculated Coverage

The results show that the cluster coverage was not affected by the insertion of the

RoF backhaul or even the insertion of another cluster to the backhaul.

It is important to note that in previous work (Assumpção e Branquinho, 2012), where there were no repeaters in the setup, the coverage distances obtained in the experimental setup were very inferior in tests with 1 and 2 RoFs, besting at approximately 21% when there was only 1 RoF and attenuation was compensated with a low-noise amplifier.

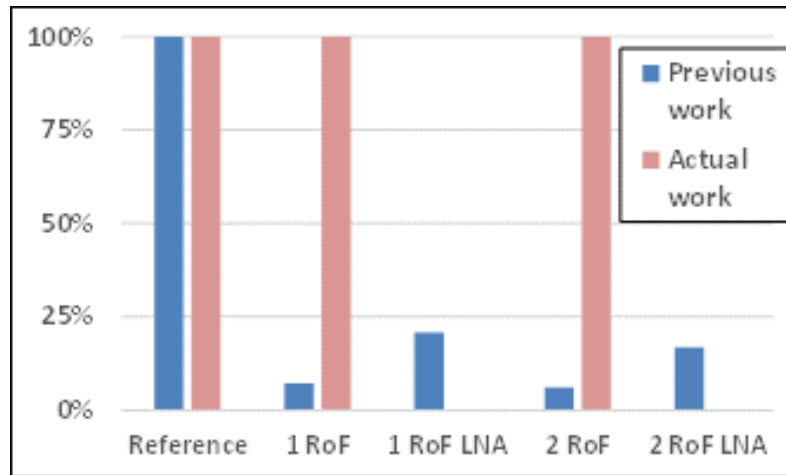


Fig. 6. Comparison between reference and other scenarios

Figure 6 shows a comparison in percentage of coverage radius in both works. Note that the radio parameters like transmission rate, packet size, filter bandwidth, etc. from this work are not the same as in previous work (Assumpção e Branquinho, 2012), preventing direct comparison on coverage distance.

By using the radio repeater, the RF signal is re-formatted before being sent to the WSN cluster, thus having no reduction in coverage area when using 1 and 2 RoFs, making it a suitable strategy to the smart cities context.

## 5 | CONCLUSIONS

The main proposal of this work is an experimental evaluation of the coverage of a Wireless Sensor Network with Radio-over-Fiber bus network backhaul that can be implemented in Smart City applications under the Internet of Things (IoT) context. The proposal was a RoF backhaul assembled in a bus topology using a radio repeater for signal distribution to the WSN.

The results show that the radio repeater device mitigate the impairments imposed by the RoF to signal quality by decoupling the wireless cluster from the RoF backhaul. The improvement in the coverage area of the WSN enables the system to be used in Smart Cities applications.

The bus topology in the RoF network is proposed as a new contribution for RoF usage, maintaining performance, reduced attenuation and strong electromagnetic immunity while reducing infrastructure cost.

For future works, we intend to extend this experiment using more RoF clusters and implementing this setup in a real use case for further tests.

## 6 | ACKNOWLEDGEMENTS

The authors wish to acknowledge the financial support from FINEP in the research laboratory LP-SiRa of PUC-Campinas.

The authors also want to acknowledge to CNPq (National Council of Scientific and Technologic Development) by granting PhD scholarship to Raphael M. Assumpção.

## REFERENCES

- ACKERMAN, E. I.; COX, C. H. . I. **State of the art in analog fiber-optic link technology**, 1998 URSI International Symposium on Signals, Systems, and Electronics. Conference Proceedings (Cat. No.98EX167), 1998.
- ANDRADE, T. P. C. DE; FONSECA, N. L. S. DA; OLIVEIRA, L. B.; BRANQUINHO, O. C. **MAC protocols for wireless sensor networks over radio-over-fiber links**, 2012 IEEE International Conference on Communications (ICC), 2012.
- ASSUMPÇÃO, R. M.; BRANQUINHO, O. C. **Design methodology for wireless sensor network working with radio over fiber.**, Microwave and Optical Technology Letters, v. 54, n. 5, p. 1277–1281, 2012.
- FERNANDO, X. N. **Signal processing for optical fiber based wireless access.**, Calgary: University of Calgary, 2001.
- FERNANDO, X. N.; ANPALAGAN, A. **On the design of optical fiber based wireless access systems**, 2004 IEEE International Conference on Communications (IEEE Cat. No.04CH37577), 2004
- HALL, T. J. et al. **Radio-over-Fibre access for sustainable Digital Cities.**, Annales des Telecommunications/Annals of Telecommunications, v. 68, n. 1–2, p. 3–21, 2013.
- HOSSEN, M.; JANG, B. J. **Extension of wireless sensor network by employing RoF-based 4G network**, In: Advanced Communication Technology, 2009. ICACT 2009. 11th International Conference on. IEEE, 2009. p. 275-278.
- RADIUINO. **Radiuino**. Disponível em: <<http://radiuino.cc/>>. Acesso em: 16 abr. 2017.
- RAMASWAMI, R.; SIVARAJAN, K. N. **Optical Networks: A Practical Perspective**. 2nd edition ed. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., 2002.
- RAPPAPORT, T. S. **Wireless communications : principles and practice**. 2nd. ed. Upper Saddle River, N.J.: Prentice Hall PTR, 2002. v. 2
- TEXAS INSTRUMENTS. **CC1101 Low-Power Sub-1 GHz RF Transceiver**, Texas instruments, 2013. Disponível em: <<http://focus.ti.com/lit/ds/symlink/cc1101.pdf>>
- TIWARI, A.; THAKUR, S. S. **Consequence of Urbanization on Wireless Sensor Network Reportage Domain.**, International Journal of Engineering Science and Technology, v. 3, n. 3, 2011.
- WIBERG, A. **Generation , Modulation , and Detection of Signals in Microwave Photonic Systems.**, Göteborg: Chalmers University of Technology, 2008.



ZANELLA, A.; BUI, N.; CASTELLANI, A.; VANGELISTA, L.; ZORZI, M. **Internet of Things for Smart Cities.**, IEEE Internet of Things Journal, v. 1, n. 1, p. 22–32, fev. 2014.

ZONU, O. **RF over Fiber Optic Transceiver**, Optical Zonu, , [s.d.]. Disponível em: <<http://www.opticalzonu.com/standalone/oz810/>>

## MODELO ELETROMAGNÉTICO DE UM ARRANJO PLANAR DE NANODIPOLOS SOBRE PLANO DE OURO ATRAVÉS DA FUNÇÃO DE GREEN 3D

### André Felipe Souza da Cruz

Universidade Federal do Pará, Faculdade de Eng.  
Elétrica.  
Tucuruí – Pará

### Nadson Welkson Pereira de Souza

Universidade Federal do Sul e Sudeste do Pará,  
Faculdade de Eng. Elétrica.  
Marabá – Pará

### Karlo Queiroz da Costa

Universidade Federal do Pará, Faculdade de Eng.  
Elétrica.  
Tucuruí – Pará

**RESUMO:** Neste trabalho é apresentado a modelagem eletromagnética tridimensional de um arranjo periódico planar de nanodipolos operando em frequência óptica sobre um plano de ouro. Os campos eletromagnéticos foram determinados através do método do potencial magnético, definido por meio da Função de Green Periódica (PGF). Foram realizadas análises paramétricas de campo próximo e distante, e análises dos coeficientes de reflexão e transmissão. A concentração de luz sobre o plano de ouro devido à ressonância plasmônica superficial é observada. Para verificar a convergência do método, foram gerados resultados gráficos da convergência de Cauchy das séries da PGF cossenoidal. Os resultados apontam que as condições físicas de contorno

são atendidas, tanto de campo próximo quanto campo distante, demonstrando que o modelo matemático obtido pode ser utilizado como ferramenta na análise e desenvolvimento de dispositivos fotônicos de sensoriamento e comunicação.

**PALAVRAS-CHAVE:** Arranjo planar, Função de Green Periódica, Campos Próximo e Distante, Método Analítico.

**ABSTRACT:** In this work is presented the three - dimensional electromagnetic modeling of a planar periodic array of nanodipoles operating at optical frequency upon a gold plane. The electromagnetic fields were determined by the magnetic potential method, defined by the Periodic Green's Function (PGF). Parametric analyzes of the near and far field were performed, as well analysis of the reflection and transmission coefficients. The concentration of light on the gold layer due to the surface plasmon resonance is observed. To verify the method convergence, graphical results of the Cauchy convergence of the cosine PGF series were generated. The results show that the physical boundary conditions are met, both in the near and far field, demonstrating that the obtained mathematical model can be used as a tool in the analysis and development of photonic devices for sensing and communication.

**KEYWORDS:** Planar Array, Periodic Green's

## 1 | INTRODUÇÃO

Nanoantenas, também conhecidas como antenas ópticas, são estruturas metálicas com dimensões de poucos nanômetros, capazes de transmitir, receber, confinar e amplificar a luz (POHL et al., 2013; SANTOS, 2012). Estas características possibilitam o controle e a modulação da luz, desempenhando assim, o papel principal em muitas aplicações, tais como em sensores (RODRÍGUES et al., 2011), microscopia óptica de alta resolução (SQALLI et al., 2002), controle e diretividade da luz (AHMEDI et al., 2010), entre outras.

Nanopartículas, com dimensões menores que o comprimento de onda de excitação, podem ser equivalentemente modeladas por nanoantenas do tipo dipolo (nanodipolos) ou por correntes elementares (HOFMANN et al., 2007). Dessa forma, os efeitos do acoplamento óptico de um arranjo de nanopartículas, sobre um substrato e suas interações de campo próximo e distante, podem ser descritas através da solução de um arranjo de nanodipolos (TAJDINI & MOSALLAEIL, 2011; BOHREN & HUFFMAN, 1983).

Especial atenção tem sido dada a nanopartículas posicionadas sobre metais nobres (ouro e prata) (COSTA et al., 2015; RASHIDI & MOSALLAEIL, 2010), pois, quando excitadas, sua re-irradiação pode provocar ondas plasmônicas de superfície na interface metálica que, por sua vez, alteram características de concentração e emissão de luz do conjunto (AHMEDI et al., 2010; SOUZA et al., 2017). Diante disto, diferentes arranjos periódicos de nanoantenas têm sido propostos (ROSTAMI et al., 2013; RASHIDI et al., 2013). A análise e designer destas estruturas são, em geral, realizados através de métodos numéricos, tais como Método dos Elementos Finitos (FEM) e Método dos Momentos (MOM), entretanto, a representação de arranjos de nanoantenas pode demandar elevado custo e tempo computacional. Contudo, métodos analíticos baseados em Funções de Green Periódicas (PGF) no domínio espectral são propostos, viabilizando a modelagem eletromagnética (CRUZ et al., 2018; DUDLEY, 1994).

Neste trabalho, é apresentado o estudo analítico *full-wave* tridimensional de um arranjo planar periódico de nanodipolos hertzianos sobre um plano metálico de ouro. Os campos eletromagnéticos foram definidos através do método do potencial magnético, determinado através da solução da equação de Helmholtz, para isso, utilizou-se o método da transformada série complexa de Fourier, com condições periódicas na coordenada  $xy$ , e o método da função de Green na coordenada  $z$ . Foram realizadas análises de campo próximo e distante para um comprimento de onda 632.8nm, e análises espectrais dos coeficientes de reflexão e transmissão TE e TM. Através da identidade de Euler, os somatórios nas séries complexas foram alterados para reduzir o custo computacional das simulações, e por fim, verificou-se a convergência de

Cauchy destas séries.

## 2 | MODELAGEM ANALÍTICA

### 2.1 Descrição Teórica e Modelagem da Estrutura

A estrutura em análise é definida por um arranjo periódico de nanodipolos hertzianos sobre uma camada de ouro infinita, conforme Fig. 1:

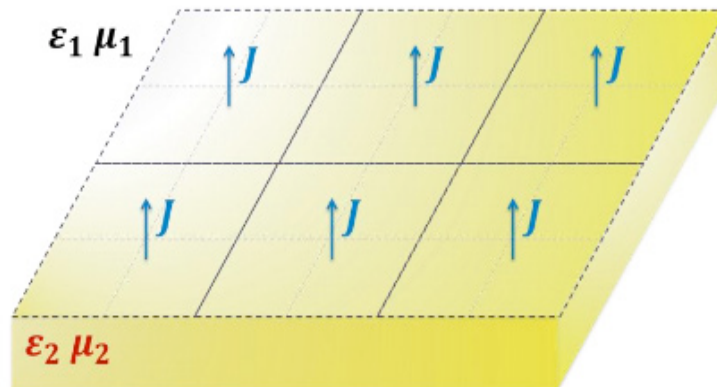


Figura 1 - Arranjo periódico de nanodipolos sobre camada de ouro.

Consideram-se nanodipolos com comprimento  $l \ll \lambda/50$  localizados a uma altura  $h$  da interface  $d$ , representados por correntes concentradas  $\mathbf{j}$ , com momentos de corrente orientados na direção  $\hat{\mathbf{a}}_r$  pelos ângulos fixos  $\theta'$  e  $\phi'$ :

$$\mathbf{J}(\theta', \phi') = I_0 l \delta(\mathbf{r} - \mathbf{r}') [\cos \phi' \sin \theta' \hat{\mathbf{a}}_x + \sin \phi' \sin \theta' \hat{\mathbf{a}}_y + \cos \theta' \hat{\mathbf{a}}_z] \quad (1)$$

onde  $\mathbf{r} = (x, y, z)$  e  $\mathbf{r}' = (0, 0, h)$ . Como a estrutura possui periodicidade bidimensional ao longo do plano  $xy$ , podemos realizar a análise de campo através de uma célula com largura  $2a$  em  $x$  e  $2b$  em  $y$  (Fig. 2a). Neste trabalho ambos os meios são considerados não magnéticos, sendo o meio 1  $(\epsilon_1, \mu_1)$ ,  $Z > d$ , o espaço livre, e o meio 2  $(\epsilon_2, \mu_2)$ ,  $z < d$ , caracterizado pela permissividade relativa complexa do ouro que, no espectro óptico, é descrita pelo modelo de Lorentz-Drude com excelente precisão (SOUZA et al., 2017).

A região de análise é delimitada por dois volumes  $V_1$  e  $V_2$ , meio 1 e meio 2, respectivamente, encerrados pelas superfícies fechadas  $S_1 = S_{11} + S_{12} + S_{13} + S_{14} + S_{15} + S_{16}$  e  $S_2 = S_{21} + S_{22} + S_{23} + S_{24} + S_{25} + S_{26}$ , conforme Fig. 2b.

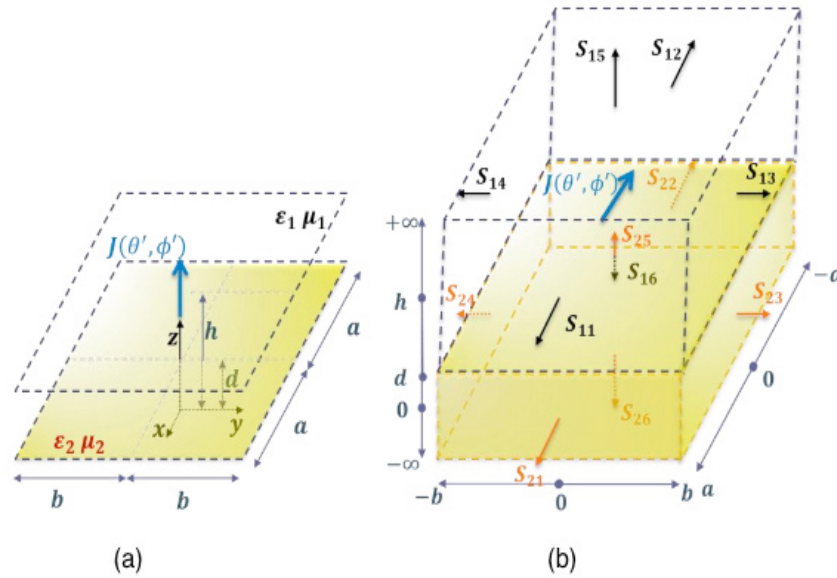


Figura 2 - (a) Célula de análise. (b) Regiões de análise V1 e V2.

Para determinar analiticamente os campos elétrico e magnético, primeiramente foi definido o potencial magnético em ambos os meios através da solução da equação diferencial vetorial de Helmholtz, não homogênea no meio 1, e homogênea no meio 2 (DUDLEY, 1994):

$$\begin{aligned}\nabla^2 \mathbf{A}_1 + k_1^2 \mathbf{A}_1 &= -\mu_1 \mathbf{J} \\ \nabla^2 \mathbf{A}_2 + k_2^2 \mathbf{A}_2 &= 0\end{aligned}\quad (2)$$

Note que o campo potencial nas superfícies  $S_{11}, S_{12}, S_{13}, S_{14}, S_{21}, S_{22}, S_{23}, S_{24}$  deve atender as condições periódicas em (3), e em regiões distantes da fonte, como nas superfícies  $S_{15}$  e  $S_{26}$ , o campo potencial deve atender as condições de limite em (4), ou seja, serem nulos no infinito.

$$\begin{aligned}\mathbf{A}_1(a, y, z)|_{S_{11}} &= \mathbf{A}_1(-a, y, z)|_{S_{12}}, & \frac{\partial}{\partial x} \mathbf{A}_1(a, y, z)|_{S_{11}} &= \frac{\partial}{\partial x} \mathbf{A}_1(-a, y, z)|_{S_{12}} \\ \mathbf{A}_1(x, b, z)|_{S_{13}} &= \mathbf{A}_1(x, -b, z)|_{S_{14}}, & \frac{\partial}{\partial y} \mathbf{A}_1(x, b, z)|_{S_{13}} &= \frac{\partial}{\partial y} \mathbf{A}_1(x, -b, z)|_{S_{14}} \\ \mathbf{A}_2(a, y, z)|_{S_{21}} &= \mathbf{A}_2(-a, y, z)|_{S_{22}}, & \frac{\partial}{\partial x} \mathbf{A}_2(a, y, z)|_{S_{21}} &= \frac{\partial}{\partial x} \mathbf{A}_2(-a, y, z)|_{S_{22}} \\ \mathbf{A}_2(x, b, z)|_{S_{23}} &= \mathbf{A}_2(x, -b, z)|_{S_{24}}, & \frac{\partial}{\partial y} \mathbf{A}_2(x, b, z)|_{S_{23}} &= \frac{\partial}{\partial y} \mathbf{A}_2(x, -b, z)|_{S_{24}} \\ \lim_{z \rightarrow +\infty} \mathbf{A}_1(x, y, z)|_{S_{15}} &= 0, & \lim_{z \rightarrow -\infty} \mathbf{A}_2(x, y, z)|_{S_{26}} &= 0\end{aligned}\quad (3)$$

$$\lim_{z \rightarrow +\infty} \mathbf{A}_1(x, y, z)|_{S_{15}} = 0, \quad \lim_{z \rightarrow -\infty} \mathbf{A}_2(x, y, z)|_{S_{26}} = 0 \quad (4)$$

Na interface  $z=d$  formada pelas superfícies  $S_{16}$  e  $S_{25}$ , o campo potencial deve atender (5) para que sejam obedecidas às condições de continuidade dos campos EM tangenciais e fluxos EM normais.

$$\begin{aligned}
A_{x_2,y_2}|_{z=d} = A_{x_1,y_1}|_{z=d} \quad , \quad \frac{1}{\mu_2} \frac{\partial}{\partial z} A_{x_2,y_2} \Big|_{z=d} = \frac{1}{\mu_1} \frac{\partial}{\partial z} A_{x_1,y_1} \Big|_{z=d} \\
\frac{1}{\mu_2} A_{z_2} \Big|_{z=d} = \frac{1}{\mu_1} A_{z_1} \Big|_{z=d} \quad , \quad \frac{1}{\varepsilon_2 \mu_2} \nabla \cdot \mathbf{A}_2 \Big|_{z=d} = \frac{1}{\varepsilon_1 \mu_1} \nabla \cdot \mathbf{A}_1 \Big|_{z=d}
\end{aligned} \tag{5}$$

Cada equação em (2) produz três equações diferenciais escalares com a mesma estrutura e condições de contorno idênticas, com exceção das condições na interface  $z=d$ , dito isso, primeiramente determinou-se a solução para uma componente genérica  $A_v$  excitada pela fonte genérica  $J_v = I_0 l f_v(\theta', \phi')$ , e posteriormente o resultado foi estendido para as componentes  $A_x, A_y$  e  $A_z$ , no meio 1 e 2.

## 2.2 Tensores de Campo Através do Método PGF

Este tópico apresenta os principais passos do formalismo utilizado na obtenção das soluções de (2), com condições (3) - (5). As equações diferenciais genéricas de  $A_v$  foram determinadas com o auxílio da Função de Green Periódica (PGF), com condições de Neumann e Limite em  $z$ , e condições periódicas em  $xy$ , todas condições homogêneas e derivadas a partir das condições de contorno dos problemas em  $A_v$  (DUDLEY, 1994). O método baseia-se em utilizar a transformada dupla série complexa de Fourier para levar o problema tridimensional espacial ao domínio espectral discreto  $m$  e  $n$ , retirando a dependência espacial de  $xy$ , e então resolvendo o problema unidimensional na variável  $z$  com constante de propagação  $k_{zu}$  (6), com  $Im\{k_{zu}\} < 0$ , nos meios 1 e 2.

$$k_{z1} = \sqrt{k_1^2 - \left[\frac{m\pi}{a}\right]^2 - \left[\frac{n\pi}{b}\right]^2} \tag{6}$$

Através da transformada inversa dupla série complexa de Fourier (DUDLEY, 1994) obtém-se a PGF no domínio espacial  $xyz$  (7), em que os indices 1 e 2 indicam o meio de análise.

$$g^{1,2}(\mathbf{r}, \mathbf{r}') = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} g_{mn}^{1,2}(z, z') \bar{u}_m(x') \bar{u}_n(y') u_m(x) u_n(y) \tag{7}$$

$$u_m(x) = (2a)^{-1/2} e^{j\frac{m\pi}{a}x} \tag{8}$$

$$u_n(y) = (2b)^{-1/2} e^{j\frac{n\pi}{b}y} \tag{9}$$

onde  $u_m(x)$  e  $u_n(y)$  são as autofunções próprias do problema periódico em  $x$  e  $y$ , com autovalores  $k_x = m\pi/a$  e  $k_y = n\pi/b$ . No domínio espectral  $mn$  as funções de

Green unidimensionais em z no meio 1 e 2 são, respectivamente:

$$g_{mn}^1(z, z') = \frac{1}{2jk_{z1}} [e^{-jk_{z1}|z-z'|} + e^{-jk_{z1}(z+z'-2d)}] \quad (10)$$

$$g_{mn}^2(z, z') = \frac{1}{2jk_{z2}} [e^{-jk_{z2}|z'-z|} + e^{jk_{z2}(z+z'-2d)}] \quad (11)$$

Aplicando a identidade de Green ao produto interno entre a equação diferencial escalar de Helmholtz, na componente genérica  $A_v$ , e a função de Green (7) nos meios  $u=1$  e  $2$ :

$$\langle -(\nabla^2 + k_u^2)A_{vu} | g^u \rangle = \langle A_{vu} | -(\nabla^2 + k_u^2)g^u \rangle + \mathcal{J}(A_{vu}, g^u)|_{S_u} \quad (12)$$

O operador  $-\nabla^2$ , para condições de contorno periódicas, Neumann e limite, possui propriedades do operador auto-adjunto (DUDLEY, 1994). Assim, a solução para as componentes genéricas do campo potencial magnético no meio 1 e 2 são obtidas por:

$$A_{v1}(\mathbf{r}) = \frac{\mu_1}{j8ab} \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \left\{ I_0 I_f(\theta', \phi') \frac{1}{k_{z1}} [e^{-jk_{z1}|z-h|} + e^{-jk_{z1}(z+h-2d)}] - \frac{2}{\mu_1 k_{z1}} e^{-jk_{z1}(z-d)} \Lambda_{v1}^{mn} \right\} e^{-j\pi \left[ \frac{m}{a}x + \frac{n}{b}y \right]} \quad (13)$$

$$A_{v2}(\mathbf{r}) = \frac{1}{j8ab} \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \left\{ \frac{2}{k_{z2}} e^{jk_{z2}(z-d)} \Lambda_{v2}^{mn} \right\} e^{-j\pi \left[ \frac{m}{a}x + \frac{n}{b}y \right]} \quad (14)$$

onde:

$$\Lambda_{v1}^{mn} = \int_{-b}^b \int_{-a}^a e^{j\pi \left[ \frac{m}{a}x' + \frac{n}{b}y' \right]} \frac{\partial}{\partial z} A_{v1}(x', y', d) dx' dy' \quad (15)$$

$$\Lambda_{v2}^{mn} = \int_{-b}^b \int_{-a}^a e^{j\pi \left[ \frac{m}{a}x' + \frac{n}{b}y' \right]} \frac{\partial}{\partial z} A_{v2}(x', y', d) dx' dy' \quad (16)$$

Aplicando as condições de fronteira (5) na interface  $z=d$  e resolvendo o sistema de equações formado por (13)-(16), o campo potencial magnético na forma tensorial é dado por (17).



$$\mathbf{A} = \mathbf{I}(\theta', \phi') \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \begin{bmatrix} A_{xx}^{mn} & 0 & 0 \\ 0 & A_{yy}^{mn} & 0 \\ -j \left[ \frac{m\pi}{a} \right] A_{zx}^{mn} & -j \left[ \frac{n\pi}{b} \right] A_{zy}^{mn} & A_{zz}^{mn} \end{bmatrix} e^{-j\pi \left[ \frac{m}{a}x + \frac{n}{b}y \right]} \quad (17)$$

com elementos  $A_{ij}^{mn}$ , sendo a componente  $i$  excitada pela componente  $j$  da fonte de corrente linear  $\mathbf{I}(\theta', \phi')$ , conforme (18).

$$\mathbf{I}(\theta', \phi') = \frac{I_0 l}{8ab} \left[ \cos \phi' \sin \theta' \hat{a}_x + \sin \phi' \sin \theta' \hat{a}_y + \cos \theta' \hat{a}_z \right] \quad (18)$$

Nos meios 1 e 2, as componentes do tensor (17) são formadas por:

$$A_{xx1}^{mn} = A_{yy1}^{mn} = -j\mu_1 \frac{1}{k_{z1}} \left[ e^{-jk_{z1}|z-h|} + R^{TE} e^{-jk_{z1}(z+h-2d)} \right] \quad (19)$$

$$A_{zx1}^{mn} = A_{zy1}^{mn} = \mu_1 S^{TEM} e^{-jk_{z1}(z+h-2d)} \quad (20)$$

$$A_{zz1}^{mn} = -j\mu_1 \frac{1}{k_{z1}} \left[ e^{-jk_{z1}|z-h|} + R^{TM} e^{-jk_{z1}(z+h-2d)} \right] \quad (21)$$

$$A_{xx2}^{mn} = A_{yy2}^{mn} = -j\mu_2 \left[ \frac{\mu_1 k_{z2}}{\mu_2 k_{z1}} \right] \frac{1}{k_{z2}} e^{j[k_{z2}(z-d) - k_{z1}(h-d)]} \quad (22)$$

$$A_{zx2}^{mn} = A_{zy2}^{mn} = \mu_2 S^{TEM} e^{j[k_{z2}(z-d) - k_{z1}(h-d)]} \quad (23)$$

$$A_{zz2}^{mn} = -j\mu_2 \left[ \frac{k_{z2}}{k_{z1}} \right] T^{TM} \frac{1}{k_{z2}} e^{j[k_{z2}(z-d) - k_{z1}(h-d)]} \quad (24)$$

onde os coeficientes de reflexão e transmissão dos modos TE e TM são:

$$R^{TE} = \left[ \frac{\mu_2 k_{z1} - \mu_1 k_{z2}}{\mu_2 k_{z1} + \mu_1 k_{z2}} \right] \quad (25)$$

$$T^{TE} = \left[ \frac{2\mu_2 k_{z1}}{\mu_2 k_{z1} + \mu_1 k_{z2}} \right] \quad (26)$$

$$R^{TM} = \left[ \frac{\varepsilon_2 k_{z1} - \varepsilon_1 k_{z2}}{\varepsilon_2 k_{z1} + \varepsilon_1 k_{z2}} \right] \quad (27)$$

$$T^{TM} = \left[ \frac{2\varepsilon_2 k_{z1}}{\varepsilon_1 k_{z2} + \varepsilon_2 k_{z1}} \right] \quad (28)$$

e STEM um coeficiente de acoplamento que relaciona os coeficientes de transmissão dos modos TE e TM (CHEW, 1995):

$$S^{TEM} = \frac{1}{2k_{z1}^2} \left[ \frac{\mu_1 \epsilon_1}{\mu_2 \epsilon_2} - 1 \right] T^{TM} T^{TE} \quad (29)$$

Definido o potencial magnético (17), os tensores de campos magnético e elétrico próximos e distante podem ser facilmente obtidos através de operações diferenciais (BALANIS, 2005).

### 2.3 Forma Alternativa da Representação da PGF

A representação espectral discreta, nas soluções das PGF em (17), são feitas na forma de somatórios duplos de menos infinito a mais infinito, o que pode demandar um razoável custo computacional nas simulações. Para reduzir tal custo é proposto alterar os somatórios utilizando a identidade de Euler, desde que os termos espectrais nas expressões de campo obedecem as seguintes condições:

$$\begin{aligned} c_{m,n}(z) &= c_{-m,-n}(z) = c_{-m,n}(z) = c_{m,-n}; \\ c_{m,0}(z) &= c_{-m,0}(z); \quad c_{0,n}(z) = c_{0,-n}(z) \end{aligned} \quad (30)$$

As identidades propostas são:

$$\begin{aligned} S &= \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} c_{mn}(z) e^{-j\pi \left[ \frac{m}{a}x + \frac{n}{b}y \right]} \\ &= \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \epsilon_{mn} c_{mn}(z) \cos \left[ \frac{m\pi}{a}x \right] \cos \left[ \frac{n\pi}{b}y \right] \end{aligned} \quad (31)$$

$$\begin{aligned} \frac{\partial S}{\partial x} &= \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} c_{mn}(z) \left[ -j \frac{m\pi}{a} \right] e^{-j\pi \left[ \frac{m}{a}x + \frac{n}{b}y \right]} \\ &= \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \epsilon_{mn} c_{mn}(z) \left[ -\frac{m\pi}{a} \right] \sin \left[ \frac{m\pi}{a}x \right] \cos \left[ \frac{n\pi}{b}y \right] \end{aligned} \quad (32)$$

$$\begin{aligned} \frac{\partial S}{\partial y} &= \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} c_{mn}(z) \left[ -j \frac{n\pi}{b} \right] e^{-j\pi \left[ \frac{m}{a}x + \frac{n}{b}y \right]} \\ &= \sum_{m=0}^{\infty} \sum_{n=0}^{\infty} \epsilon_{mn} c_{mn}(z) \left[ -\frac{n\pi}{b} \right] \cos \left[ \frac{m\pi}{a}x \right] \sin \left[ \frac{n\pi}{b}y \right] \end{aligned} \quad (33)$$

onde  $\epsilon_{mn}$  é o número de Neumann duplo:

$$\epsilon_{mn} = \begin{cases} 1, & m = 0, n = 0 \\ 2, & m \neq 0, n = 0 \\ 2, & m = 0, n \neq 0 \\ 4, & m \neq 0, n \neq 0 \end{cases} \quad (33)$$

### 3 | RESULTADOS

A partir do modelo teórico desenvolvido, foram preparados algoritmos computacionais no software Matlab, que permitem avaliar numericamente os coeficientes de reflexão e transmissão TE e TM na interface ar/ouro, o campo próximo e o diagrama de radiação de um arranjo de nanodipolos. Os resultados dos coeficientes de transmissão e reflexão foram coletados no intervalo  $\lambda$  entre 450nm e 750nm. Para análises de campo próximo e diagrama de radiação foram gerados resultados gráficos em  $\lambda=632.8\text{nm}$ , equivalente ao utilizado em circuitos ópticos. No programa desenvolvido foram utilizadas as séries cossenoidais propostas em (31) - (33), estas foram truncadas utilizando  $m=0:M$  e  $n=0:N$ . Por fim, verificou-se a característica de convergência das séries cossenoidais nas PGF do campo potencial A.

#### 3.1 Coeficientes de Reflexão e Transmissão TE e TM

Os modos TE e TM, ou mesmo modos híbridos, podem ser excitados na estrutura através da orientação dos nanodipolos sobre o plano de ouro. Os campos radiados pelos nanodipolos são compostos por somas discretas de ondas planas, das quais a componente espectral fundamental  $m=n=0$ , com incidência normal (onda predominante), é a escolhida para as análises gráficas dos coeficientes de reflexão e transmissão apresentados na Fig. 3.

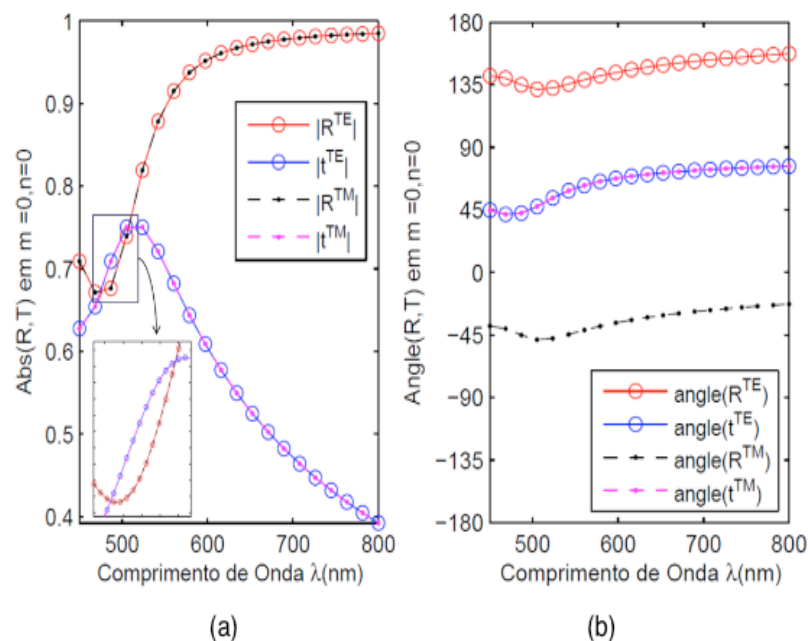


Figura 3 - Coeficientes  $R^{TE}$ ,  $t^{TE}$ ,  $R^{TM}$  e  $t^{TM}$ : (a) Módulo (b) Fase, para  $m=n=0$ .

A partir da Fig. 3a nota-se que os coeficientes de reflexão  $R^{TE}$  e  $R^{TM}$ , assim como os coeficientes de transmissão  $t^{TE}$  e  $t^{TM}$ , são iguais, o que significa que em amplitude as ondas refletidas e transmitidas, para incidência normal, nos modos TM e TE são iguais. Na Fig. 3b, verifica-se que os coeficientes de reflexão  $R^{TE}$  e  $R^{TM}$  estão defasados de  $180^\circ$  em todos os comprimentos de onda, estando de acordo com o esperado para incidência normal sobre bons condutores (BALANIS, 2005). Os coeficientes de transmissão  $t^{TE}$  e  $t^{TM}$  possuem o mesmo ângulo de defasagem (curva ao centro), indicando que a transmissão nestes modos sempre está em fase.

Na Fig. 3a, nota-se também que para comprimentos de onda menores, o ouro perde sua característica de condutor perfeito, ou seja, há a presença da onda transmitida. Na maior parte do espectro em análise, a transmissão é menor que a reflexão, com exceção de uma pequena faixa entre 473.64nm e 507.89nm, onde a transmissão se torna ligeiramente maior que a reflexão. O pico de transmissão ocorre no comprimento de onda  $\lambda \approx 514.14\text{nm}$ . Os coeficientes de reflexão efetiva foram calculados a partir de (25) e (27), e os coeficientes de transmissão efetiva foram calculados através das relações:

$$t^{TM} = \left[ \frac{k_{z2}\epsilon_1}{k_{z1}\epsilon_2} \right] T^{TM} = 1 - R^{TM} \quad (35)$$

$$t^{TE} = \left[ \frac{k_{z2}\mu_1}{k_{z1}\mu_2} \right] T^{TE} = 1 - R^{TE} \quad (36)$$

### 3.2 Campos Elétrico e Magnético Próximos

Foram gerados gráficos das componentes  $E_z$  e  $H_y$  em três células, considerando que os nanodipolos nos arranjos possuem comprimento  $l \ll \lambda/50$  e estão localizados no centro da célula, a uma distância  $a = 3\lambda$ , em x, e  $b=3\lambda$ , em y, das bordas da célula, posicionados em uma altura  $h = \lambda/10$  da interface metal/ar em  $z = d = 0$ . Todos excitados por uma corrente normalizada em  $I_0 = 1A$ .

A Fig. 4 mostra a distribuição de campo no plano xy em  $z = 0.15\lambda$  (altura focal) para um arranjo de nanodipolos horizontais ( $\theta' = 90^\circ$  e  $\phi' = 45^\circ$ ). A Fig. 5 mostra a distribuição no plano xz em  $y=0$  para um arranjo de nanodipolos verticais ( $\theta' = 0^\circ$ ).

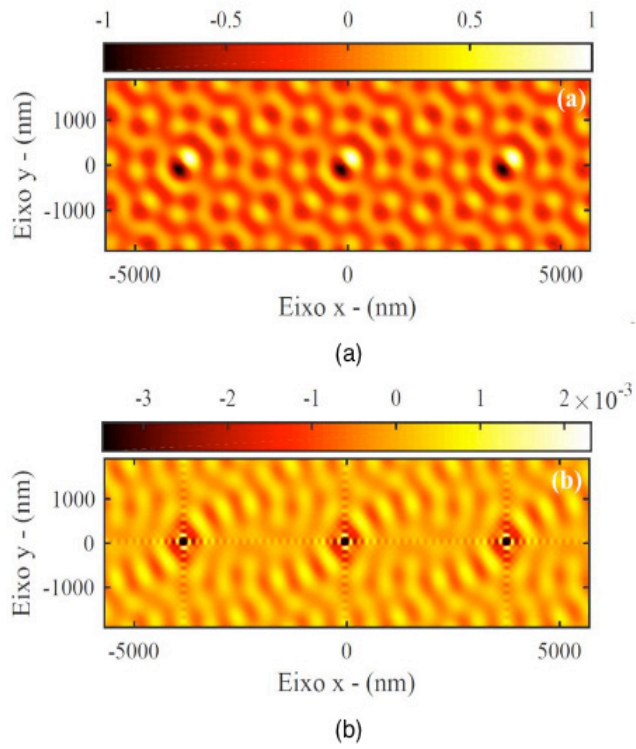


Figura 4 - Componentes de campos no plano xy em  $z = 0.15\lambda$ , polarização horizontal: (a) $Re\{E_z\}$ , (b) $Re\{H_y\}$ .

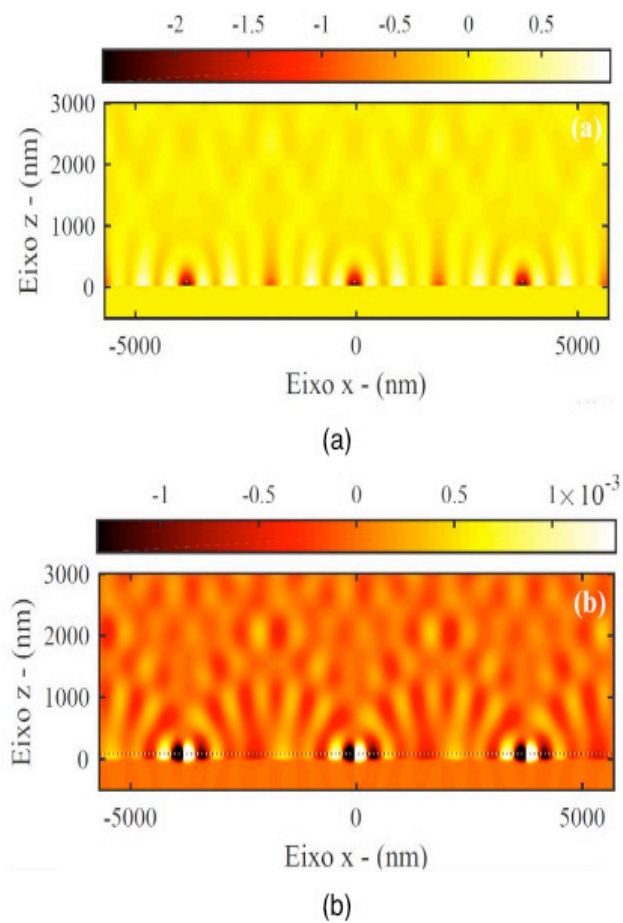


Figura 5 - Componentes de campos no plano xz em  $y=0$ , polarização vertical: (a) $Re\{E_z\}$ , (b) $Re\{H_y\}$ .

Primeiramente, verifica-se que as condições de contorno periódicas para

ambos os campos são atendidas. Nas Fig. 4 e Fig.5, os modos predominantes são, respectivamente, o TE e o TM. Note, na Fig. 5, o surgimento de ondas plasmônicas sobre a camada metálica, justamente devido à predominância do modo TM que é exigência necessária ao surgimento destas ondas (SOUZA, et al., 2017; MAIER, 2007) O aparecimento destas ondas é de grande interesse em aplicações de sensoriamento e em células fotovoltaicas devido a concentração da intensidade de luz localizada sobre a superfície dos metais.

Por fim, utilizando todas as componentes de E e H, os campos elétrico e magnético totais em uma célula de período  $2a = 6\lambda$  são demonstrados na Fig. 6. Na Fig. 7 é mostrada a visão do topo da célula, nesta as características rotacionais do campo magnético, e radiais do campo elétrico são observadas, corroborando com a teoria clássica destes campos (BALANIS, 2005).

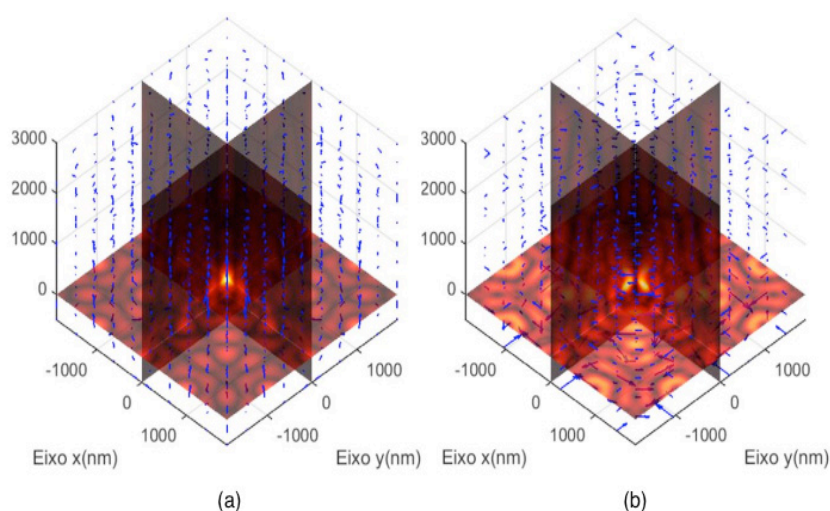


Figura 6 – Campo Total na Célula de Análise, visão tridimensional: (a) E; (b) H.

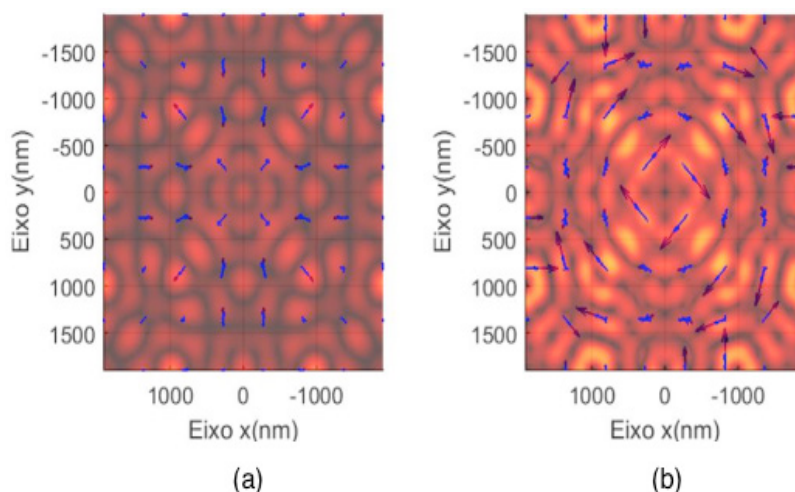


Figura 7 - Campo Total na Célula de Análise, visão do topo: (a) E; (b) H.

### 3.3 Diagramas de Radiação

Utilizando a definição de far-field (BALANIS, 2005), os diagramas de radiação



foram gerados e são demonstrados na Fig. 8. O espaçamento entre os nanodipolos no arranjo é calculado para  $a = b = 0.25\lambda$ ,  $0.5\lambda$  e  $\lambda$ . Estes são posicionados em  $h = \lambda/10$  e  $h = \lambda$ , e orientados na vertical e horizontal.

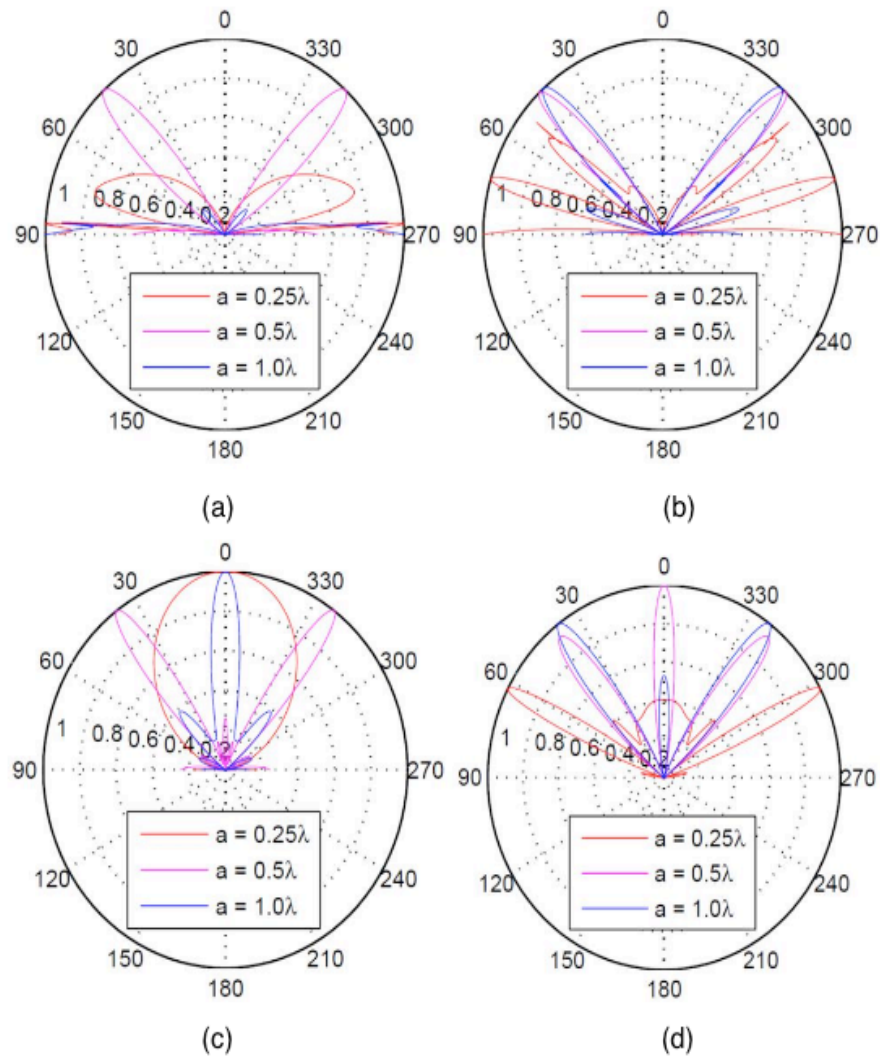


Figura 8 - Diagramas de Radiação com nanodipolos orientados na: Vertical: (a) ,

A partir das Fig. 8a e 8b, podemos verificar que quando os nanodipolos são orientados verticalmente, ( $\theta' = 0^\circ$ ), a radiação é concentrada na região da interface metal/ar. Esta concentração de ondas é proveniente da excitação de ondas SPP, e é amplificada à medida que aproximamos o arranjo da interface, confinando a radiação nesta região. No caso das Fig. 8c e 8d, quando os nanodipolos são orientados horizontalmente ( $\theta' = 90^\circ$ ) praticamente toda radiação é longitudinal e perpendicular ao arranjo. Os resultados demonstram a capacidade de controle da intensidade da luz em diferentes direções a partir da mudança da distribuição e inclinação do arranjo de nanodipolos.

### 3.4 Convergência do Método

Nos resultados de campos as séries duplas foram truncadas em  $M=N=20$  para Fig. 4,  $M=N=50$  para Fig.5. As simulações foram realizadas com razoável



custo computacional, aproximadamente 35 minutos e 40 segundos, contudo, fez-se necessário verificar a convergência do método e a real necessidade do número de termos nas séries. Na Fig. 9 é demonstrada o erro relativo para a componente  $A_z(0,0,z)$  em termos do número de iterações  $m \times n$ , para três pontos em  $z$ .

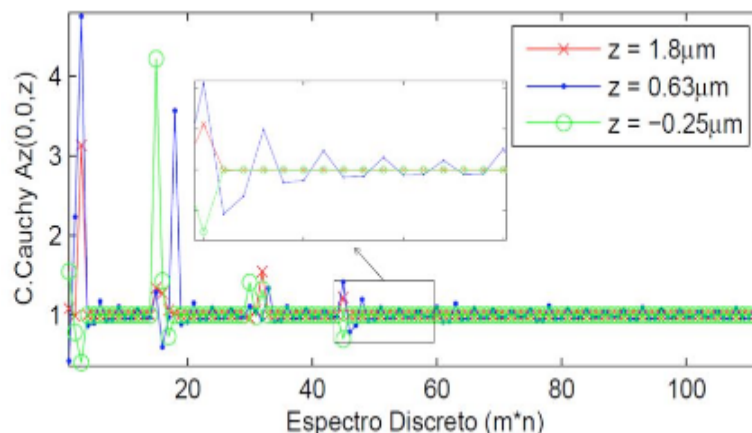


Figura 9 - Erro relativo da série cossenoidal na componente  $A_z(0,0,z)$ , onde  $z = 1.8\mu m$  (sobre a fonte),  $z = 0.63\mu m$  (próximo a fonte) e  $z = -0.25\mu m$  (abaixo da interface metal/ar).

Após verificar que a convergência do método ocorre em  $M=N=7$ , as simulações foram refeitas, demandando um tempo de 2 minutos e 16 segundos, sucedendo resultados praticamente iguais aos obtidos nas Fig. 4 e Fig. 5.

## 4 | CONCLUSÕES

Neste trabalho, foi apresentada a modelagem eletromagnética de um arranjo periódico planar de nanodipolos, sobre um plano de ouro, através do método da função de Green periódica, com o principal intuito de propor um modelo que possa ser utilizado na otimização de projetos de nanodispositivos. A partir dos resultados obtidos podemos verificar as características de campos próximo e radiado. Na Fig. 9, verificamos que a convergência das séries ocorre de formas diferentes para diferentes pontos no espaço, sendo muito rápida distante da fonte, e lenta nas proximidades, principalmente pelo fato do método tentar aproximar uma fonte delta de Dirac. A utilização da série dupla em termos de cosseno propicia a redução em quatro vezes o número de termos em relação a série complexa de Fourier convencional.

## REFERÊNCIAS

AHMADI, A; GHADARGHADR, S; MOSALLAEI, H. **An optical reflectarray nanoantenna: The concept and design**. Opt. Express 18, pp. 123-133, 2010.

BALANIS, C. A. **Antenna Theory: Analysis and Design. MICROSTRIP ANTENNAS**, third edition, John Wiley sons,(2005).

- BOHREN, C. F; Huffman, D. R. **Absorption and Scattering of Light by Small Particles**, Wiley: New York, 1983.
- CHEW, W. C; **Waves and Fields in inhomogeneous media**. IEEE press,(1995).
- COSTA, K. Q; COSTA J. S; DMTRIEV, V; DEL ROSSO, T; PANDOLI O; AUCELIO R. Q. **Analysis of surface plasmon resonance sensor coupled to periodic array of gold nanoparticles**, in {IMOC: SBMO/IEEE MTT-S International}, Porto de Galinhas, PE, Brazil, 2015, pp. 15.
- DUDLEY, D. G. **Mathematical Foundations for Electromagnetic Theory**, IEEE Press, 1994
- HOFMANN, H. F; KOSAKO, T; KADOYA, Y. **Design parameters for a nano-optical Yagi Uda antenna**. New Journal of Physics 9, 2007.
- MAIER, S. A. **Plasmonics: Fundamentals and Applications**, ed. New York, NY, USA: Springer, 2007, pp. 65--87.
- POHL, D. W; ALU, A; ENGHETA, N; MARQUIER, F; **Optical Antennas**. Forthcoming Publications: Science, PRL, 2013.
- RASHIDI, A; MOSALLAEI, H. **Array of plasmonic particles enabling optical near-field concentration: A non-linear inverse scattering design approach**, Phys. Rev. B, 82, 035117, 2010.
- RASHIDI, A; MOSALLAEI, H; Mittra, R. **Numerically efficient analysis of array of plasmonic nanorods illuminated by an oblique incident plane wave using the characteristic basis function method**, J. Comput. Theor. Nanosci. 10, pp 427-445, 2013.
- RODRÍGUEZ, C. P. J; MARTÍNEZ M. M; RODRÍGUEZ, F. F. J; TOMÁS, N. B; ORTUÑO, R; PERANSÍ, L. S; MARTÍNEZ, A. **Demonstration of near infrared gas sensing using gold nanodisks on functionalized silicon**. Opt. Express 19, pp. 7664-7672, 2011.
- ROSTAMI, M; D. ANSARI O.B; MOSALLAEI H. **Dispersion diagram analysis of arrays of multishell multimaterial nanospheres**, PIERB, 48, pp. 77-98, 2013.
- SANTOS, T. L. T. **Análise comparativa das propriedades ressonantes de nanopartículas e nanoantennas bowtie de ouro de diferentes geometrias**, 2012. Tese, Universidade Federal do Pará, Belém.
- SOUZA, N. W. P; COSTA, J. S; SANTOS R. C; CRUZ, A. F. S; DEL ROSSO T; COSTA K. Q. **Modal analysis of surface plasmon resonance sensor coupled to periodic array of core-Shell metallic nanoparticles**, Resonance, Prof. Jan Awrejcewicz (Ed.), InTech, 2017.
- SQALLI, O; UTKE, I; HOFFMANN, P; WEIBLE, F. M. **Gold elliptical nanoantennas as probes for near field optical microscopy**. J. of Appl. Physics, vol. 92, pp. 1078-1083, July 2002.
- TAJDINI, M. M ; MOSALLAEI, H. **Characterization of large array of plasmonic nanoparticles on layered substrate: dipole mode analysis integrated with complex image method**. Applied EM and Optics Laboratory Northeastern University, 2011.

## AVALIAÇÃO DE FADIGA MUSCULAR LOCALIZADA EM SINAIS ELETROMIOGRÁFICOS UTILIZANDO TAXA DE AMOSTRAGEM VARIÁVEL NO TEMPO

**Jean Kevyn Correia Pessoa**

Instituto de Educação Superior de Brasília – Iesb  
Brasília – Distrito Federal

**Pedro Henrique Melgaço de Oliveira  
Martins**

Instituto de Educação Superior de Brasília – Iesb  
Brasília – Distrito Federal

**Thiago Raposo Milhomem de Carvalho**

Instituto de Educação Superior de Brasília – Iesb  
Brasília – Distrito Federal

**RESUMO:** Este trabalho objetiva avaliar os padrões espectrais de sinais eletromiográficos obtidos durante indução de fadiga muscular aplicando-se sensoriamento compressivo com taxa de amostragem variável no tempo, propondo-se uma técnica de subamostragem sem uso de transformadas, próprias para sinais de contrações dinâmicas cíclicas. Assim, avaliaram-se os efeitos dessa técnica sobre os valores de frequência média dos ciclos de contração (parâmetro indicador da fadiga). A adequação da proposta foi avaliada pelo coeficiente de correlação ( $\rho$ ) entre os valores de frequência média obtidos com e sem a técnica, obtendo-se  $\rho \cong 0,9$  para uma taxa média de 250 Hz (75% inferior à de Nyquist).

**PALAVRAS-CHAVE:** Eletromiografia de superfície, Sensoriamento compressivo, Fadiga muscular localizada.

**ABSTRACT:** This study evaluates the spectral patterns of electromyography signals obtained during fatigue induction with compressive sensing on a time variant sampling rate. It proposes a subsampling technique without the use of transforms, designed for cyclical dynamic contraction signals. The effects of the proposed technique were evaluated on the mean frequency values of the dynamic contraction cycles (fatigue indication parameter). The proposed methodology was evaluated through the correlation coefficient ( $\rho$ ) between the mean frequency values obtained with and without the technique, obtaining  $\rho \cong 0,9$  for an average rate of 250 Hz (75% inferior to the Nyquist rate).

**KEYWORDS:** *Surface electromyography, compressive sensing, Localized muscle fatigue.*

### 1 | INTRODUÇÃO

A avaliação das funções musculares e a aferição objetiva dos fenômenos envolvidos na contração muscular constituem etapas essenciais ao entendimento do movimento humano e dos processos fisiológicos que ele envolve. Neste sentido, a eletromiografia é uma técnica importante para a monitoração das atividades musculares, realizada por meio da medição dos sinais elétricos emanados pelo músculo (MERLETTI R, PARKER PA, 2004).

Como modalidade não invasiva de aferição da atividade mioelétrica tem-se a eletromiografia de superfície (EMG-S), técnica já difundida em aplicações clínicas e científicas (MERLETTI R, PARKER PA, 2004; MILHOMEM TR, 2015). O uso da EMG-S é especialmente interessante quando o sinal eletromiográfico é obtido durante contrações musculares em protocolos experimentais que envolvem a indução da fadiga muscular localizada, como, por exemplo, na atividade de ciclismo (DELUCA CJ, 1984; MERLETTI R, PARKER PA, 2004; MILHOMEM TR, 2015). Nessas situações em que ocorre a fadiga muscular, avaliam-se, tradicionalmente, o comportamento dos parâmetros de amplitude e espectrais do sinal de EMG-S ao longo do tempo. Um dos parâmetros espectrais clássicos – utilizado também neste artigo – para avaliação da fadiga muscular é a frequência média (MNF), que, como parâmetro indicador da fadiga muscular, deve apresentar comportamento decrescente ao longo do tempo, realizando-se sua estimativa em janelas de comprimento adequadamente escolhido – e.g. via Transformada de Fourier de Tempo Curto (STFT) –, a depender da modalidade de contração (DELUCA CJ, 1984; MILHOMEM TR, 2015).

Em comparação a outros tipos de sinais – como voz, imagem, vídeo etc. –, na literatura científica disponível encontram-se relativamente poucos trabalhos dedicados especificamente a propostas de técnicas de compressão de sinais eletromiográficos. Em Berger, PA *et al.* (2008), Costa, MVC *et al.* (2009) e Trabuco, MH *et al.* (2018), os autores descrevem técnicas para compressão de sinais de EMG-S – quando obtidos em contrações cíclicas dinâmicas – em duas dimensões, explorando correlação bidimensional. Entretanto, para a exploração da correlação bidimensional no processo de compressão, conforme se descreve nesses trabalhos, há a necessidade de se ter o sinal eletromiográfico completo previamente adquirido. Em Trabuco, MH *et al.* (2014), os autores propõem um algoritmo de compressão utilizando decomposição em wavelets e alocação dinâmica de bits em sub-bandas. Em comum, todos esses trabalhos citados fazem uso de transformadas ortogonais – e da representação esparsa obtida no domínio transformado – no processo de compressão. No presente trabalho, faz-se uma abordagem diferente, formulando-se uma técnica implementada inteiramente no domínio do tempo. Além disso, a técnica proposta é formulada de forma que possa ser posteriormente implementada em tempo real – durante a aquisição do sinal de EMG-S.

Assim, neste trabalho são abordados sinais de EMG-S obtidos de ciclistas durante contrações dinâmicas cíclicas com indução da fadiga muscular localizada, propondo-se uma técnica de sensoriamento compressivo com taxa de amostragem variável no tempo sem o uso de transformadas, objetivando-se obter uma nova taxa média de amostragem inferior à de Nyquist – próxima a 1 kHz, para eletromiografia (HERMENS HJ *et al.*, 2000) –, de forma que a avaliação da fadiga muscular localizada via MNF não seja significativamente afetada. Assim, a técnica proposta descreve uma modalidade de sensoriamento compressivo para armazenamento do sinal de EMG-S que preserve as características espectrais do sinal de que depende a aferição objetiva da fadiga.

Desta maneira, a técnica proposta representa uma abordagem – no domínio do tempo e compatível com sua aplicação em tempo real – para a otimização dos recursos computacionais dedicados aos sinais eletromiográficos que, não raramente, apresentam longa duração (DELUCA CJ, 1984).

A metodologia utilizada compreende o protocolo experimental para aquisição dos sinais com ciclistas voluntários, a instrumentação dedicada utilizada, as técnicas de subamostragem – via taxa variável no tempo – proposta no trabalho e de avaliação da fadiga (respectivamente, nas seções 2-A, 2-B, 2-C e 2-D).

## 2 | MATERIAIS E MÉTODOS

### A. *Protocolo Experimental*

O procedimento experimental com os ciclistas voluntários foi realizado num protocolo de duas etapas. Na primeira etapa, após assinar termo de esclarecimento e consentimento do protocolo experimental, o voluntário realizava aquecimento no cicloergômetro sem resistência por 3 minutos. Imediatamente após o aquecimento, ele realizava um teste com velocidade fixa de 35 km/h. A potência de resistência à pedalada do cicloergômetro iniciava a 150 W – potência razoavelmente leve para ciclistas amadores – e era incrementada em 50 W a cada minuto, até a ocorrência da exaustão. Após 48 h, na segunda etapa, o ciclista pedalava a 70% da potência máxima atingida na primeira etapa com velocidade constante de 35 km/h, também até a exaustão. Este protocolo experimental foi previamente proposto e utilizado por Milhomem TR (2015) e Milhomem TR *et al.* (2014), sendo justificado pelos autores. Neste trabalho foram utilizados sinais coletados de 3 indivíduos do sexo masculino ( $26,3 \pm 4,0$  anos,  $77,0 \pm 6,5$  kg e  $176,5 \pm 6,3$  cm de altura). O estudo foi aprovado pelo Comitê de Ética da Faculdade de Ciências da Saúde da Universidade de Brasília (parecer nº 069/10).

### B. *Instrumentação para aquisição dos sinais*

Foi utilizado um cicloergômetro vertical da marca ErgoFit, onde foram acoplados pedais propostos em Carmo JC *et al.* (2001), com adaptações. Para o encaixe e fixação da sapatilha utilizou-se uma base de pedal disponível no mercado e, para a identificação dos ciclos de pedalada do ciclista, um *trigger* (sensor magnético) foi fixado no cicloergômetro de forma a ser ativado quando cada pedalada se completasse. Pois, segundo Milhomem TR *et al.* (2014), como a velocidade do teste foi aproximadamente constante e a ação conjunta das pernas funciona como um volante, considera-se o ângulo do pedivela num instante qualquer como sendo proporcional à quantidade de amostras capturadas neste intervalo de tempo.

O sinal de EMG-S foi captado do músculo vasto lateral utilizando um eletromiógrafo da marca Delsys, modelo Bagnoli-2. Esse equipamento possui eletrodos com pré-amplificação e filtro passa-faixa (20 Hz – 400 Hz) com ganho total de 60 dB. Observaram-

se as recomendações de assepsia e posicionamento de eletrodos do SENIAM (HERMENS HJ *et al.*, 2000). Os sinais de *trigger* e de EMG-S foram digitalizados à taxa de 2,4 kHz com um conversor A/D de 12 bits da marca *National Instruments*, modelo PCI 6024E (Austin-TX, USA). O processamento dos sinais foi realizado em ambiente MATLAB (versão 7.1), utilizando-se sua *toolbox* de processamento de sinais. Antes do processamento dos sinais de EMG-S, estes foram filtrados com um filtro passa-faixa de Butterworth de 4ª ordem, entre 20 Hz e 500 Hz para a redução dos artefatos de movimentação de eletrodo e demais ruídos. A figura 1 mostra um trecho do sinal eletromiográfico correspondente a 5 ciclos de pedalada, a correspondente ativação do *trigger* (A) e espectro desse trecho de sinal de EMG-S considerado (B). O espectro foi calculado via transformada rápida de Fourier sobre trechos de 1024 amostras – com duração correspondente de 427 ms, o suficiente para considerar um ciclo completo de contração.

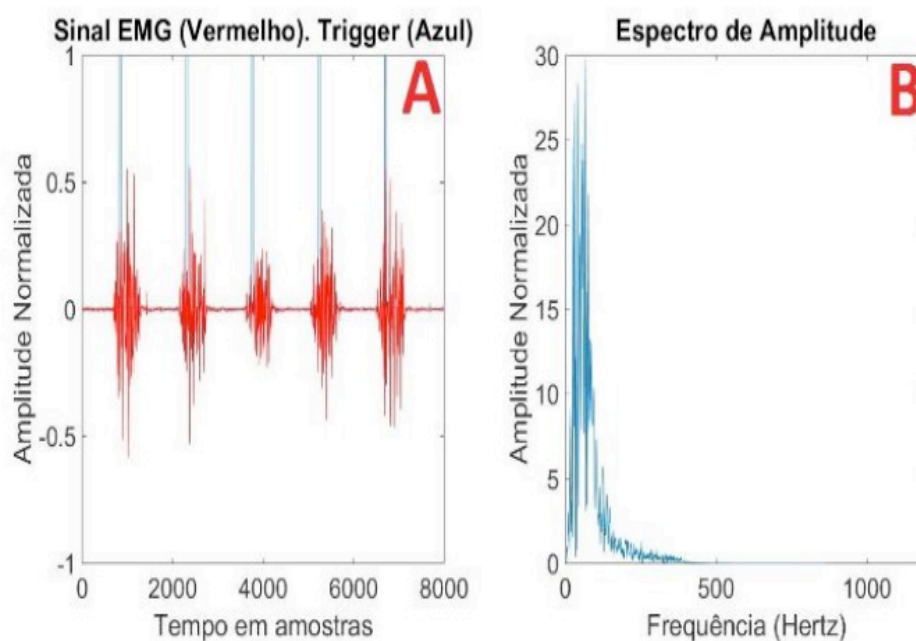


Fig. 1. (A) ciclos do sinal de EMG-S (vermelho) e sinal de *trigger* utilizado como referência (azul); (B) correspondente espectro de amplitude do sinal de EMG-S.

### C. Subamostragem do sinal de EMG-S

O sensoriamento compressivo via taxa de amostragem variável no tempo foi realizado calculando-se a energia local do sinal de EMG-S. Desta maneira, o método proposto toma a taxa de amostragem instantânea – a ser utilizada na subamostragem – com valor proporcional à amplitude local. A estimativa da amplitude local  $A[k]$  em torno da  $k$ -ésima amostra do sinal eletromiográfico foi definida como

$$A[k] = \sqrt{\frac{1}{N_g} \sum_{n=k-N_g+1}^k |x[n]g[k-n]u[k-n]|^2} \quad (1)$$



em que  $u[n]$  é a sequência degrau unitário e  $g[n]$  representa uma gaussiana em torno da origem com  $N_g$  coeficientes, dada por

$$g[n] = e^{-\frac{1}{2}\left(\frac{n}{\sigma_0}\right)^2} u[-n + N_g] \quad (2)$$

onde  $\sigma_0 = \frac{N_g - 1}{2\alpha}$ , sendo tomado o valor padrão  $\alpha = 2,5$ . Assim, a subamostragem instantânea após a  $k$ -ésima amostra do sinal de EMG-S original é proporcional a  $A[k]$ . Desta forma, a próxima amostra do sinal de EMG-S a ser tomada, após  $x[k]$ , é  $x[k + f(A[k])]$ , em que  $f(A[k])$  é uma função de amplitude local  $A[k]$ , dada por:

$$f(A[k]) = \frac{A_{m\acute{a}x}}{A[k]} \quad (3)$$

onde  $A_{m\acute{a}x}$  é uma estimativa para a amplitude local máxima que o sinal de EMG-S assume, tomada como a raiz da média quadrática (RMS) em torno do pico de amplitude do primeiro ciclo do sinal. Caso  $f(A[k])$  não resulte num inteiro, toma-se o inteiro mais próximo. Assim, o método impõe um atraso correspondente a um ciclo de pedalada (aproximadamente 0,5 s). A duração desse trecho para a estimativa de  $A_{m\acute{a}x}$  – em torno do pico de amplitude – foi tomado com a mesma duração da gaussiana  $g[n]$ , isto é,  $N_g$  amostras. O valor RMS de um trecho qualquer  $s[n]$  de  $N$  amostras é dado por

$$s_{RMS} = \sqrt{\frac{1}{N} \sum_{n=1}^N |s[n]|^2} \quad (4)$$

Assim, obtém-se uma taxa variável após a subamostragem, de forma que esta taxa seja maior em torno das regiões de maior amplitude do eletromiograma e menor nos trechos de menor amplitude e em que, portanto, não há ativação muscular significativa, contendo pouca informação. Desta maneira, obtém-se relação sinal-distorção (SDR) local maior onde há informação significativa no sinal, de forma que a SDR global não seja significativamente afetada após o processo de subamostragem com taxa variável. O cálculo da SDR (dB) para uma sequência qualquer  $s[n]$  corrompida com ruído aditivo  $r[n]$  e comprimento de  $N$  amostras se dá pela expressão

$$SDR_{dB} = 10 \log_{10} \left( \frac{\frac{1}{N} \sum_{n=0}^{N-1} |s[n]|^2}{\frac{1}{N} \sum_{n=0}^{N-1} |r[n]|^2} \right) \quad (5)$$

Após a subamostragem, o sinal é reconstruído utilizando interpolação de primeira ordem – por eficiência e simplicidade – seguida da aplicação de um filtro passa-baixa



para suavização do sinal. Esse filtro passa-baixa consiste de uma janela gaussiana de duração igual à utilizada no método para a estimativa da amplitude local  $A[k]$ .

Assim, ao aplicar o método descrito ao sinal de EMG-S original  $x[n]$ , de comprimento igual a  $N_x$  amostras obtém-se o sinal  $\bar{x}$ , reconstruído a partir das amostras tomadas nos trechos de significativa energia. Como a taxa de amostragem inicial utilizada para a simulação do método foi  $f_s = 2,4$  kHz, tem-se uma nova taxa média  $\bar{f}_s$  dada pela relação

$$\bar{f}_s = \frac{N_{\bar{x}}}{N_x} f_s \quad (6)$$

A SDR obtida com a aplicação do método é calculada via equação 5, tomando  $s[n] = x[n]$  e  $r[n] = x[n] - \bar{x}[n]$ .

#### D. Avaliação da fadiga

Para a avaliação da fadiga foi tomado como indicador o valor de frequência média (MNF) calculado sobre cada ciclo do sinal de EMG-S, conforme proposto em Milhomem TR (2015), explorando a característica aproximadamente ciclo-estacionária do sinal de EMG-S dinâmico durante contrações cíclicas realizadas a uma cadência constante (BONATO P *et al.*, 2001; HUG F, DOREL S., 2009; MILHOMEM TR, 2015). Para uma sequência qualquer de amostras, o índice da transformada discreta de Fourier (DFT) correspondente à MNF é dado pela relação

$$k_{MNF} = \sum_{k=0}^{M-1} k |S[k]|^2 / \sum_{k=0}^{M-1} |S[k]|^2 \quad (7)$$

em que  $S[k]$  é o  $k$ -ésimo coeficiente da DFT de  $s[n]$  e  $k_{MNF}$  é o índice correspondente à MNF. A partir da taxa de amostragem  $f_s$ , obtém-se a frequência média  $f_{média}$  em Hertz. Em contrações musculares em que há indução de fadiga e realizadas até a exaustão, espera-se obter uma sequência decrescente de valores para  $f_{média}$ , quando calculados sobre os ciclos do sinal de EMG-S (MILHOMEM TR, 2015). A fadiga muscular foi estimada, por meio da MNF, sobre: (1) o sinal de EMG-S original e (2) sobre o sinal obtido após a aplicação da técnica descrita de subamostragem; obtendo-se, assim, duas sequências de valores de MNF que descrevem o processo da fadiga muscular (i.e. com e sem aplicação da técnica de subamostragem).

Como figura de mérito, de modo a avaliar a fidelidade da avaliação da fadiga após o processo de subamostragem, foi calculado o coeficiente de correlação normalizado  $\rho$  entre ambas as sequências de valores de MNF, com e sem a aplicação da técnica de subamostragem, por meio da relação

$$\rho = \frac{1}{N-1} \sum_{i=1}^N \left( \frac{f_{m\u00e9dia}[i] - \mu}{\sigma} \right) \left( \frac{\overline{f_{m\u00e9dia}}[i] - \bar{\mu}}{\bar{\sigma}} \right) \quad (8)$$

em que  $f_{m\u00e9dia}[i]$  \u00e9 o valor de MNF calculado sobre o  $i$ -\u00e9simo ciclo do sinal de EMG-S original e  $\overline{f_{m\u00e9dia}}[i]$  \u00e9 o correspondente valor calculado sobre o sinal de EMG-S ap\u00f3s a t\u00e9cnica de subamostragem proposta;  $N$  \u00e9 a quantidade de ciclos de contra\u00e7\u00e3o realizados no experimento at\u00e9 que ocorresse a exaust\u00e3o;  $\mu$  e  $\sigma$  s\u00e3o a m\u00e9dia e o desvio padr\u00e3o de  $f_{m\u00e9dia}[i]$  e  $\bar{\mu}$  e  $\bar{\sigma}$  s\u00e3o a m\u00e9dia e o desvio padr\u00e3o de  $\overline{f_{m\u00e9dia}}[i]$  ( $i = 1, 2, \dots, N$ ).

Uma s\u00edntese do procedimento descrito neste trabalho \u00e9 ilustrada no diagrama da figura 2. Nesse diagrama tem-se como entrada o sinal de EMG-S coletado do ciclista volunt\u00e1rio, mostrando-se as etapas at\u00e9 a avalia\u00e7\u00e3o da fadiga sobre os sinais com e sem a t\u00e9cnica de subamostragem proposta e os respectivos par\u00e2metros utilizados para avaliar o m\u00e9todo – SDR obtida na reconstru\u00e7\u00e3o ap\u00f3s subamostragem, coeficiente de correla\u00e7\u00e3o entre as sequ\u00eancias indicadoras de fadiga (valores de MNF) e a nova taxa m\u00e9dia resultante, que indica o fator de compress\u00e3o obtido.

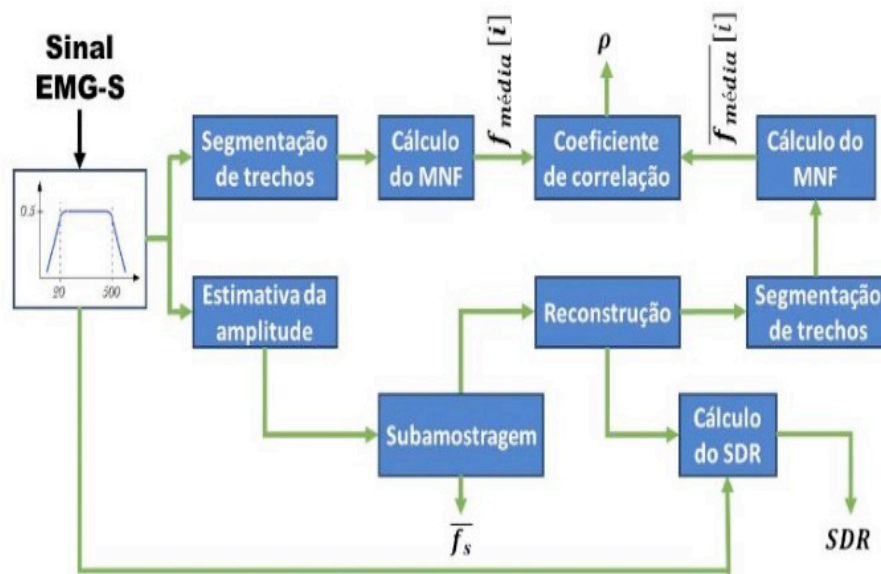


Fig. 2. Diagrama de blocos ilustrando a abordagem geral proposta.

### 3 | RESULTADOS E DISCUSS\u00c3O

A indu\u00e7\u00e3o da fadiga muscular localizada foi confirmada pelos valores de MNF obtidos. A figura 3 a seguir exemplifica o padr\u00e3o t\u00edpico de MNF ao longo do exerc\u00edcio, com o gr\u00e1fico para a sequ\u00eancia  $f_{m\u00e9dia}[i]$  obtido para um dos ciclistas volunt\u00e1rios (em azul) e respectiva reta de ajuste linear para a sequ\u00eancia.

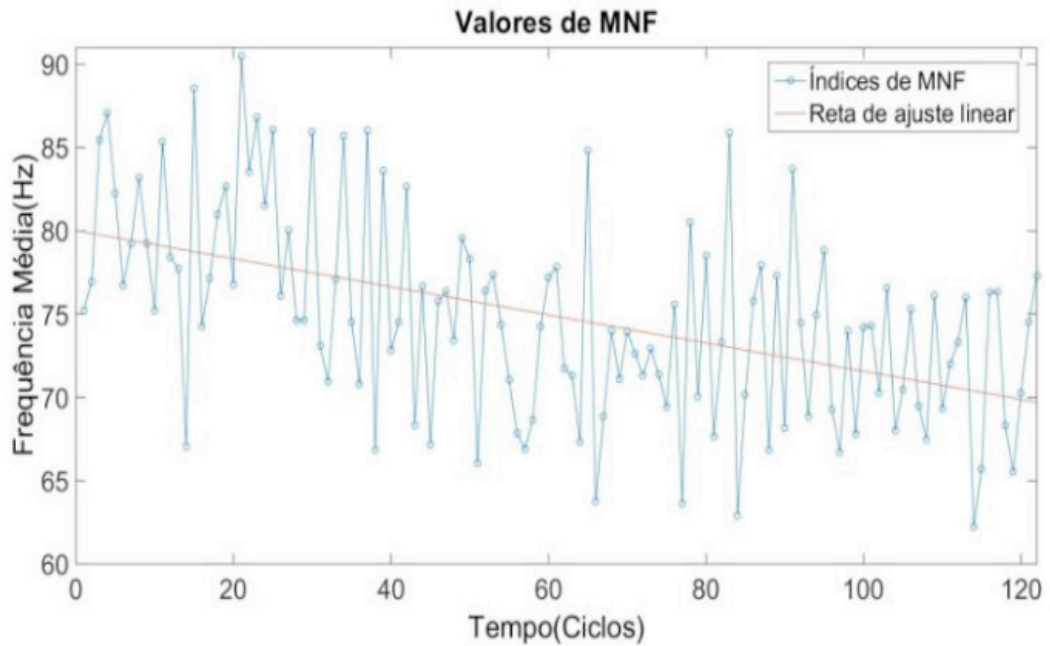


Fig. 3. Frequência média (azul) e reta de ajuste linear (vermelho).

A ocorrência da fadiga muscular é confirmada por meio do coeficiente negativo obtido no ajuste linear, que é descrito, após a regressão de primeira ordem, por uma equação da forma:

$$F(X) = aX + b \quad (9)$$

Assim,  $a < 0$  deve indicar o comportamento decrescente da MNF ao longo do tempo, sugerindo a indução da fadiga no exercício. Comumente toma-se como índice de fadiga a razão  $\frac{a}{b}$ , isto é, o decréscimo médio (indicado por  $a$ ) normalizado pelo valor inicial de frequência média, indicando-se uma diminuição percentual média da MNF. Tal razão, para os três voluntários em cujos sinais de EMG-S a técnica proposta foi aplicada, é mostrada na tabela 1.

Sinal (voluntário)	S1	S2	S3
Índice ( $\frac{a}{b} \times 10^{-4}$ )	-2,4	-2,5	-11,0

Tabela 1. Índices de fadigas obtidos.

As figuras 4 e 5 mostram, em azul, um trecho do sinal de EMG-S obtido de um dos voluntários, de um ciclo de contração (pedalada) e os correspondentes instantes de tempo (em vermelho) em que são selecionadas as amostras no processo de subamostragem proposto. Note-se que, nas regiões de maior amplitude do sinal, há maior densidade de amostras consideradas, tendo-se nessas regiões, portanto, maior

taxa instantânea na subamostragem (figura 4). Na figura 5, tem-se exemplo de trecho do sinal de EMG-S – correspondente a um ciclo de contração – antes (em azul) e após (em preto) a reconstrução utilizando o método proposto. A análise conjunta dos gráficos contidos nas figuras 4 e 5 permitem perceber que, assim, os trechos do sinal que contribuem com a SDR de forma mais significativa não são significativamente degradados.

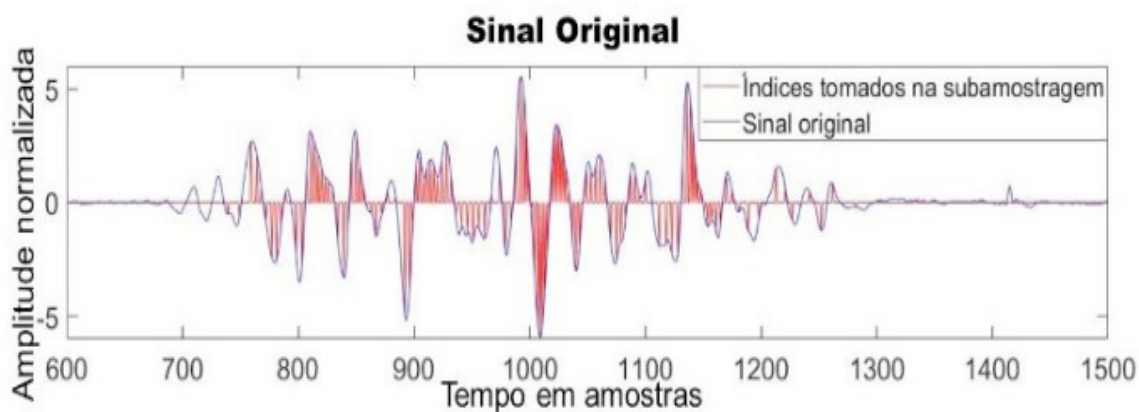


Fig. 4. Ciclo do sinal de EMG-S (azul) e amostras tomadas na subamostragem (vermelho);

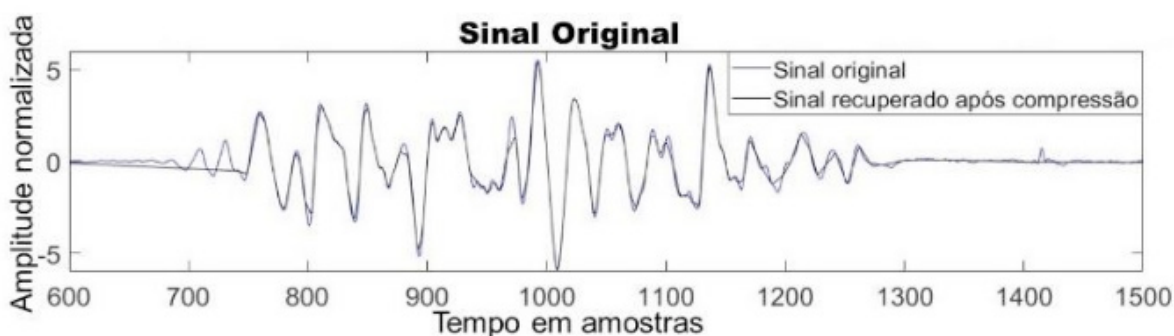


Fig. 5. Ciclo do sinal de EMG-S (azul) e sinal interpolado após subamostragem (preto).

A figura 6 ilustra um trecho do sinal de EMG-S (azul) em que a amplitude local apresenta aumento, ocasionando o aumento da taxa instantânea no processo de subamostragem (em vermelho).

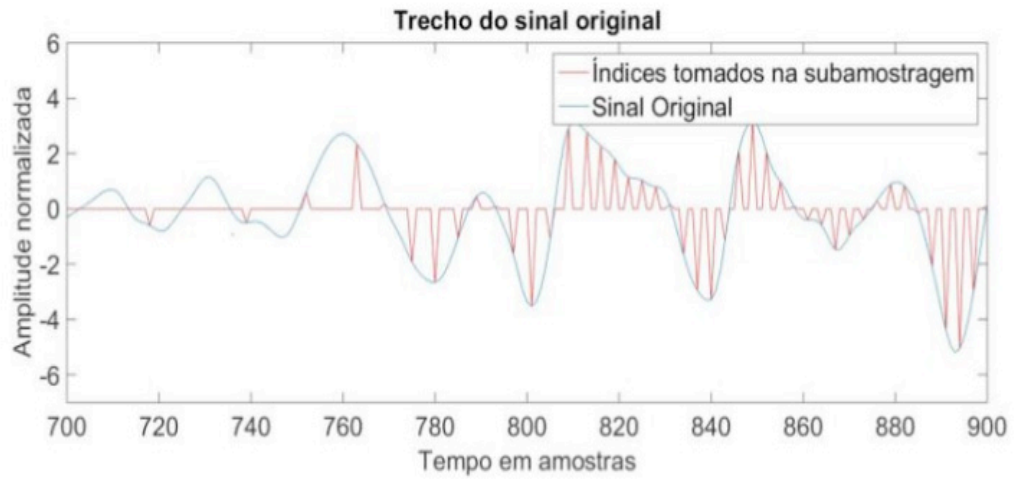


Fig. 6. Amostras tomadas num trecho do sinal de EMG-S com o método proposto.

A figura 7-A apresenta os valores obtidos para o coeficiente de correlação  $\rho$  entre  $\overline{f_{média}}[i]$  e  $f_{média}[i]$  em função da duração (quantidade  $N_g$  de amostras) da janela utilizada na estimativa da amplitude local do eletromiograma. Nota-se que os valores de correlação mais próximos da unidade foram obtidos com janelas de 400 a 800 amostras, aproximadamente, com durações aproximadas correspondentes entre 170 ms e 330 ms. A figura 7-B mostra esses valores de correlação obtidos em função da taxa média  $\overline{f_s}$  resultante da subamostragem. Verifica-se que, para taxas resultantes entre 250 Hz e 450 Hz – taxas inferiores à metade da de Nyquist –, os parâmetros espectrais não foram significativamente alterados ( $\rho \cong 1$ ).

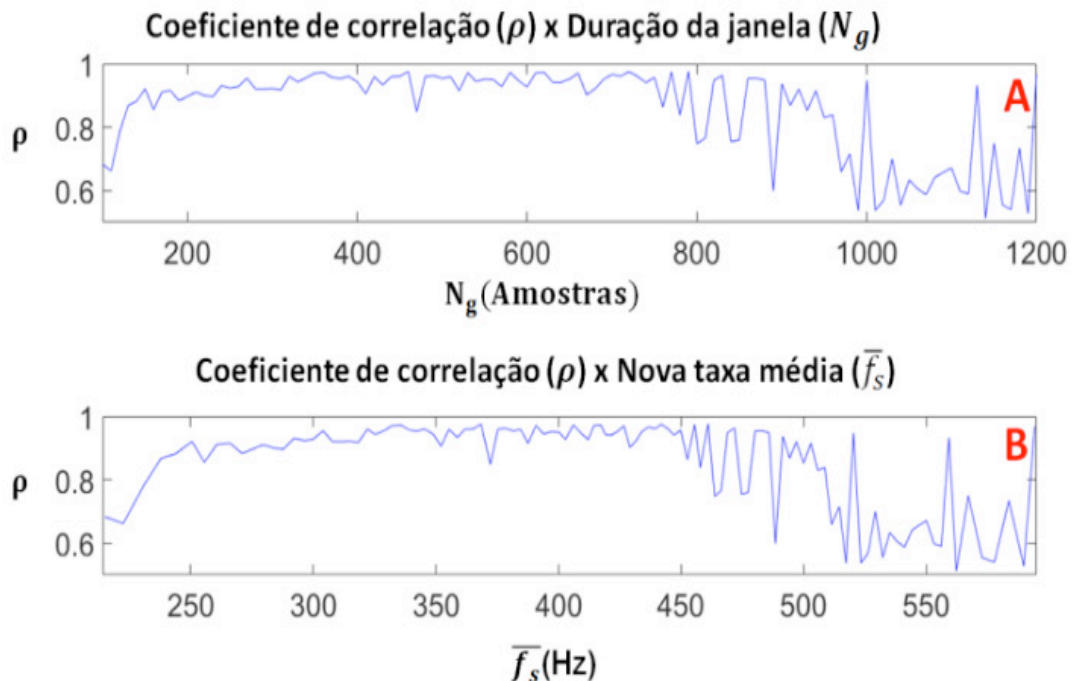


Fig. 7. (A) Coeficiente de correlação ( $\rho$ ) em função da duração de janela ( $N_g$ ); (B) Coeficiente de correlação em função da nova taxa média ( $\overline{f_s}$ ) obtida após a subamostragem.

A figura 8 mostra, também em função da duração  $N_g$  da janela para estimativa

da amplitude local e da nova taxa média  $\bar{f}_s$  obtida com a subamostragem, os valores de SDR (dB) obtidos na reconstrução do sinal de EMG-S a partir de suas amostras mantidas no processo de subamostragem.

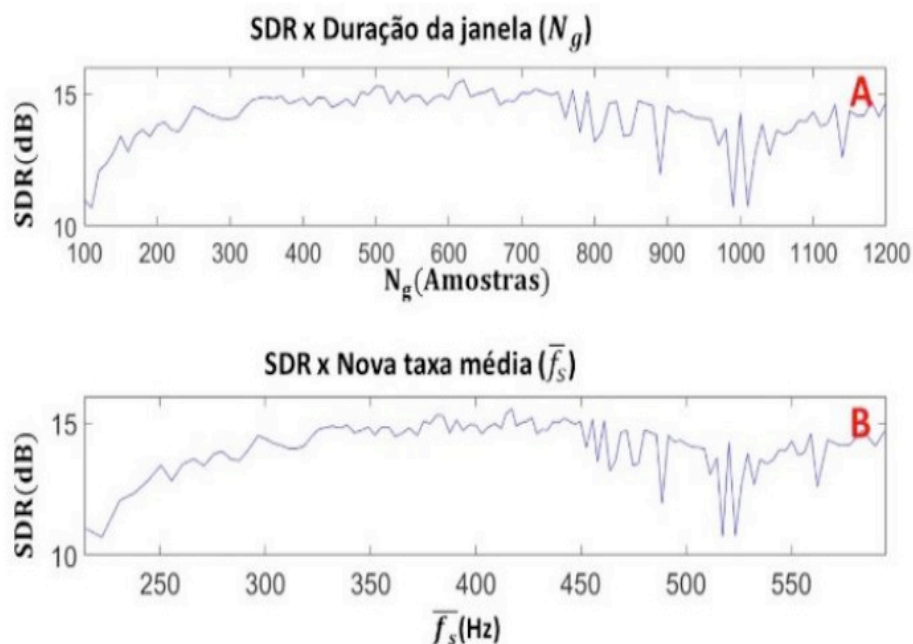


Fig. 8. (A) SDR em função da duração de janela ( $N_g$ ); (B) SDR em função da Nova taxa média ( $\bar{f}_s$ ) obtida após subamostragem.

Confirma-se que os melhores resultados foram obtidos para  $400 \leq N_g \leq 800$  (cujas durações correspondentes são 170 ms e 330 ms, aproximadamente) e para  $\bar{f}_s$  entre 250 Hz e 450 Hz.

Possíveis razões para esses resultados residem no fato de que, ao utilizar janela  $g[n]$  de duração muito menor do que 400 amostras para a  $k$ -ésima estimativa de amplitude local,  $A[k]$ , tem-se uma estimativa excessivamente sensível às variações locais devidas a não estacionariedades rápidas, que podem corresponder, por exemplo, à variação da posição do eletrodo para aquisição da EMG-S, que altera a forma local do sinal mioelétrico (MERLETTI R, PARKER PA, 2004; MILHOMEM TR, 2015). Além disso, o uso da janela com duração demasiadamente pequena pode causar problemas em torno de picos mais elevados no sinal de EMG-S ocorridos durante o recrutamento de novas unidades motoras – o que ocorre no processo de fadiga muscular localizada (DELUCA CJ, 1984; MERLETTI R, PARKER PA, 2004) – e em torno de cruzamentos no zero dentro do trecho de ativação muscular dos ciclos do sinal de EMG-S. Por outro lado, durações muito grandes para  $g[n]$  devem ocasionar excessiva demora na atualização da taxa de amostragem. Neste caso, ter-se-ia um aumento demasiadamente lento na taxa de amostragem instantânea no início da ativação muscular – trecho do sinal de EMG-S em que a taxa deve ser aumentada – e uma diminuição também excessivamente lenta após o final da ativação muscular



– onde a taxa deve ser diminuída – nos ciclos de contração. Tais fatores, resultantes de escolhas de valores muito grandes para  $N_g$ , implicam, respectivamente: 1) no aumento do erro de reconstrução do sinal logo que se inicia a ativação muscular, e 2) no desnecessário uso de maiores taxas de amostragem instantâneas após cessar a ativação. Assim, empiricamente, verificam-se melhores resultados para  $400 \leq N_g \leq 800$ .

Desta maneira, obtêm-se os melhores resultados com as correspondentes novas taxas médias  $\bar{f}_s$  entre os valores citados de 250 Hz e 450 Hz, às quais correspondem os fatores de compressão iguais a, aproximadamente, 89,6% e 81,3%.

## 4 | CONCLUSÕES

Neste trabalho foi proposto um método de sensoriamento compressivo utilizando taxa de amostragem variável no tempo sem o uso de transformadas, de forma que tal sensoriamento compressivo possa ser realizado a taxas inferiores à de Nyquist sem significativa degradação dos parâmetros espectrais que indicam a fadiga muscular localizada no sinal de EMG-S.

Os resultados obtidos indicam que foi possível a redução taxa de maneira a possibilitar fatores de compressão de até, aproximadamente, 10 vezes, considerando-se a taxa de amostragem original de aquisição utilizada de 2,4 kHz, mantendo-se a relativa integridade dos valores de MNF do eletromiograma.

Assim, o método proposto consiste numa abordagem direcionada à otimização dos recursos computacionais dedicados ao tratamento de sinais de EMG-S, que, em função do protocolo experimental, podem apresentar durações excessivamente longas.

Ainda, o método, aplicável em tempo real a sinais de EMG-S obtido durante contrações cíclicas, explora este aspecto e apresenta relativa simplicidade por seu tratamento no domínio do tempo, evitando-se a busca de domínios de representação esparsa para o sinal, como já tradicionalmente feito nas técnicas de sensoriamento compressivo em geral.

## 5 | AGRADECIMENTOS

Os autores agradecem aos ciclistas voluntários que participaram dos protocolos experimentais e aos demais que auxiliaram no processo da coleta dos sinais.

## REFERÊNCIAS

Berger, PA; Nascimento, FAO; Menezes, LRAX.; Rocha, AF; Carvalho, JL. “**Biomedical Signal Compression: An Overview**”. In: Nilmini Wickramasinghe; Eliezer Geisler. (Org.). Encyclopedia of Healthcare Information Systems. Hershey - New York: IGI Global, 2008, v. 1, p. 157-1165.

Bonato P, Roy SH, Knaflitz M, deLuca CJ, (2001). “**Time-frequency parameters of de surface myoelectric signal for assessing muscle fatigue during cyclic dynamic contractions.**” IEEE Transactions on Biomedical Engineering, vol. 48, nº 7, pp. 745-753.



Carmo JC, Nascimento FAO, Costa JC e Rocha AP. **Instrumentação para aquisição e avaliação das forças exercidas nos pedais por ciclistas**. Revista Brasileira de Biomecânica. 2001; 2(3):31-37.

Costa, MVC; Carvalho, JL; Berger, PA; Rocha, AF; Nascimento, FAO. **“Compression of Surface Electromyographic Signals Using Two-Dimensional Techniques. Recent Advances in Biomedical Engineering”**. 1ed.Vienna: In-Tech, 2009, v. 1, p. 17-38.

DeLuca CJ (1984). **“Myoelectrical Manifestations of Localized Muscular Fatigue in Humans”**, Crit. Rev. Biomed. Eng. (1984), v. 11, n. 4, pp. 251-79.

Hermens HJ, Freriks B, Disselhorst-Klug C, Rau G. **Development of recommendations for SEMG sensors and sensor placement procedures**. Journal of Electromyography and Kinesiology. 2000; 10(5):361-74.

Hug F, Dorel S. (2009); **“Electromyographic analysis of pedaling: A review”**; Journal of Electromyography and Kinesiology 19 (2009) 182–198.

Merletti R, Parker PA (2004) **“Electromyography, physiology, engineering, and noninvasive applications”**, New Jersey: John Wiley & Sons, Inc.

Milhomem TR **“Proposta e análise de indicadores eletromiográficos espectrais e temporais em protocolos dinâmicos com indução de fadiga muscular localizada”**, Tese (doutorado)— Universidade de Brasília, Faculdade de Tecnologia, Departamento de Engenharia Elétrica, 2015

Milhomem TR, Rocha Junior VA, Carmo JC, Nascimento FAO. (2014) **“Desempenho em Ciclismo: Instrumentação e Parâmetros Biomecânicos e Eletromiográficos.”** In: XXIV Congresso Brasileiro de Engenharia Biomédica, 2014, Uberlândia. Anais do XXIV CBEB, v. 24. pp. 809-812.

Trabuco, MH; Costa, MVC.; Macchiavello, B; Nascimento, FAO. **“S-EMG Signal Compression in One-Dimensional and Two-Dimensional Approaches”**. In: IEEE Journal of Biomedical and Health Informatics, v. 22, p. 1104-1113, 2018.

Trabuco, MH; Costa, MVC; Nascimento, FAO. **“S-EMG signal compression based on domain transformation and spectral shape dynamic bit allocation”**. In: Biomedical Engineering Online (Online), v. 13, p. 22-47, 2014.

## **SOBRE O ORGANIZADOR**

**Henrique Ajuz Holzmann** - Professor da Universidade Tecnológica Federal do Paraná (UTFPR). Graduação em Tecnologia em Fabricação Mecânica e Engenharia Mecânica pela Universidade Tecnológica Federal do Paraná. Mestre em Engenharia de Produção pela Universidade Tecnológica Federal do Paraná Doutorando em Engenharia e Ciência do Materiais pela Universidade Estadual de Ponta Grossa. Trabalha com os temas: Revestimentos resistentes a corrosão, Soldagem e Caracterização de revestimentos soldados.

Agência Brasileira do ISBN  
ISBN 978-85-7247-449-8

