

UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

**META-HEURÍSTICA PARA MAPEAMENTO BBU-RRH
E BALANCEAMENTO DE CARGA ENTRE BBUS,
APLICADA A REDES DE ACESSO CENTRALIZADO**

RITA DE CÁSSIA PORFÍRIO DA CUNHA

DM: 24/2022

UFPA / ITEC / PPGEE
Campus Universitário do Guamá
Belém-Pará-Brasil
08/2022

UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

META-HEURÍSTICA PARA MAPEAMENTO BBU-RRH
E BALANCEAMENTO DE CARGA ENTRE BBUS,
APLICADA A REDES DE ACESSO CENTRALIZADO

RITA DE CÁSSIA PORFÍRIO DA CUNHA

UFPA / ITEC / PPGEE
Campus Universitário do Guamá
Belém-Pará-Brasil
08/2022

UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

RITA DE CÁSSIA PORFÍRIO DA CUNHA

META-HEURÍSTICA PARA MAPEAMENTO BBU-RRH E
BALANCEAMENTO DE CARGA ENTRE BBUS, APLICADA A
REDES DE ACESSO CENTRALIZADO

Dissertação submetida à Banca Examinadora do Programa de Pós-Graduação em Engenharia Elétrica da UFPA para a obtenção do Grau de Mestre em Engenharia Elétrica na área de Concentração em Computação Aplicada.

Orientador: Prof. Dr. Diego Lisboa Cardoso

UFPA / ITEC / PPGEE
Campus Universitário do Guamá
Belém-Pará-Brasil
08/2022

**Dados Internacionais de Catalogação na Publicação (CIP) de acordo com ISBD
Sistema de Bibliotecas da Universidade Federal do Pará
Gerada automaticamente pelo módulo Ficat, mediante os dados fornecidos pelo(a) autor(a)**

C972m CUNHA, RITA DE CASSIA PORFIRIO.
META-HEURÍSTICA PARA MAPEAMENTO BBU-RRH E
BALANCEAMENTO DE CARGA ENTRE BBUS, APLICADA A
REDES DE ACESSO CENTRALIZADO / RITA DE CASSIA
PORFIRIO CUNHA. — 2022.
75 f. : il. color.

Orientador(a): Prof. Dr. Diego Lisboa Cardoso
Dissertação (Mestrado) - Universidade Federal do Pará,
Instituto de Tecnologia, Programa de Pós-Graduação em
Engenharia Elétrica, Belém, 2022.

1. 5G. 2. C-RAN. 3. Alocação de Recursos. 4. DPSO. 5.
Balanceamento BBU-RRH. I. Título.

CDD 004.6782

UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

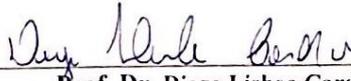
**“META-HEURÍSTICA PARA MAPEAMENTO BBU-RRH E BALANCEAMENTO DE CARGA
ENTRE BBUS, APLICADA A REDES DE ACESSO CENTRALIZADO”**

AUTORA: RITA DE CÁSSIA PORFÍRIO DA CUNHA

DISSERTAÇÃO DE MESTRADO SUBMETIDA À BANCA EXAMINADORA APROVADA PELO
COLEGIADO DO PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA, SENDO
JULGADA ADEQUADA PARA A OBTENÇÃO DO GRAU DE MESTRE EM ENGENHARIA
ELÉTRICA NA ÁREA DE COMPUTAÇÃO APLICADA.

APROVADA EM: 19/08/2022

BANCA EXAMINADORA:



Prof. Dr. Diego Lisboa Cardoso

(Orientador – PPGEE/UFPA)

Prof. Dr. Marcelino Silva da Silva

(Avaliador Interno – PPGEE/UFPA)



Prof. Dr. Jorge Antonio Moraes de Souza

(Avaliador Externo - UFRA)

VISTO:

Prof. Dr. Carlos Tavares da Costa Júnior

(Coordenador do PPGEE/ITEC/UFPA)

AGRADECIMENTOS

Neste momento é difícil agradecer todas as pessoas que de algum modo fizeram ou fazem parte da minha vida. Foi um momento de muita dedicação e comprometimento, repleto de desafios e oportunidades, aonde fui em busca de mais uma importante etapa em minha vida.

Primeiramente agradeço a Deus, por proporcionar todas as oportunidades que tive e que tornaram minha vida excelente, com uma família maravilhosa, amigos sinceros e, ainda, por renovar a cada momento a minha força e disposição ao longo dessa jornada.

Aos meus pais, Raimundo [in memoriam] e Maria Deolinda Porfirio, muito obrigado por toda educação e valores que sempre me deram. Pelo amor, incentivo e apoio incondicional em todos os momentos. Sempre foram a minha base e tudo o que sou!

Aos maiores presentes que recebi na vida, meus filhos, Hirla, Haroldo Filho e Heitor, sempre me incentivando, motivando e apoiando nos mais variados momentos. Foi Deus que, num momento lindo, me deu vocês e por vocês faço o impossível!

Ao Reginaldo Cavalcante Mesquita Junior, meu namorado. Obrigada pela sua incansável boa vontade em me ajudar e me incentivar a prosseguir, por todo seu carinho, atenção e, principalmente, a sua parceria em todas as horas, vibrando com as minhas conquistas.

À minha família, por todo incentivo e apoio durante esses anos, que nunca mediram esforços para me ajudar.

Agradeço a todos os meus professores que de diversas maneiras tiveram uma imensa contribuição na minha formação, mas principalmente ao meu orientador, Prof. Dr. Diego Lisboa Cardoso, que através de sua orientação, competência, profissionalismo dedicação e incentivo, me guiou e proporcionou o suporte necessário para o desenvolvimento dessa dissertação.

Agradeço aos amigos que fiz no LPO, por todo companheirismo no decorrer de toda essa jornada: Daniel, Carlos, Igor, Paulo Henrique, Lucivaldo, Paulo Tássio, Nilze, Ermínio, Áurea, Welton, Ladislav, Matheus Leto e, em particular, ao Rafael Fogarolli e a Mariane Gonçalves, que são amigos especiais e, por isso, eu sempre estarei unida a cada um de vocês e vocês sempre estarão presentes em minha vida.

Por fim, a todos aqueles que contribuíram direta ou indiretamente, para a realização desta dissertação, os meus sinceros agradecimentos.

*Bem-aventurado o homem que acha sabedoria, e o
homem que adquire conhecimento.*

(Provérbios 3:13)

SUMÁRIO

1	Introdução	1
1.1	Motivações	3
1.2	Objetivos	5
1.3	Organização do texto	6
2	Redes Móveis	7
2.1	Redes móveis de quinta geração (5G)	7
2.2	Rede de acesso centralizado (C-RAN)	10
2.2.1	Tipos de arquiteturas C-RAN.....	13
2.2.2	Benefícios de arquiteturas C-RAN.....	15
2.3	Rede Definida por Software [SDN] e Virtualização de Funções de Rede (NFV)	17
2.3.1	SDN	17
2.3.2	NFV	20
4.1 2.4	Considerações finais do capítulo.....	22
3	Meta Heurística.....	23
3.1	Inteligência Computacional (IC)	23
3.1.1	Meta-heurística	23
3.1.2	Algoritmo de Otimização por Enxame de Partículas (PSO)	25
3.1.3	Método da Otimização Discreta por Nuvem de Partículas (DPSO).....	30
3.1.4	Função Aptidão	31
3.1.5	Otimização Bio-Inspirada.....	31
4.2 3.3	Considerações finais do capítulo.....	31
4	Trabalhos Relacionados.....	32
4.1	Pesquisas que abordam soluções para a otimização do balanceamento de carga em BBU's	32
4.2	Considerações finais do capítulo	38
5	OTIMIZAÇÃO BIO-INSPIRADA PARA MAPEAMENTO BBU-RRH E BALANCEAMENTO DE CARGAS NAS BBU's	39
4.1	Arquitetura do sistema	39
4.2	Modelagem BBU-RRH	40
5.1.1	KPI.....	40
5.1.2	Função Objetivo	42
5.1.3	Benchmark e Desvio Padrão.....	43
5.2	Simulação do modelo.....	44
5.3	Algoritmo DPSO	46

5.4 Considerações finais do capítulo.....	48
6 Análise dos resultados	49
6.1 Resultados	49
6.2 Considerações finais do capítulo	53
7 Conclusão	54
7.1 Contribuições	55
7.2 trabalhos futuros.....	55
8 Referências	57

LISTA DE ILUSTRAÇÕES

Figura 1	Dispositivos x tráfego	1
Figura 2	Flutuação da carga de tráfego na BS durante o período de um dia	4
Figura 3	Evolução da Comunicação Móvel	7
Figura 4	Arquitetura Simplificada de uma Rede 5G	9
Figura 5	Arquitetura C-RAN tradicional	12
Figura 6	Arquitetura C-RAN Totalmente Centralizada	14
Figura 7	Arquitetura C-RAN Parcialmente Centralizada	15
Figura 8	Modelos de Rede Convencional e de Rede Definida por Software	18
Figura 9	Arquitetura SDN	19
Figura 10	Virtualização do BBU pool em C-RAN	21
Figura 11	Movimento de uma partícula	27
Figura 12	Topologias do PSO	28
Figura 13	Processo de execução do PSO	29
Figura 14	Conceito genérico de uma auto-organização em C-RAN.....	39
Figura 15	Demonstração do cenário proposto	42
Figura 16	Fluxograma da implementação do DPSO usando função objetivo.....	47
Figura 17	Nível de Balanceamento - Medida de Dispersão	53

LISTA DE TABELAS

Tabela 1	Síntese dos Principais Trabalhos Relacionados	37
Tabela 2	Parâmetros do DPSO	45
Tabela 3	Número médio de usuários (300 usuários)	49
Tabela 4	Número médio de usuários (600 usuários)	50
Tabela 5	Número médio de usuários (900 usuários)	50
Tabela 6	Número médio de usuários (1200 usuários),.....	50
Tabela 7	Número médio de usuários (1500 usuários),.....	51
Tabela 8	Número médio de usuários (1800 usuários)	51

LISTA DE ABREVIATURAS E SIGLAS

1G	Primeira Geração Móvel
2G	Segunda Geração Móvel
3G	Terceira Geração Móvel
4G	Quarta Geração Móvel
5G	Quinta Geração Móvel
ANATEL	Agência Nacional de Telecomunicações
AMPS	Sistema avançado de telefonia móvel
BBUs	Baseband Unit
BS	Base Station
CAPEX	Capital Expenditure
CDI	Cell Differentiation and Integration
CRAN	Cloud Radio Access Network
DPSO	Discrete particle swarm Optimization
DPS	Dynamic Point selection
EB	Exabytes
EE	Energy Efficiency
EDGE	Enhanced Data Rates For GSM Evolution
EUA	Estados Unidos da America
FDMA	Frequency Division Multiple Access
FIFO	First in, first out
GPRS	General Packet Radio Services
GSM	Global System for Mobile Communication
H-CRAN	Heterogeneous Cloud Radio Access Network
IA	Inteligência Artificial
IC	Inteligência Computacional
KPIs	Key Performance Indicator
LN	Lógica Nebulosa
LTE	Long Term Evolution
MATLAB	MATrix LABoratory
MOPTS	Million Operations Per Timeslot
NS-3	Network Simulato
OPEX	Operational Expenditure

OFDMA	Orthogonal Frequency Division Multiple Access
PSO	Optimization Particle Swarm
QoS	Quality of Service
RAT	Radio Access Technology
RNA	Redes Neurais Artificiais
RRH	Remote Radio Head
RSL	Received Signal Level
SE	Spectral Efficiency
SCNs	Small Cell Networks
SON	Self-Optimising Network
SOCRAN	Self-optimized Cloud Radio Access Network
SNR	signal-to-noise ratio
SDN	Software Defined Network
SUI	Stanford University Interim
TACS	Sistema de comunicação de acesso total
TCO	Cost of Ownership
TDMA	Time Division Multiple Access
UBs	User Bases
UE	User equipment
UMTS	sistema móvel de telecomunicações universais
W-CDMA	acesso múltiplo por divisão de código em banda larga
Wifi	Wireless Fidelity

RESUMO

O crescimento da demanda por acesso à informação, gerado por aplicativos multimídia, é um dos desafios da nova geração de redes móveis. A quinta geração (5G) visa atender requisitos cada vez mais estritos dos usuários, tais como latências e baixo consumo de energia. Uma das arquiteturas proposta para suprir as demandas que surgem com o 5G e suportar esse tráfego é a *Cloud Radio Access Network (C-RAN)*, a qual centraliza o poder de processamento para resolver o desbalanceamento de carga, alocando recursos de acordo com a demanda da rede. Essa arquitetura propõe o compartilhamento de recursos enquanto aborda questões de escalabilidade de processamento. Recentemente, os algoritmos de otimização meta-heurísticos vêm sendo amplamente empregados para resolver problemas dessa natureza. As meta-heurísticas são utilizadas por serem mais poderosas que os métodos convencionais, que se baseiam nas lógicas formais ou na programação matemática, além de o tempo necessário para execução ser menor que o dos algoritmos exatos. Neste contexto, objetiva-se com este estudo, desenvolver um modelo de alocação de recursos otimizado que realiza o balanceamento de carga entre *Baseband Units (BBUs)* e *Remote Radio Heads (RRHs)*, baseado em método de Otimização por Enxame de partículas (PSO). Para este fim, foi usado uma variação do algoritmo PSO, o *Discrete Particle Swarm Optimization (DPSO)*, que otimiza a função objetivo proposta. Resultados apontam desempenho superior desta função objetivo em comparação ao benchmarking utilizado, tanto em cenários de alta, como em baixa densidade de tráfego.

PALAVRAS-CHAVE: 5G, C-RAN, Alocação de Recursos, DPSO, Balanceamento BBU-RRH.

ABSTRACT

The growing demand for information access, generated by multimedia applications, is one of the challenges of the new generation of mobile networks. The fifth generation (5G) aims to meet increasingly stringent user requirements, such as latencies and low power consumption. One of the proposed architectures to supply the demands that arise with 5G and to support this traffic is the Cloud Radio Access Network (C-RAN), which centralizes processing power to solve the load imbalance, allocate resources accordingly based on network demand. This architecture proposes resource sharing while addressing processing scalability issues. Recently, metaheuristic optimization algorithms have been widely used to solve problems of this nature. Meta-heuristic algorithms are used because they are more powerful than conventional methods, which are to on formal logic or mathematical programming, in addition to the fact that the time required for execution is less than the exact algorithms' one. In this context, the objective of this study is to develop an optimized resource allocation model that performs load balancing between Baseband Units (BBUs) and Remote Radio Heads (RRHs), based on the Particle Swarm Optimization (PSO) method. For this purpose, a variation of the PSO algorithm, the Discrete Particle Swarm Optimization (DPSO) was used, to optimize the proposed objective function. Results indicated a point to superior performance of this objective function in comparison to the adopted benchmarking, both in high and low traffic densities.

Keywords: 5G, C-RAN, Resource Allocation, DPSO, Balancing BBU-RRH.

1 INTRODUÇÃO

A comunicação móvel, nos últimos 50 anos, está em constante transformação, buscando atender as diversas necessidades do mercado, onde cada década trouxe uma nova tecnologia móvel com inovações revolucionárias. Os requisitos de largura de banda móvel evoluíram das chamadas de voz e mensagens de texto ao vídeo em alta.

As projeções globais de conexão móvel para 2023, de acordo com [CISCO, 2020], é que haverá quase 30 bilhões de dispositivos/conexões, 45% de todos os dispositivos em rede terão conexão móvel, 55% estarão conectados por cabo ou Wi-Fi e a quinta geração (5G) da conexão móvel representará mais de 10,6% das conexões móveis do mundo, sua velocidade será 575 Mbps, 13 vezes superior à média da conexão móvel atual, indo de 13 Mbps (2018) para 44 Mbps (2023), e as velocidades médias de banda larga subirão de 46 Mbps para 110 Mbps.

Em [ERICSSON, 2018], haverá um aumento do tráfego de dados móvel em torno de 43% por ano, chegando a 107 *exabytes* (EB) por mês no fim de 2023. Isso é possível verificar na Figura 1, que mostra o crescimento do número de dispositivos e taxa de dados esperado em 2023.

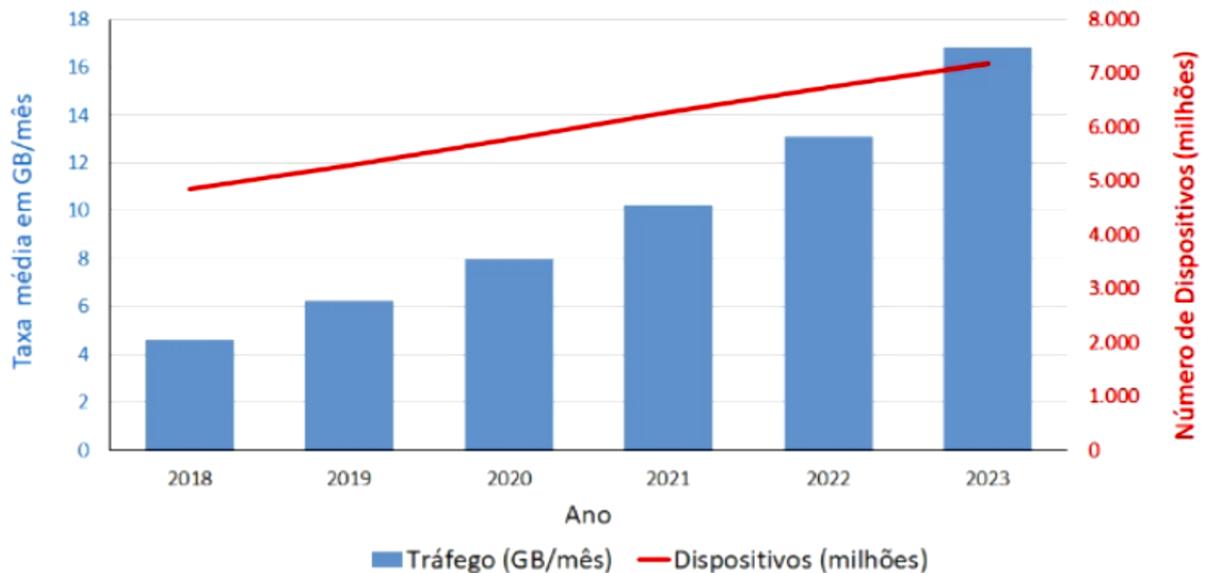


Figura 1 - Dispositivos x tráfego, [ERICSON, 2018]

Com capacidades avançadas de desempenho, o 5G deve propiciar infraestruturas móveis mais dinâmicas. Essa grande demanda de usuários e serviços na rede 5G, traz consigo uma inconstância na vazão da rede devido a movimentação dos usuários, surgindo um fenômeno conhecido como efeito maré [YAN et al., 2018]. Este fenômeno afeta as redes, que

necessitam ter seus recursos orquestrados para atender esta demanda tão dinâmica, caso esta tarefa seja negligenciada, gera-se um desbalanceamento na rede, e com isso os recursos da rede são mal alocados, quando os recursos estão subutilizados durante situações de tráfego desequilibrado, usuários podem ser bloqueados [CHEN et al., 2018], o que pode comprometer as métricas de avaliação de *Quality of Service* (QoS), que são métricas amplamente baseadas no desempenho da rede.

Alcançar a auto-otimização na rede de forma automatizada para extrair todo seu potencial, se tornou crucial para as *Mobile Network Operator* (MNOs), a ponto de serem direcionadas a investirem cada vez mais em infraestruturas de rede e busquem novos mecanismos para melhorar a taxa de dados, capacidade de cobertura e melhor qualidade e eficiência dos serviços, com um padrão suficientemente alto para atender às expectativas do usuário final [KHAN; ALHUMAIMA; AL-RAWESHIDY, 2015], sem a necessidade de aquisição de novos sites ou equipamentos.

O surgimento de tecnologias como *Software Defined Networks* (SDN) e *Centralized/Cloud Radio Access Network* (C-RAN) oferecem soluções tecnológicas promissoras para suprir a demanda que surge com o 5G. Essas tecnologias permitem que a implementação eficiente dos recursos de rede e o agendamento flexível sejam compartilhados.

A tecnologia SDN separa o plano de controle do plano de dados, permitindo que um controlador SDN adquira uma visão global atualizada, não apenas de toda a rede, mas também de todos os fluxos que competem pelo tráfego. Isso eleva a flexibilidade e a escalabilidade do sistema, tornando as redes programáveis, adaptáveis e econômicas [KITINDI et al., 2017]. Essas características permitem também, a integração de soluções inovadoras baseadas em técnicas de Inteligência Computacional (IC).

A C-RAN, tem uma visão global da rede alocando diferentes recursos conforme às variações de carga, possibilitando a orquestração dos recursos, tal como a alteração dos links lógicos entre *Radio Remote Head* (RRH) e *Baseband Unit* (BBU), balanceando assim os recursos de *base band* (BB) [BHAUMIK et al., 2012]. No entanto, esta tecnologia ainda está passando por vários desenvolvimentos para um serviço mais eficiente.

Essa arquitetura é constituída por três componentes principais, ou seja, redes de Unidade de Rádio Remota (RRH), um pool de BBU centralizado e rede de transporte ou *fronthaul* [PAN et al., 2018]:

- *Radio Remote Head* (RRH) - A rede RRH ajuda a conectar dispositivos sem fio a pontos de acesso ou torres semelhantes às redes celulares tradicionais. Pool de

- *Baseband Unit* (BBU) - O pool de BBU é instalado em um local central e funciona como um data center ou uma nuvem. Com base nos requisitos de rede atuais, seus vários nós BBU atribuem recursos dinamicamente a RRHs.
- *Links de Fronthaul* ou rede de transporte - *Fronthaul* é a camada que conecta um BBU com um conjunto de RRHs. Essas camadas fornecem links de alta largura de banda para gerenciar as necessidades de várias RRHs, faz a conexão RRHs-BBUs.

Destacam-se como sendo os principais benefícios de uma arquitetura C-RAN: (i) economia nas despesas operacionais devido à manutenção centralizada; (ii) melhora no desempenho da rede devido a técnicas avançadas de processamento de sinal coordenado; (iii) processamento das tarefas de forma mais ágil e eficiente; e (iv) redução nos gastos ao alocar diferentes recursos conforme às variações de carga [BHAUMIK et al., 2012].

De acordo com [GUPTA et al., 2015], a importância do C-RAN nas redes 5G, está nos diversos cenários de implantação destas redes, além de estar abrindo caminhos para enormes oportunidades no mercado global, propondo o compartilhamento de recursos enquanto aborda questões de escalabilidade de processamento.

1.1 Motivações

Devido ao aumento da utilização de comunicações móveis nos últimos anos, a arquitetura C-RAN, tradicionalmente utilizada para processar requisições de usuários, passou a necessitar de maior capacidade para suportar altas taxas de dados provenientes de ambientes de alta mobilidade [CMRI, 2010; YASEEN; AL-KHALIDI; AL-RAWESHIDY, 2017]. Com a chegada do 5G, esse desafio se tornou ainda maior, dado o aumento significativo do volume de dados móveis impulsionado, principalmente, pelos serviços e aplicações multimídias e pelo crescente número de dispositivos que se conectam a rede [AL-FALAHY et al., 2017].

Na arquitetura C-RAN, as BSs precisam ter uma pré-configuração com alta capacidade para lidar com os picos de uso sem interrupções. Entretanto quando se trata de cenários práticos de redes móveis, o tráfego raramente está em seu pico, pois a carga muda gradualmente ao longo do dia, sendo possível observar as alterações em um padrão geométrico de tempo, denominado *Tidal Effect* [YAN et al., 2018; RAZA et al., 2016]. A Figura 2 retrata a flutuação da carga na BS, demonstrando que, durante a noite, as BSs nas áreas residenciais são muito utilizadas, enquanto as BSs nas áreas empresariais permanecem ociosas e consumindo muita energia [CHECKO et al., 2014].

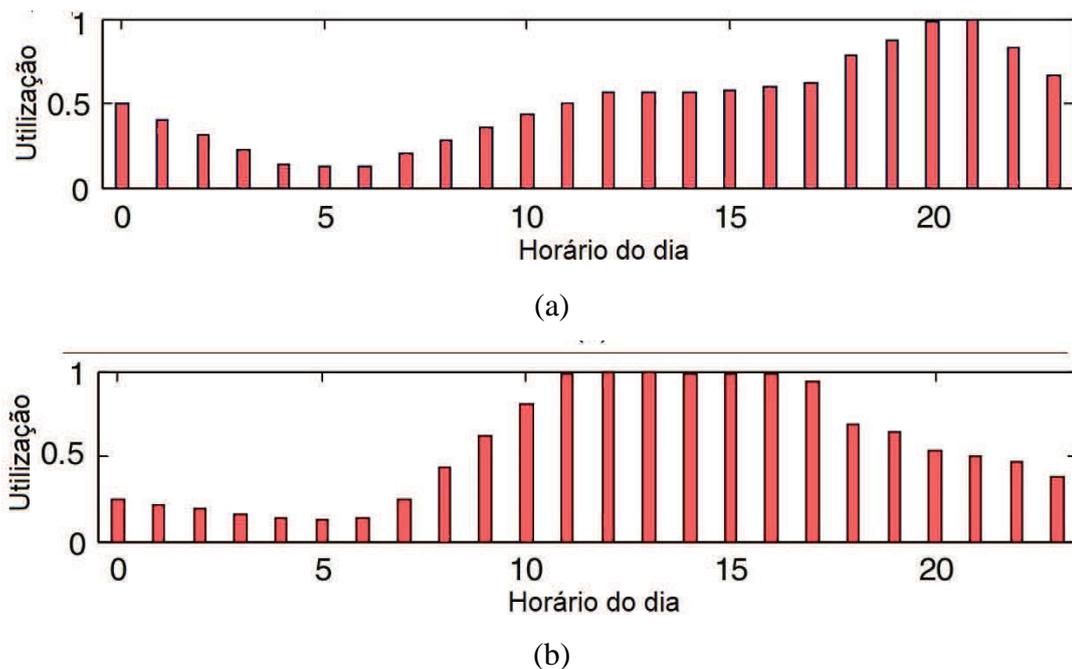


Figura 2 - Flutuação da carga de tráfego na BS durante o período de um dia: (a) área residencial; (b) área empresarial., adaptado de [RAZA et al. 2016].

Se tornou categórico resolver esse problema, de modo a liberar os recursos e diminuir custos de processamento e energia, especialmente porque, de acordo com informações divulgadas pela [CISCO, 2020], conforme visto anteriormente, a transformação digital global continuará a ter um impacto significativo nas demandas e requisitos das redes, sobretudo em razão do aumento de usuários da internet. já que usuários pessoais e corporativos no mundo inteiro continuam e continuarão gerando novas demandas e expectativas para a conexão em redes móveis.

Esse aumento contínuo de usuários de internet, dispositivos e conexões, com demandas cada vez maiores na rede, mostra a necessidade de uma arquitetura de rede móvel de maior capacidade para suportar altas taxas de dados provenientes de ambientes de alta mobilidade [YASEEN; AL-KHALIDI; AL-RAWESHIDY, 2017].

A arquitetura C-RAN se torna uma solução eficiente no que diz respeito às questões de gerenciamento dos recursos computacionais. Nessa arquitetura, ocorre um provisionamento dinâmico de acordo com as necessidades correntes de tráfego [SIGWELE et al., 2015], reduzindo os custos, em razão dos recursos computacionais serem alocados de forma mais eficiente, aumentando a economia de energia e compartilhando o processamento dos sinais nos BBU's para diferentes *Base Stations* (BSs) [NIKAEIN et al., 2015]. O processamento do sinal digital da RRH é executado em uma *BBU Pool*, localizada na nuvem, a qual encontra-se em

uma distância máxima recomendada de aproximadamente até 40Km da RRH [CHANCLOU et al., 2013].

Levar a efeito a arquitetura C-RAN não é, no entanto, tarefa trivial, pois há diversos desafios a serem superados. A necessidade de lidar de forma mais eficiente com as diferentes demandas de carga, utilizando-se de abordagens que realizem balanceamento de carga e a escalabilidade de recursos tem sido bastante explorados, abrindo espaço para oportunidades de pesquisa na literatura. A exemplo, podemos citar a orquestração dos recursos ou o provisionamento dinâmico, de acordo com as necessidades correntes de tráfego [SIGWELE; PILLAI; HU. 2015], a fim de aumentar o compartilhamento do processamento dos sinais nas BBU's para diferentes BSs [NIKAEIN et al., 2015; RAZA et al., 2016]. Porém, conseguir lidar com a reestruturação para melhor distribuir as requisições a serem executadas por cada um dos recursos compartilhados, de maneira que ocorra um balanceamento de carga entre as BBU's para um processamento ágil e com baixos custos de infraestrutura, é um procedimento complexo e, ao mesmo tempo, vital para manter o equilíbrio e a qualidade na entrega das requisições [MAROTTA et al., 2015; DUAN et al., 2017].

O balanceamento de carga e a escalabilidade de recursos, além de ser uma opção de baixo custo para operadoras em relação as despesas operacionais ou *Operational Expenditure* (OPEX) e despesas de capital ou *Capital Expenditure* (CAPEX), por se tratarem de fatores significativos para a redução de custos e o aumento do QoS, resolvem o problema de desbalanceamento e reduz os custos e a economia de energia, em razão dos recursos computacionais serem alocados de forma mais eficiente.

1.2 Objetivos

Baseado nos problemas ocasionados pela dinamicidade do tráfego citados anteriormente, esta dissertação tem como objetivo geral a otimização do balanceamento da carga nas BBU's para arquiteturas C-RAN/5G, através do mapeamento BBU-RRH, com a finalidade de alocar recursos capaz de fornecer balanceamento de carga de rede, a fim de melhorar a qualidade e eficiência dos serviços, com um padrão suficientemente alto para atender às expectativas do usuário final.

Assim, os objetivos específicos desta dissertação incluem:

- Estudar os elementos essenciais de uma arquitetura C-RAN para garantir similaridade a cenários realistas, especialmente em cenários de alta densidade e mobilidade de usuários;

- Apresentar as principais pesquisas que envolvam o mapeamento BBU-RRH, visando o balanceamento de carga nas BBUs, para arquiteturas C-RAN/5G;
- Apresentar uma Função objetivo (FO), como mecanismo para equilibrar a carga da rede nos recursos de processamento;
- Propor o uso do DPSO como um algoritmo capaz de garantir o balanceamento de carga, ao mesmo tempo que reduz os custos das operadoras em relação ao OPEX e CAPEX, também que satisfaça as restrições de QoS da rede;
- Avaliar o desempenho do modelo proposto.

1.1 Organização do texto

Este capítulo apresentou uma breve descrição da solução que está sendo proposta nesta dissertação. Nos capítulos seguintes, a metodologia de desenvolvimento será melhor apresentada bem como os resultados que demonstram os benefícios obtidos. Os capítulos 2 e 3 descrevem os fundamentos técnicos e teóricos que norteiam o desenvolvimento da proposta dessa dissertação. No capítulo 2, são apresentados os conceitos e a evolução das redes móveis, bem como as principais tecnologias que envolvem uma rede centralizada, e no capítulo 3 serão tratados os fundamentos sobre meta-heurística. O capítulo 4 aborda os trabalhos relacionados à pesquisa, encontrados na literatura que serviram de base para o desenvolvimento da proposta. O capítulo 5 descreve em detalhes a solução proposta. O capítulo 6 descreve a métodos de desenvolvimento da solução proposta e destaca os resultados obtidos. Por fim, o capítulo 7 apresenta as considerações finais e o capítulo 8 mostra os trabalhos prospectivos e novas oportunidades de pesquisa trabalhos futuros.

2 REDES MÓVEIS

Neste capítulo, serão apresentados os conceitos fundamentais sobre redes móveis. Será feita uma breve introdução sobre a evolução da comunicação móvel, apresentando a arquitetura da geração atual de redes móveis. Uma descrição da arquitetura de C-RAN, seus tipos e benefícios, será apresentada; bem como as principais características da rede definida por software e virtualização de funções de Rede.

2.1 Redes móveis de quinta geração (5G)

O desenvolvimento da comunicação móvel segue em ascensão, entretanto, nos últimos 40 anos, testemunhou o avanço de quatro importantes gerações das comunicações sem fio. Os avanços iniciaram com a telefonia móvel na década 80 e foram evoluindo seus serviços oferecidos juntamente com a tecnologia de comunicação [DAHLMAN et al., 2018]. A cada geração, alguns padrões, capacidades técnicas e novos recursos, que a diferenciam de gerações anteriores, foram especificados conforme ilustrado pela figura 3.

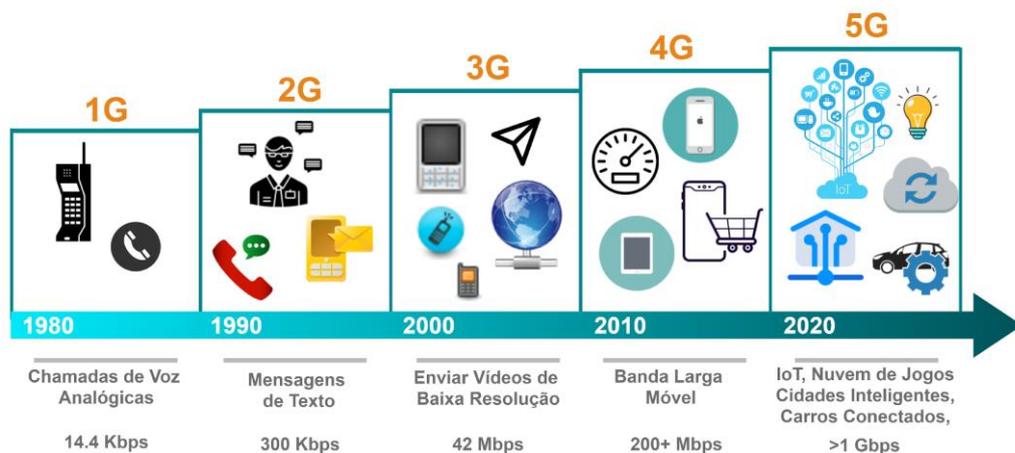


Figura 3 - Evolução da Comunicação Móvel, [DAHLMAN et al, 2018]

A primeira geração (1G) foi criada no começo dos anos 80 e surgiu com o objetivo simples de possibilitar ligações de voz em um aparelho sem fio, possibilitando a execução de chamadas em movimento. Contudo, pela baixa capacidade de tráfego e pelo alto custo, não se sabia ao certo qual era o potencial de crescimento da nova tecnologia, pois era limitada apenas a transmissão de voz e era pouco acessível ao grande público. Por esse motivo, não houve uma padronização dos sinais, ou seja, a Europa tinha diversos padrões diferentes, por exemplo, o *Total Access Communications System* (TACS), para o Reino Unido, a Áustria, a Espanha, a

Irlanda e a Itália; o *Nordic Mobile Telephone 450* (NMT450), para a Suécia, a Noruega, a Finlândia e a Dinamarca; e a Radiocom2000, para a França [BABA et al., 2018].

Já os Estados Unidos (EUA) adotaram o padrão *Advanced Mobile Phone System* (AMPS), implementado em 1983. Operava na faixa de 800 MHz. Basicamente, as bandas eram divididas entre canais de 30 kHz, o que foi possível graças ao *Frequency Division Multiple Access* (FDMA). Os sinais recebidos de um transmissor cobriam uma área específica, chamada de célula. Quando um usuário saía dessa célula e passava a integrar outra, o sinal era transferido sem transição ou interferência. Esse foi o grande divisor de águas para a tecnologia G e o que possibilitou o avanço da mobilidade na telefonia. Esse padrão também foi seguido por diversos países da América, entre eles o Brasil.

Enquanto os Estados Unidos precisavam de uma melhor performance, a Europa visava uniformizar os sistemas. Foi assim que, no começo dos anos 90, a segunda geração (2G) de telefone móvel passou a utilizar tecnologia digital na transmissão de rádio, deixando a analógica de lado e permitindo integração com circuitos digitais, com isso além da transmissão de voz permitiu também a transmissão limitada de dados [BABA et al., 2018].

Os EUA, então, criaram os padrões IS-54, IS-136 e IS-95, para aumentar a capacidade de tráfego. Já a Europa unificou os padrões criando o *Global System for Mobile* (GSM — sistema global para celular) e esse foi dividido em 3 fases:

- Primeira fase: ofereciam-se chamadas de voz, SMS, dados síncronos e assíncronos, bem como transmissão de pacotes assíncronos;
- Segunda fase: passou-se a oferecer serviços de e-mail, ampliação no número de usuários [e diminuição da qualidade da voz], melhora no SMS e serviços de dados para informações gerais (clima e esportes, por exemplo);
- Terceira fase: iniciou-se a implementação dos pacotes de dados com altas taxas de transmissão, padrão GPRS.

No começo dos anos 2000, a terceira geração (3G) introduziu a internet de alta qualidade ao serviço móvel, além da possibilidade de acesso à internet sem fio. O sistema móvel universal de telecomunicações (UMTS) foi a primeira tecnologia que surgiu na 3G, que começou a ser utilizada no Brasil em 2007. A sua implementação foi baseada no acesso múltiplo por divisão de código em banda larga (W-CDMA), um padrão de radiofrequência que possibilitou conexões de até 2 Mbps para *download* e *upload*.

A quarta geração (4G), utiliza a tecnologia LTE ou *Long Term Evolution* (evolução a longo prazo) e trouxe uma experiência melhorada do acesso a internet móvel pelo usuário, priorizando o tráfego de dados em vez de tráfego de voz, o que a torna mais rápida e estável.

Foi a primeira versão a ser comercializada com taxas de pico que atingiam os 300 Mbps para download e 75 Mbps para *uplink*, além disso, a latência também é muito mais baixa em comparação com as gerações anteriores [DAHLMAN et al., 2018].

Com tamanho tráfego de dados e com a implementação cada vez maior de dispositivos inteligentes que utilizam a internet para as mais variadas operações e diferente das mudanças realizadas nas gerações anteriores, a rede móvel de quinta geração (5G) tem o propósito de ser uma arquitetura de rede mais heterogênea, com características superiores às gerações anteriores. De forma bastante simplificada, uma rede 5G pode ser entendida como um conjunto de 3 partes ilustradas na Figura 4, o núcleo (*core*), a RAN e os *User Equipment* (UEs), as duas primeiras partes são virtualmente invisíveis para o usuário.

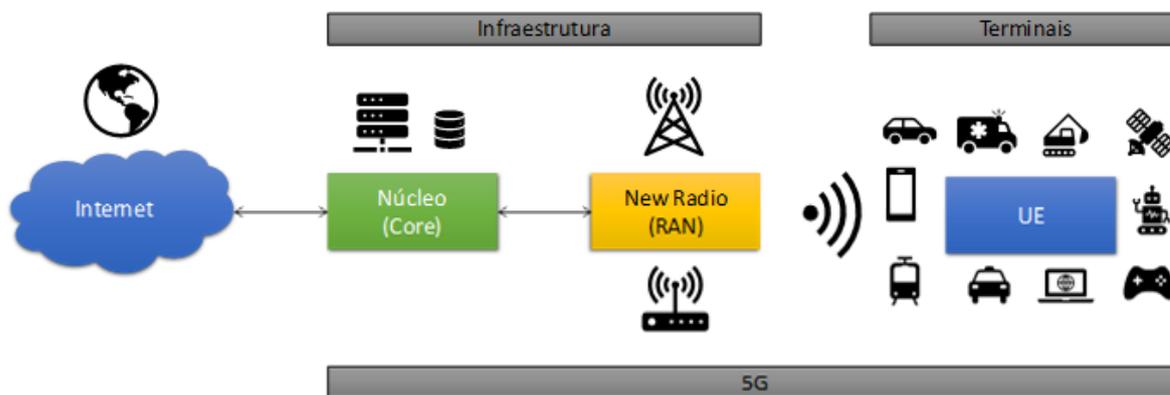


Figura 4: Arquitetura Simplificada de uma Rede 5G, adaptado de [CHINA, 2011]

A rede 5G funciona a partir da adaptação das antenas já utilizadas, ocupando as frequências entre 600 MHz e 700 MHz, 26 GHz e 28 GHz, 38 GHz e 42 GHz, tendo os smartphones representando mais de 90% de todo tráfego móvel no mundo [DAHLMAN et al., 2018]. É esperado que o 5G entregue uma conexão robusta, com:

- a) Taxa de dados até 10 vezes superior ao 4G atual;
- b) Latência menor;
- c) Maior largura de banda;
- d) Oferecer ao usuário tempo praticamente ininterrupto de conexão;
- e) Número maior de dispositivos conectados ao mesmo tempo;
- f) Menor custo de operação.

O 5G não busca ser uma melhora de seus antecessores, mas busca ser sim um salto tecnológico em relação à taxa de dados, latência, qualidade de conexão, permitindo uma evolução nos serviços suportados [CISCO, 2017]. Nesse cenário, a comunicação vai além da interpessoal, inclui também a *Machine-to-machine* (M2M), o que gera um impacto fundamental na arquitetura da rede.

2.2 Rede de Acesso Centralizado (C-RAN)

O C-RAN é um paradigma promissor para redução de gastos em bens de capitais e operacionais, bem como eficiência energética e espectral e aproveitamento dos recursos computacionais [WU et al., 2015]. A arquitetura *Radio Access Network* (RAN) tradicional, consiste em muitas BSs que se conectam apenas a um número fixo de antenas de setor, onde as funções de processamento de frequência de rádio (RF) e BB são integradas dentro de uma BS e o módulo de antena geralmente está integrado ao módulo de rádio [CHECKO et al., 2015].

A arquitetura C-RAN foi inicialmente proposta pela *China Mobile* (maior operadora celular chinesa) em 2009, com o intuito de habilitar as futuras demandas por aplicações com altas taxas de dados. O “C” em C-RAN denota tanto “*centralized*” (centralizado) como “*cloud*” (aplicação em nuvem), que são dois aspectos diferentes, mas relacionados, porque a C-RAN coloca todos os BBUs em um local centralizado, normalmente imaginado como a nuvem com a funcionalidade de BBU virtualizada em plataformas de processamento comuns.

Contudo, como os usuários são móveis, o tráfego de cada BS tem sua taxa de utilização média bastante baixa, como consequência disto, durante o dia, as BSs nas áreas de centros urbanos ficam sobrecarregadas, enquanto as das áreas residenciais ficam ociosas, porém continuam consumindo uma quantidade significativa de energia e processamento [CHINA, 2011].

Segundo [CHECKO et al., 2015], a maneira mais proeminente de otimizar a capacidade da arquitetura RAN tradicional é adicionando mais células, criando uma estrutura complexa de redes heterogêneas e de *SmallCells* ou implementando técnicas como MIMO (*Multiple Input, Multiple Output*) e *Massive MIMO*, onde inúmeras antenas atendem simultaneamente a vários usuários na mesma frequência. No entanto, essa abordagem resulta em níveis crescentes de interferência entre células e altos custos financeiros, o que torna a arquitetura RAN tradicional muito cara para as MNOs se manterem competitivas no mercado.

O custo financeiro de uma rede móvel, também conhecido como *Total Cost of Ownership* (TCO), é a soma das despesas de CAPEX e OPEX. O CAPEX refere-se principalmente aos gastos relevantes para a construção da rede, que podem abranger desde o

planejamento da rede até a aquisição do local, como por exemplo, o hardware de RF e BB, licenças de software e custos de instalação. Já o OPEX cobre os custos necessários para operar a rede, ou seja, aluguel do local, energia elétrica e custos com manutenção e operação [CHECKO et al., 2015].

Para sanar esse desafio, a *China Mobile Research Institute* (CMRI) introduziu, em abril de 2010, uma nova tecnologia centralizada de RAN totalmente baseada em computação em nuvem, denominada de C-RAN. O objetivo era de fornecer, desde um nível sofisticado de cooperação e comunicação entre BSs, como também o compartilhamento dinâmico dos recursos de computação e espectro [POMPILI et al., 2015; CHECKO et al., 2015]. Tais como, processamento centralizado, rádio colaborativo e infraestrutura com eficiência energética. Essa arquitetura combinou todos os recursos computacionais das BSs tradicionais com a *Network Function Virtualization* (NFV) em um conjunto centralizado de BBUs denominado de BBU pool, o que permitiu a comunicação entre as BSs com baixa latência e a troca de dados em alta velocidade.

O BBU pool opera executando o processamento de BB com a ajuda de processadores de uso geral, na qual todo o processamento ocorre por um único e poderoso processador centralizado, isso significa que os recursos de rádio de diferentes BBUs podem ser compartilhados [HOSSAIN et al. 2019]. Esses sinais são coletados dos RRHs e transmitidos para a plataforma em nuvem por meio de uma rede de transmissão óptica denominada *fronthaul*. Isso possibilita o processamento centralizado de alto nível e de baixo consumo energético [WU et al. 2015]. A Figura 5 apresenta a arquitetura C-RAN tradicional, em seguida, resume-se às principais funções de cada componente da arquitetura.

Os RRHs são responsáveis por transmitir sinais de RF aos *User Equipments* (UEs) no *downlink* e encaminhar os sinais de BB dos UEs ao BBU pool no processamento de *uplink* através de uma *Optical Transport Networks* (OTN). Por centralizar a maioria das funções de processamento de sinal no BBU pool, os RRHs podem ser relativamente simples e podem ser distribuídos geograficamente de maneira econômica [WANG et al., 2014].

O BBU pool é um *datacenter* responsável por grande parte do processamento dos sinais das camadas L1/L2 em um grande número de BBUs virtualizadas que operam em plataformas NFV, geralmente são instanciadas para processar sinais de BB oriundos dos RRHs [SUNDARESAN et al., 2016].

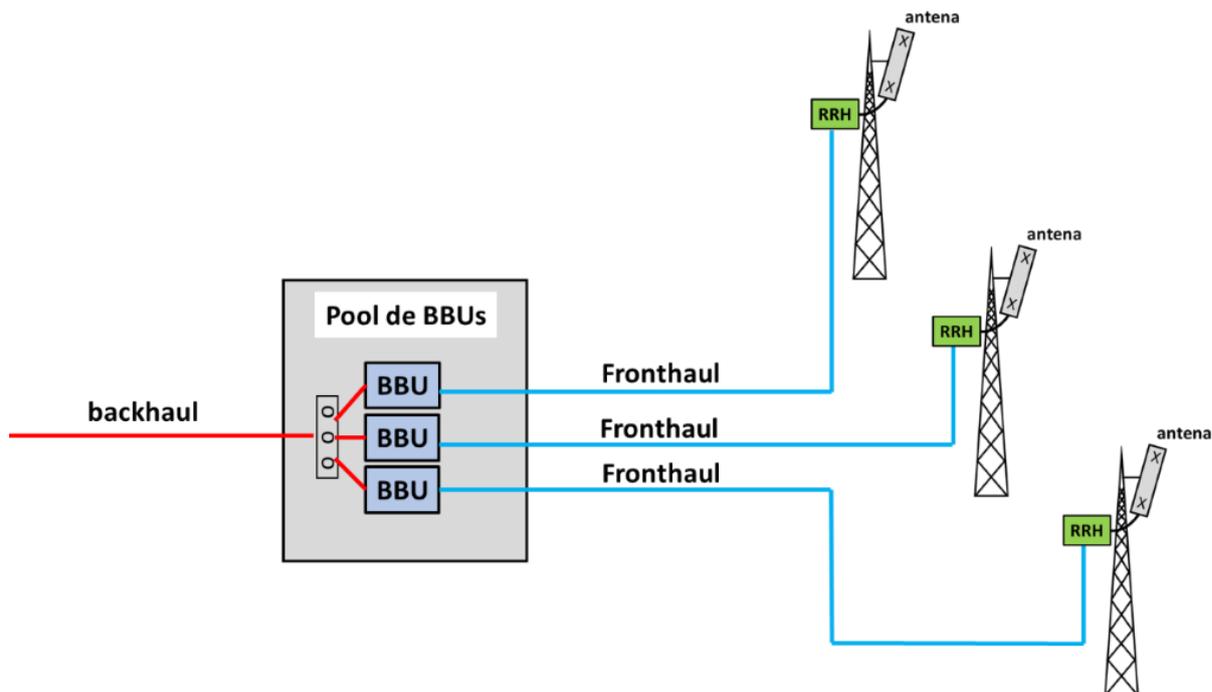


Figura 5 – Arquitetura C-RAN tradicional, adaptado de [CHINA, 2011].

De acordo com [CHINA, 2014; WANG et al., 2014], o mapeamento BBU-RRH, ou seja, a relação proporcional de BBUs por RRHs, pode ser implementada de maneira centralizada ou distribuída. Na forma distribuída, um RRH se conecta diretamente à sua BBU exclusiva. Essa abordagem é simples e fácil de ser implementada. Porém, não é possível explorar os benefícios e vantagens do processamento conjunto dos sinais e do controle centralizado da C-RAN, já em uma abordagem centralizada, todos os RRHs podem se conectar a um comutador central.

Este comutador tem a tarefa de agendar, com flexibilidade, os recursos de processamento no BBU pool para um RRH ou um conjunto de RRHs. Essa prática apresenta muitas vantagens em termos de compartilhamento flexível de recursos e eficiência energética por meio de programação paralela, o que permite a implementação de algoritmos eficientes de prevenção e gerenciamento de interferências em redes heterogêneas (*HetNets*). Além disso, fornece os meios para ativar e desativar dinamicamente os RRHs, de acordo com as flutuações do tráfego em diferentes cenários.

O *Fronthaul* é responsável por interligar os RRHs ao BBU pool e é considerado um dos grandes desafios da arquitetura C-RAN. Segundo [WANG et al. 2014], os links de *fronthaul* se dividem em duas categorias: ideal, sem restrições de largura de banda e não ideal, com algum tipo de restrição. Geralmente podem ser idealizados usando diferentes tecnologias que incluem a comunicação por fibra óptica, comunicação celular ou comunicação por ondas milimétricas.

Para [SALMAN, 2016], a comunicação por fibra óptica é considerada ideal no C-RAN, pois fornece o mais alto requisito de largura de banda, no entanto, com altos custos de implementação. Por outro lado, a comunicação por ondas milimétricas é mais acessível e de fácil implantação, porém com algumas considerações relevantes como uma largura de banda menor, dificuldade em se propagar através de obstáculos e enlaces muito suscetíveis a bloqueios [WEI et al, 2014].

1.2.1 Tipos de Arquiteturas C-RAN

Com base no desempenho do *fronthaul* algumas funcionalidades do grupo BBU-RRH podem ser divididas, o que causará um grande impacto nos detalhes da arquitetura C-RAN. Desta forma, uma C-RAN pode ser classificada como, totalmente centralizada e parcialmente centralizada. A divisão depende de onde são tratadas as funções das camadas L1 (PHY), L2 (*Media Access Control* (MAC)) e L3 (*Network*). As funções da camada L1, incluem o processamento de BB, enquanto a L2 é responsável pelo mecanismo de controle de acesso ao meio. Já a camada L3, assegura os processos de roteamento e endereçamento dos pacotes [POMPILI, 2015; CHECKO et al., 2015]. Abaixo, é apresentado um resumo das principais funcionalidades do BBU pool em cada tipo de arquitetura.

Totalmente centralizado: Conforme observado na Figura 6, em uma arquitetura totalmente centralizada, todas as funcionalidades de L1, L2 e L3 são implementadas pela BBU. Nesse tipo de arquitetura, as BBUs lidam com todas as funções de gerenciamento e processamento dos recursos e, portanto, podem se beneficiar significativamente de operações e manutenções simplificadas. Além disso, é possível implementar o compartilhamento máximo de recursos, por exemplo, a execução de técnicas de controle de interferência intercelular. Para [SALMAN, 2016], esta opção é a considerada a configuração ideal de C-RAN, entretanto, devido à sobrecarga de comunicação entre a BBU e as RRHs, seu desempenho é limitado ao desempenho do *fronthaul*.

Parcialmente centralizado: Em estruturas parcialmente centralizadas, as funções da camada L1 são realizadas nos RRHs, enquanto as funções de L2 e L3 são realizadas nas BBUs. Essa configuração tem a vantagem de exigir uma largura de banda muito menor entre BBU e RRH, separando o processamento de BB da BBU e integrando-o no RRH [CHECKO et al., 2015; SALMAN, 2016]. Contudo, como o processamento de BB é integrado ao RRH, ele tem menos flexibilidade na atualização e menos conveniência para o processamento de sinal colaborativo, por exemplo, o compartilhamento de recursos da camada L1 entre diferentes

RRHs [CHINA, 2011; POMPILI, 2015]. A Figura 7, apresenta um exemplo desse tipo de arquitetura.

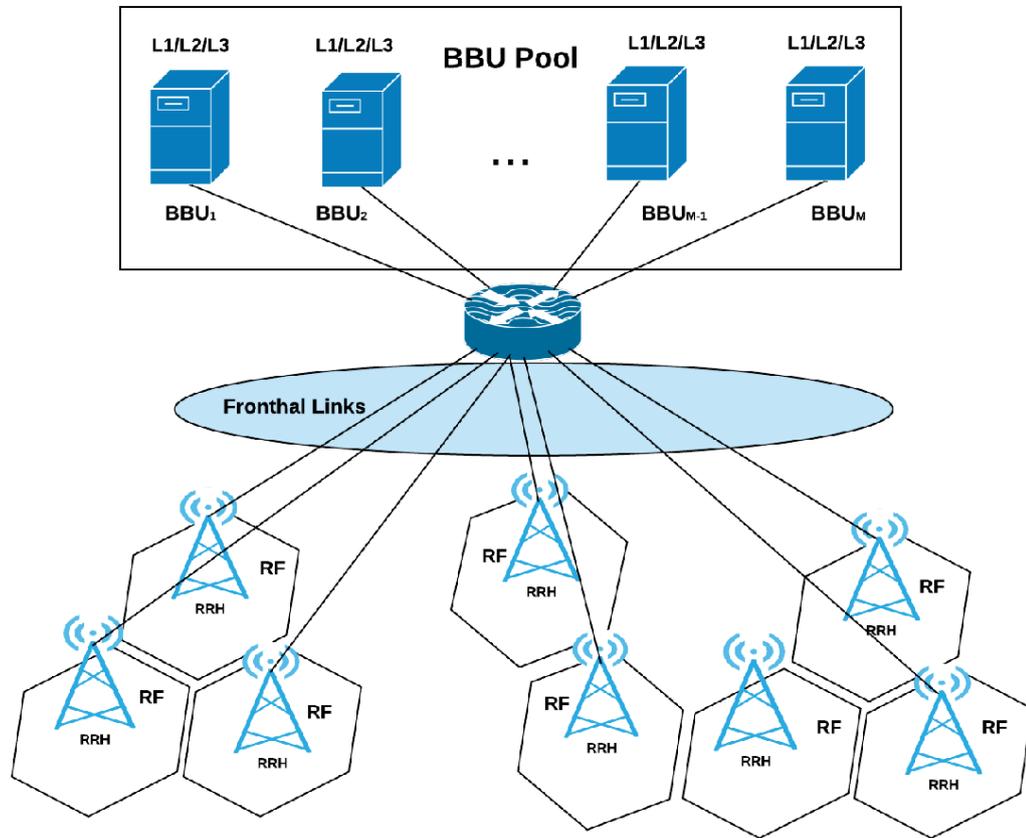


Figura 6 – Arquitetura C-RAN Totalmente Centralizada, [CHINA, 2011]

Ainda de acordo com [CHINA, 2011], independentemente do tipo de arquitetura C-RAN implementada, as MNOs podem fazer atualizações rapidamente em sua rede apenas adicionando e conectando novos RRHs ao BBU pool. Se a carga da rede aumentar, as MNOs precisarão apenas atualizar o Hardware do BBU pool para acomodar o aumento da capacidade de processamento. Além disso, combinando a arquitetura totalmente centralizada com uma plataforma de virtualização de processadores de uso geral, é possível desenvolver técnicas baseadas em SDN que permitirão a atualização dos padrões de interfaces de rádio apenas por software, facilitando a alocação de recursos físicos e o consumo energético da rede.

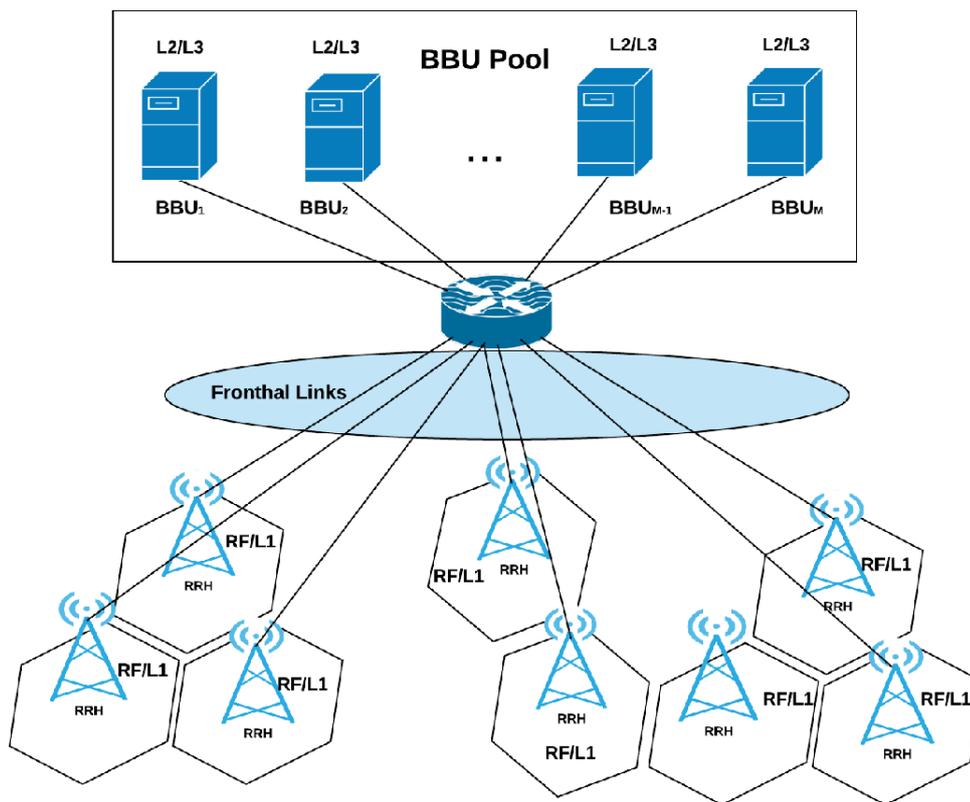


Figura 7 – Arquitetura C-RAN Parcialmente Centralizada, [CHINA, 2011]

1.2.2 Benefícios de Arquiteturas C-RAN

Além das vantagens citadas acima, a arquitetura C-RAN traz outros benefícios, Como:

- **Redução de custos:** Na RAN tradicional a implantação e o comissionamento de BSs são caros e demorados [HOSSAIN et al., 2019]. Por outro lado, na C-RAN, todas as BBUs e equipamentos de suporte local de uma grande região estão localizados em um datacenter comum, o que simplifica a gestão, operação, manutenção e o compartilhamento dos equipamentos [POMPILI, 2015]. Segundo [HOSSAIN et al., 2019], essa abordagem proporciona uma redução de CAPEX de até 15%. Além disso, os recursos computacionais no C-RAN são virtualizados em poucos datacenters de alto poder de processamento, deixando funções mais simples nos RRHs, isso permite a implantação de técnicas de balanceamento de carga e que podem proporcionar redução de grande parte dos custos de OPEX.
- **Eficiência energética:** A C-RAN reduz o número de BSs distribuídas e as libera do compromisso de fornecer serviços 24h, 7 dias por semana. Portanto, o consumo de energia de aparelhos de ar-condicionado e de outros equipamentos de suporte podem ser drasticamente reduzidos [HOSSAIN et al., 2019]. Além disso, como as técnicas

de redução de interferência cooperativa podem ser aplicadas entre os RRHs, é permitida uma maior densidade de RRHs. Isto é, células menores, com menor potência de transmissão, podem ser implantadas visando maior reutilização e capacidade de frequência, enquanto a cobertura da rede não é afetada [CHECKO et al., 2015]. O consumo de energia e o congestionamento de carga na rede, podem ser minimizados ao alocar dinamicamente os recursos de processamento de muitas BBUs e nestes casos, algumas podem até ser desativadas em períodos de ociosidade [KHAN et al., 2015].

- **Gerenciamento aprimorado de interferência:** Em C-RAN, as BBUs podem facilmente compartilhar sinalização, dados de tráfego e informações referentes ao estado do canal de usuários ativos. O compartilhamento de informações sobre faixas do espectro utilizadas entre o UE e RRHs permite multiplexar mais fluxos no mesmo canal com pouca ou até nenhuma interferência [CHECKO et al., 2015].
- **Adaptabilidade ao tráfego não uniforme:** As redes celulares modernas exibem uma quantidade substancial de diversidade tempo/espaço. Por exemplo, durante o dia, os usuários se movem entre diferentes áreas (residenciais e comerciais) [HOSSAIN et al., 2019]. Isso significa que quando os usuários passam das áreas comerciais para as áreas residenciais, uma enorme quantidade de energia de processamento é desperdiçada nas áreas das quais os usuários foram transferidos, implicando um desperdício de energia de processamento fora do horário de pico [KHAN et al., 2015]. Como todo o processamento é centralizado no BBU pool, a CRAN é extremamente adequada ao tráfego distribuído não uniforme, pois embora o RRH em serviço mude dinamicamente de acordo com o movimento dos UEs, a BBU em serviço ainda está no mesmo BBU pool [CHINA, 2011]. Esta abordagem permite que o tráfego distribuído gerado seja compartilhado por mais de uma BBU.
- **Facilidade no gerenciamento:** A interconexão de baixa latência, alta velocidade e a combinação dos paradigmas de NFV e SDN, facilitam a criação de BSs virtuais. Portanto, para que as operadoras adicionem ou atualizem regras ou SLAs, não é necessário a substituição de nenhum equipamento, mas simplesmente atribuir uma nova BBU [KITINDI et al., 2017; CHECKO et al., 2015].

1.3 *Software defined networks (SDN) e Network functions virtualization (NFV)*

A *Software defined networks* (SDN) e a *Network functions virtualization* (NFV), visam atender requisitos como flexibilidade na configuração e implementação de funções de rede, e, conseqüentemente, facilitar procedimentos de operação e manutenção. Ambas as arquiteturas fazem uso de abstração de rede, entretanto, não da mesma maneira. Enquanto, a SDN procura separar as funções de controle de rede das funções de encaminhamento de rede, a NFV procura abstrair o encaminhamento e outras funções de rede do hardware no qual é executado [SDN vs. NFV, 2022].

1.3.1 *Software defined networks (SDN)*

O conceito SDN não é recente e teve origem ainda na década de 1990, foi projetada para lidar com ambientes de rede dinâmicos e exigentes, de modo a torná-la flexível e ágil [CHEN et al., 2015]. Tecnologias SDN são promovidas e padronizadas pela *Open Network Foundation* (ONF) e seu sucesso vem da abstração sistemática de problemas complexos de rede, transformando os problemas anteriores das redes distribuídas em um problema lógico centralizado. A Figura 8 compara as redes convencionais, figura 8(a)), e SDN, figura 8(b).

O paradigma SDN visa desmembrar o plano de controle e o plano de dados. Essa separação não só fornece aos operadores de rede o uso eficiente dos recursos como também simplifica o gerenciamento. Além disso, SDN emergiu como tecnologia capaz de proporcionar uma mudança em direção a um modelo centralizado de fluxo, que utiliza hardware programável de baixo custo, um controlador logicamente centralizado, e aplicações que usam as informações expostas pelo controlador para orquestrar os serviços da rede [KARAKUS; DURRESI, 2017b]. Para [CHAVES; GARCIA E MADEIRA, 2016], SDN pode ser visto como um facilitador para a implantação de NFV, e juntos, podem simplificar as redes 5G.

É destacado em [NUNES, 2014] que, “a separação do hardware de encaminhamento da lógica de controle permite uma implantação mais fácil de novos protocolos e aplicações, da visualização e gerenciamento da rede. Ao invés de impor políticas de funcionamento em diversos dispositivos, a rede é reduzida a hardware de encaminhamento simples e ao controlador de rede, responsável pela tomada de decisão”.

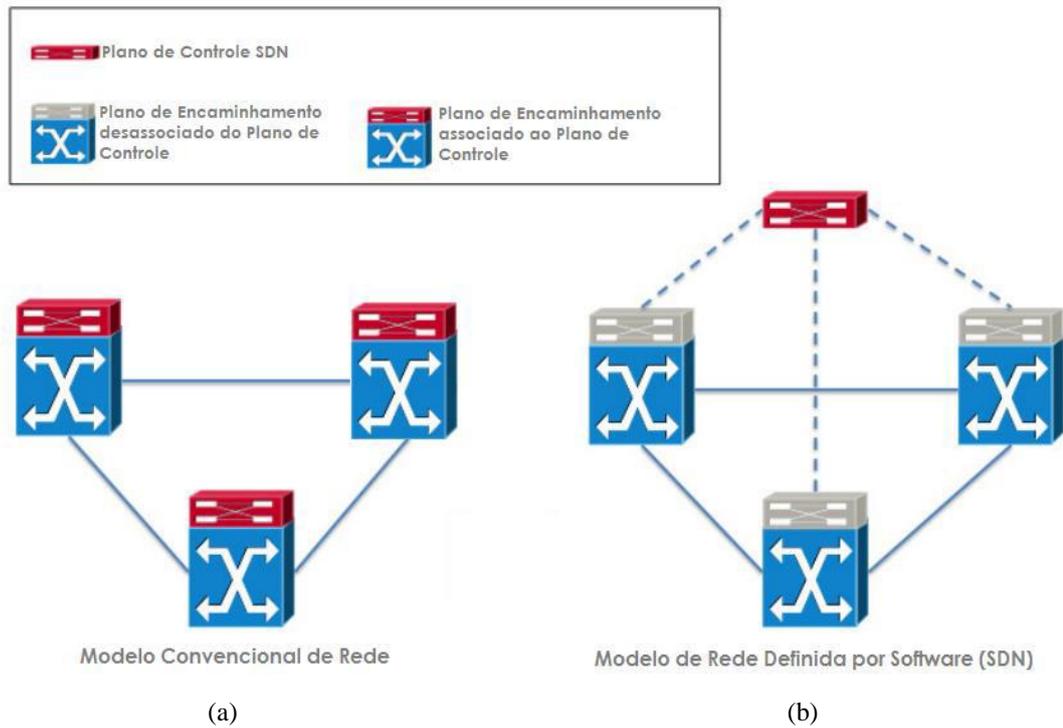


Figura 8 - Modelos de Rede Convencional (a) e de Rede Definida por Software (b), [NUNES at al, 2014]

Em uma rede tradicional, os fluxos (ou pacotes) são tratados principalmente com base em uma ou poucas combinações de atributos de cabeçalhos de pacotes, como prefixos *Internet Protocol* (IP) de destino mais longos, endereços MAC de destino ou uma combinação de endereços IP e números de porta *Transmission Control Protocol* (TCP) e *User Datagram Protocol* (UDP). Com a SDN os fluxos são gerenciados com base em mais atributos de cabeçalhos de pacotes por meio de uma Interface de plano de dados do controlador. A Figura 9 apresenta a divisão vertical dos três planos da arquitetura SDN. Abaixo apresenta-se as funções principais de cada um. A separação dos planos facilita o gerenciamento da rede, uma vez que torna o plano de controle diretamente programável. Dessa forma, caso haja necessidade de redefinir as regras de encaminhamento, bastará modificar o plano de controle [KARAKUS; DURRESI, 2017b].

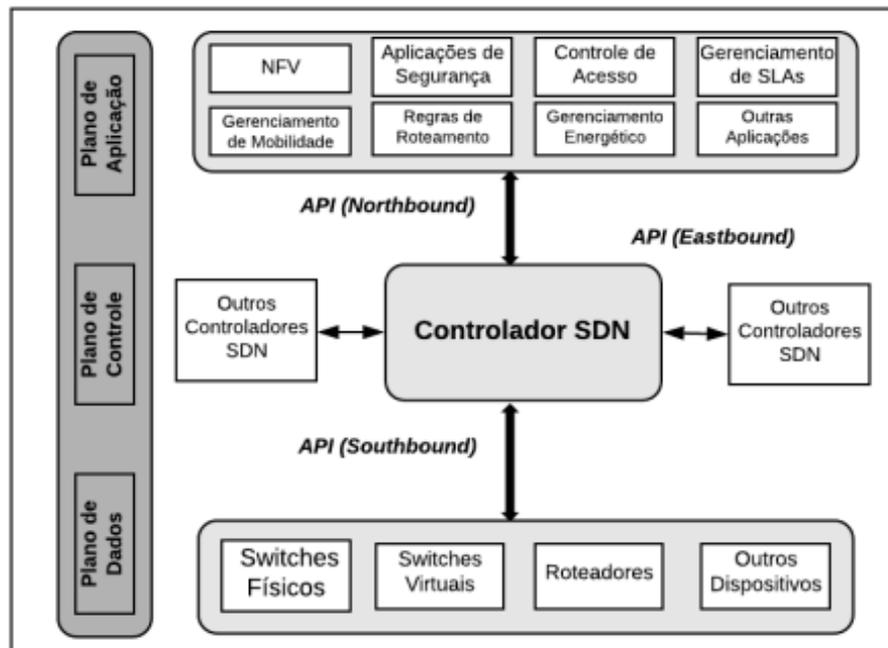


Figura 9. Arquitetura SDN, adaptado de [NGUYEN; DO; KIM, 2016]

- **Plano de dados:** É o plano inferior responsável por todas as atividades de dados, incluindo encaminhamento, fragmentação e remontagem. Consiste em dispositivos de rede, como roteadores, comutadores virtuais e pontos de acesso acessíveis e gerenciados através de controladores SDN. Os elementos de rede e os controladores podem se comunicar através de conexões seguras em portas lógicas, como a conexão *Transport Layer Security* (TLS) [KARAKUS e DURRESI, 2017a].
- **Plano de controle:** Compreende um conjunto de controladores SDN baseados em software que fornecem funcionalidades de controle e supervisão ao encaminhamento dos fluxos. O controlador é o principal componente responsável por estabelecer tabelas de fluxo, políticas de manipulação de dados, coletar informações e abstrair a complexidade da rede [KARAKUS e DURRESI, 2017a]. Possui *Application Programming Interface* (APIs) que permitem a comunicação entre todos os componentes da SDN. A API *Eastbound* permite a comunicação entre controladores e um plano de controle; Já a *Southbound*, permite a comunicação entre controladores e dispositivos de rede; por fim, a *Northbound*, fornece a comunicação entre controladores e aplicações [KARAKUS; DURRESI, 2017b].
- **Plano de aplicação:** Consiste em uma ou mais aplicações que interagem com o controlador para obter uma visão abstrata da rede em um processo de tomada de decisão. É uma camada de aplicações que define regras e oferece diferentes serviços,

como *firewall*, controle de acesso, sistemas de detecção de intrusão, QoS, roteamento, serviços de *proxy* e balanceadores de carga. Essa camada é responsável por abstrair o gerenciamento de controle de rede SDN através da API *northbound* [BENZEKKI; FERGOUGUI; ELALAOUI, 2016].

Alguns desafios devem ser superados pelas SDNs, dentre eles pode-se destacar, principalmente, os relacionados à confiabilidade, escalabilidade e segurança.

A confiabilidade desempenha um papel importante em qualquer desenvolvimento de software [RANA et al., 2019]. Caso ocorram falhas no sistema, os usuários devem ser informados e a solução executada automaticamente. Na SDN, apenas o controlador central é responsável por toda a rede, se a rede central falhar, poderá gerar impactos, como indisponibilidade, em toda a rede [RANA et al., 2019]. Para aumentar a confiabilidade, discussões acerca das funções do controlador devem ser exploradas.

A escalabilidade está relacionada ao crescimento da rede e ao seu gerenciamento. A visão centralizada da rede proporcionada pela SDN facilita sua configuração e manutenção, entretanto, controladores SDN podem, potencialmente, se tornar um gargalo na operação da rede. Com a expansão da rede, mais eventos e solicitações são enviadas ao controlador e, em algum momento, este pode não conseguir lidar com todas as solicitações [RANA et al., 2019].

Quanto à segurança, as interfaces abertas da rede SDN podem trazer novos tipos de ataques à rede que podem reduzir o seu desempenho. Soluções devem ser propostas na estrutura SDN a fim de garantir a integridade do software e a detecção e mitigação de ameaças à rede [RANA et al., 2019].

Diferentes trabalhos na literatura atual discutem como as redes móveis podem se beneficiar da tecnologia SDN, trazendo agilidade e flexibilidade para as MNOs que oferecem este serviço [NAMAL et al., 2016; AHMAD et al., 2015; AKYILDIZ; WANG e LIN, 2015].

1.3.2 *Network functions virtualization* (NFV)

A *Network functions virtualization* (NFV) é uma técnica emergente de virtualização que se baseia no conceito de assumir funções de rede que tradicionalmente são executadas em um *hardware* dedicado [KITINDI et al., 2017]. A NFV é promovida e mantida pelo *European Telecommunications Standards Institute* (ETSI) e oferece muitos benefícios as operadoras de rede, incluindo redução de custos de equipamentos, redução do consumo de energia e escalabilidade [NGUYEN; DO; KIM, 2016]. No cenário de NFV, os recursos físicos de *hardware*, incluindo armazenamento, processamento e rede, são virtualizados através de uma

camada de software denominada de *Hypervisor*, o que facilita o gerenciamento de funções de rede tradicionais.

Segundo [ZHANG, 2018], em ambientes NFV, as complexas funções de rede executadas em hardware especializado são decompostas em unidades funcionais menores e orquestradas dinamicamente em uma infraestrutura virtualizada de nuvem e borda. Essas funções de rede são executadas como software em uma Máquina Virtual (VM)”.

Essa proposta simplifica o provisionamento e gerenciamento dos serviços de rede, dado que não se faz necessário a utilização de hardware específico para cada função da rede, tornando possível utilizar hardwares genéricos e compartilhados para essas funções. Ademais, essa abordagem reduz custo operacional e proporciona maior flexibilidade, personalização e escalabilidade, e, por conseguinte, reduz a complexidade da implementação de novos serviços.

No contexto do C-RAN, a NFV é implementada no nível do BBU pool. Cada BBU é um nó virtual, enquanto a comunicação entre nós é realizada por meio de *links* de rede virtuais. O *pool* opera em uma máquina física que compartilha processador, memória e recursos de rede entre várias BBUs [SALMAN, 2016]. Isso permite que as MNOs forneçam suporte eficiente a muitos RRHs simultaneamente, adequando o tráfego da rede a demandas flutuantes causadas pelo “efeito maré” [CHINA, 2011].

Para [NGUYEN; DO; KIM, 2016], as tecnologias SDN e NFV são paradigmas complementares entre si e visam acelerar uma abordagem baseada em *software* para tornar a rede mais escalável, flexível, eficiente e inovadora. A Figura 10 apresenta o compartilhamento de recursos de máquinas físicas em um BBU *pool* de BBUs virtuais.

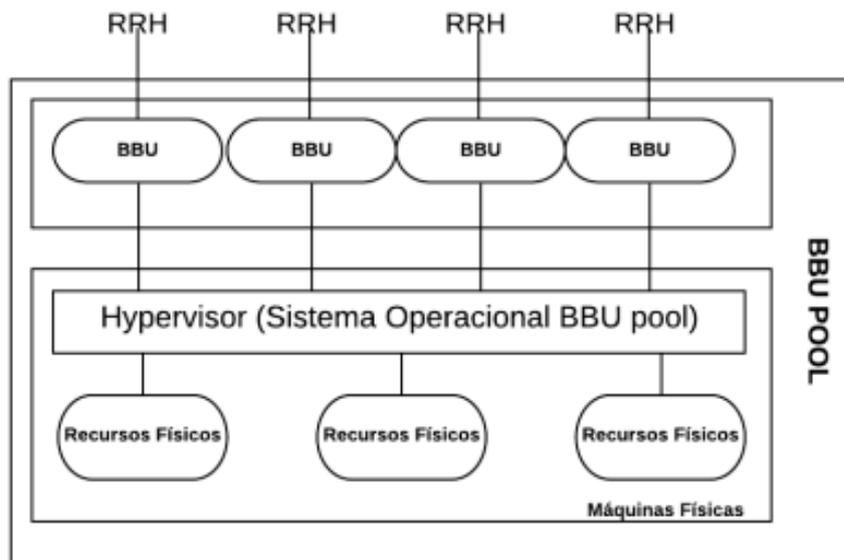


Figura 10. Virtualização do BBU pool em C-RAN, adaptado de [SALMAN, 2016].

Assim como a SDN, a NFV também tem desafios a serem superados, dentre eles se destacam a otimização da operação e manutenção da rede, onde se destacam principalmente os relacionados à [ZHANG, 2014]:

- **Eficiência:** a plataforma NFV deve fornecer um *Service Level Agreement* (SLAs) de desempenho e disponibilidade idênticos aos oferecidos por serviços dedicados.
- **Escalabilidade:** a plataforma deve suportar um grande número de VNFs e escalar a rede eficientemente à medida que o número de aplicações/volume de tráfego aumenta.
- **Disponibilidade:** a plataforma deve prover disponibilidade fim-a-fim de todos os recursos da rede.
- **Elasticidade:** a plataforma NFV deve poder aproveitar os benefícios da execução de instâncias na nuvem, permitindo que a mesma instância da rede atenda a vários usuários finais, a fim de maximizar a utilização dos recursos.

2.4 Considerações finais do capítulo

Este capítulo teve como objetivo fundamental apresentar os conceitos utilizados no desenvolvimento desse trabalho relacionados a redes móveis, referenciando as tecnologias relacionadas ao cenário das gerações móveis, com ênfase para o 5G, C-RAN, SDN E NFV. No próximo capítulo, será tratado o conceito de inteligência computacional, metaheurística, e Algoritmo *Particle Swarm Optimization* (PSO), elementos fundamentais para o desenvolvimento deste trabalho.

2. META HEURÍSTICA

Neste capítulo, será conceituado o termo meta-heurística e apresentado o algoritmo de otimização por enxame de partículas (PSO) e sua variante, o método da Otimização Discreta por Nuvem de Partículas (DPSO).

3.1 Inteligência Computacional (IC)

A Inteligência Computacional (IC) pode ser entendida, de um modo simplista, como uma especialização de uma área mais ampla de conhecimento, que é a Inteligência Artificial (IA). Nas palavras de [DUCH, 2007], definições de ramos da Ciência não são forjadas a priori, mas desenvolvem-se aos poucos, num processo de compartilhamento e de agrupamento de interesses comuns. Reconhecendo que o termo tem significados diferentes para diferentes pessoas, esse autor formula uma definição ampla para a Inteligência Computacional como sendo “um ramo da Ciência da Computação que estuda problemas para os quais não existem algoritmos computacionais eficientes”.

Em sua pesquisa sobre publicações científicas que possuem em seu título o termo “Inteligência Computacional”, abrangendo um período de mais de duas décadas (desde 1984 até 2007), [DUCH, 2007] constata que, inicialmente, o termo era praticamente indistinto de “Inteligência Artificial”, mas, ao longo do tempo, foi tomando contornos mais definidos, passando a englobar principalmente Redes Neurais Artificiais (RNA), a Lógica Nebulosa (LN) e a Computação Evolucionária (CE), em especial os Algoritmos Genéticos (AG). Essa tendência foi igualmente percebida por [VANECK et al., 2006]

Algoritmos Genéticos, AGs, são métodos de otimização e busca, inspirados nos mecanismos de evolução de populações de seres vivos. Foram introduzidos por [HOLLAND, 1975] e popularizados por um dos seus alunos, Goldberg (1989). Estes algoritmos seguem o princípio da seleção natural e sobrevivência do mais apto, declarado em 1859 pelo naturalista e fisiologista inglês Charles Darwin em seu livro "A Origem das Espécies". De acordo com Charles Darwin, "Quanto melhor um indivíduo se adaptar ao seu meio ambiente, maior será sua chance de sobreviver e gerar descendentes" [LACERDA et al., 1999].

3.2 Meta-Heurística

O nome Heurística é derivado da palavra grega *heuriskein*, que significa descobrir. Hoje esse termo é usado para descrever um método “que, baseado na experiência ou julgamento, parece conduzir a uma boa solução de um problema, mas que não garante produzir uma solução

ótima” [FOULDS, 1984]. Também podemos considerar esse termo como associado a um conhecimento circunstancial, não verificável, nem matematicamente verificável.

Já o termo “meta-heurística” deriva da composição de duas palavras gregas: “heurística”, já explicada no parágrafo anterior, e o prefixo “meta”, que significa “após”, indicando um nível superior de descoberta [BECCENERI at al., 2008].

Muitas definições podem ser dadas sobre o termo “meta-heurística” na computação, como, por exemplo: “Uma meta-heurística é um conjunto de conceitos que podem ser usados para definir métodos heurísticos, os quais podem ser aplicados a um conjunto amplo de diferentes problemas. Em outras palavras, uma meta-heurística pode ser vista como uma ferramenta algorítmica geral que pode ser aplicada a diferentes problemas de otimização, com modificações relativamente pequenas para torná-la adaptável a um problema específico”. Resumindo, podemos dizer que meta-heurísticas são mecanismos de alto nível para explorar espaços de busca, cada uma usando um determinado tipo de estratégia [BECCENERI at al., 2008].

Um dos desafios na aplicação de uma meta-heurística é encontrar o equilíbrio ideal entre diversificação e intensificação. A estratégia de busca de uma meta-heurística depende da filosofia empregada por ela, e seu objetivo é escapar dos mínimos locais a fim de proceder a exploração do espaço de busca por soluções ainda melhores [BECCENERI at al., 2008].

Para problemas de otimização que pertencem à classe de complexidade N P-Difícil, encontrar boas soluções, que não são soluções ótimas, pode ser útil em muitas situações. Meta-heurísticas são procedimentos genéricos ou frameworks que guiam outras heurísticas, ou seja, procedimentos computacionais, usualmente de busca local, de otimização global para busca de boas soluções (podendo ser a solução ótima) que se adequam a diversos problemas de otimização [BECCENERI at al., 2008].

As duas principais características em uma meta-heurística são a capacidade de diversificação e intensificação. A diversificação está relacionada a variabilidade de amostragem de soluções no espaço de busca e é importante na fuga de ótimos locais. A intensificação consiste na capacidade de busca intensa do método em uma região considerada promissora do espaço de busca. O equilíbrio desses dois fatores é fundamental para uma meta-heurística eficiente [BECCENERI at al., 2008].

Talvez o modo mais intuitivo de se classificarem meta-heurísticas seja com base nas origens do algoritmo. Nesse sentido, podemos diferenciar os algoritmos inspirados na natureza, como os Algoritmos Genéticos [HOLLAND, 1975]. Diz-se que uma meta-heurística é bioinspirada quando suas regras de busca tentam simular alguns aspectos do comportamento de

seres vivos como, por exemplo, o Algoritmo dos Pássaros ou Algoritmo *Particle Swarm Optimization* (PSO) [EBERHART et al., 2001], uma meta-heurística adequada para problemas de otimização contínua.

3.2.1 Algoritmo *Particle Swarm Optimization* (PSO)

Introduzido por [EBERHART et al., 1995], o *Particle Swarm Optimization* (PSO), surgiu de experiências com algoritmos modelados a partir da observação do comportamento social de determinados tipos de pássaros que, através da troca de informações entre indivíduos de uma população, determinam qual trajetória cada um deles deverá tomar no espaço de busca. E para criar a técnica e assim aplicá-la em problemas de otimização contínua, James Kennedy e Russell Eberhart seguiram o modelo de C. W. Reynolds, utilizando as regras de alinhamento e coesão.

O PSO enquadra-se como uma metaheurística, ou seja, uma heurística superior às convencionais, pois possui características de busca que têm maior chance de atingir a solução ótima global, diferentemente dos métodos clássicos como os baseados em gradiente [SERAPIAO, 2009]. E por ser um método heurístico estocástico, não determinístico, não garante, portanto, que a solução ótima será encontrada, mas sim que tenderá a uma solução ótima ou quase ótima.

O método é composto de partículas representadas por N indivíduos, também chamados de vetores, que definem a velocidade atual de cada partícula, e de vetores de localização, atualizados segundo sua velocidade atual, seu aprendizado pessoal e o aprendizado adquirido pelo bando, esse algoritmo engloba conceituações simples e pode ser implementado em poucas linhas de programação, requerendo apenas operadores matemáticos simples, sem grande custo computacional.

Na simulação do comportamento das aves, onde cada ave fará sua contribuição no processo de busca, dependendo de sua *fitness function*, ou função aptidão, sendo essa a métrica para avaliação do desempenho da solução, de modo que os indivíduos com maior aptidão são as melhores soluções da população [EIBEN et al., 2003]. A aptidão geral dos indivíduos é usada para a identificação do centro de alimentos, que é considerado como a melhor estimativa global. A implementação da otimização de enxame de partículas é detalhada da seguinte forma:

- Passo 01: Inicialização dos parâmetros [I_{max}], número de partículas [M], velocidades iniciais das partículas [V] e as posições iniciais das partículas [X];
- Passo 02: Repetir o Passo 03 ao Passo 07 de $i = 1$ para I_{max} ;
- Passo 03: Calcular o **fitness** para cada partícula;
- Passo 04: Atualizar a melhor posição da partícula P_{id}^t ;
- Passo 05: Atualizar a melhor posição para todo o enxame P_{gd}^t ;
- Passo 06: Atualizar a posição de X e a velocidade V ;
- Passo 07: Verificar se o critério de parada foi satisfeito.

No PSO, um indivíduo é usado para representar a potencial solução do problema a ser otimizado. Os parâmetros utilizados neste algoritmo são: o tamanho da população [N], o peso de inércia para limitar o valor da velocidade das partículas [w], o coeficiente cognitivo referente à importância da informação local ou melhor posição encontrada pela própria partícula [c_1], o coeficiente social referente à importância da informação global ou melhor posição encontrada dentre todas as partículas [c_2], e a taxa de velocidade [v_r]. Além destes parâmetros, o algoritmo utiliza as variáveis aleatórias $rand_1$ e $rand_2$, que são geradas uniformemente no intervalo [0-1] e são responsáveis pelo não-determinismo do algoritmo [KENNEDY, 2011].

No modelo matemático adotado para o PSO, um “enxame de partículas” é gerado aleatoriamente em um determinado espaço de busca. Dentro desse espaço, cada partícula corresponde a uma possível solução, representada por sua posição no espaço de busca para um dado problema. As partículas têm associado um valor de velocidade, e realizam um deslocamento sob a ação de três vetores que se somam. Essas influências são inércia, memória e comparação [EBERHART et al., 2001]. A Figura 11 demonstra um exemplo de deslocamento de uma partícula, que se encontrava na posição $x_i - 1$, para a posição x_i fazendo uso dos vetores mencionados.

A velocidade de cada partícula, deve ser atualizada de acordo com a equação 1.

$$v_{k+1}^i = w \cdot v_k^i + c1 \cdot r1 \cdot (p^i - x_k^i) + c2 \cdot r2 \cdot (p_k^s - x_k^i) \quad (1)$$

Onde v_k^i é a velocidade atual da partícula; p^i é a melhor posição encontrada pela partícula i ; p_k^s é a melhor posição dentre todas as partículas na iteração k ; w é um parâmetro que representa a inércia da partícula e controla sua capacidade de exploração do espaço de

soluções, onde um valor alto de w determina uma busca global enquanto um valor baixo determina uma busca local, usualmente estes valores oscilam entre 0,4 e 1,4; $c1$ e $c2$ são parâmetros de confiança, definem o quanto uma partícula confia em si ($c1$) ou no bando ($c2$), usualmente escolhe-se $c1 = c2 = 2$ e $r1$ e $r2$ são números aleatórios compreendidos entre zero e um [ENGELBRECHT, 2007].

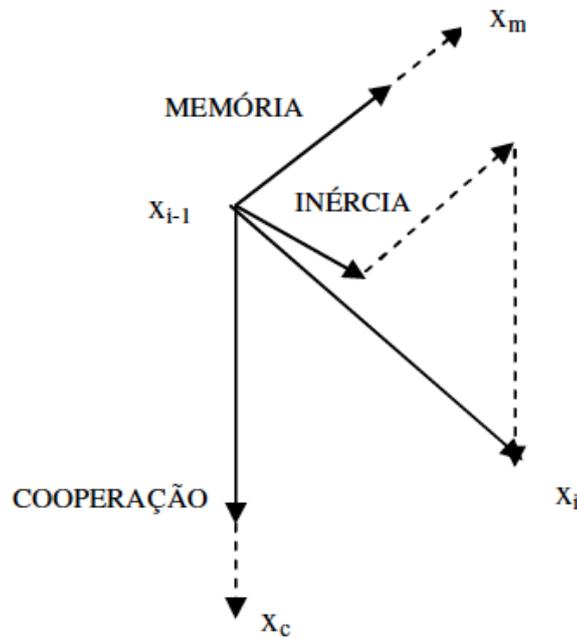


Figura 11 – Movimento de uma partícula, [EBERHART at al., 2001]

Para o cálculo da posição futura de cada partícula no algoritmo, é utilizada a equação 2:

$$x_{k+1}^i = x_k^i + v_{k+1}^i \tag{2}$$

Onde:

x_{k+1}^i é a posição de cada partícula i na iteração $k+1$;

v_{k+1}^i é o vetor velocidade da partícula (considera-se que já esteja multiplicado por um intervalo de tempo unitário).

Para a inicialização da população do enxame normalmente é obtida com as partículas dispostas aleatoriamente sobre o espaço de busca, e cada uma possui um vetor de velocidade aleatório inicial. As equações 3 e 4 mostram como são obtidos a posição e o vetor de velocidades iniciais [PRADO at al., 2005].

$$x_0^i = x_{min} + r_1[x_{max} - x_{min}] \tag{3}$$

$$v_0^i = \frac{x_{min} + r_2[x_{max} - x_{min}]}{\Delta t} \tag{4}$$

Onde:

r_1 e r_2 são números aleatórios entre 0 e 1;

x_{min} é o valor mínimo permitido para as variáveis de projeto;

x_{max} é o valor máximo permitido para as variáveis de projeto.

A depender de como as partículas se comunicam entre si e do problema a ser tratado, a busca pela solução ótima pode priorizar tanto a velocidade de convergência, a qualidade da solução ou ambas. Logo, a escolha da topologia do algoritmo, ou a forma de organizar a população, influencia na avaliação da velocidade das partículas. As principais topologias utilizadas como mecanismos de comunicação entre as partículas são: a topologia global e a topologia local. A figura 12 apresenta a estrutura de tais topologias.

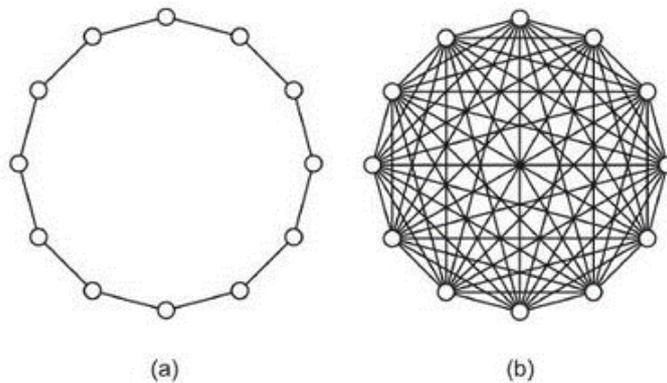


Figura 12 – Topologias do PSO [a] topologia local e [b] topologia global, adaptado de [EBERHART et al., 1995]

Na topologia global, todas as partículas possuem informações sobre todas as demais. Já na topologia local, uma partícula só possui informações de sua vizinha esquerda e direita. A escolha da topologia pode evitar que o algoritmo caia em mínimos locais.

O processo de execução do PSO é exibido na Figura 13, onde o fluxograma inicia com a criação de um enxame de partículas em que cada partícula corresponde a uma solução candidata. Em seguida as partículas averiguam a área da solução aleatoriamente com velocidades distintas, após passar pela *fitness* do algoritmo as partículas são direcionadas aos seus melhores valores de aptidão. A velocidade de uma partícula individual é alterada de modo

estocástico em cada iteração e sua atualização decorre da experiência da melhor posição histórica da respectiva partícula e a melhor experiência de posição de partículas vizinhas, isto é, a melhor posição global.

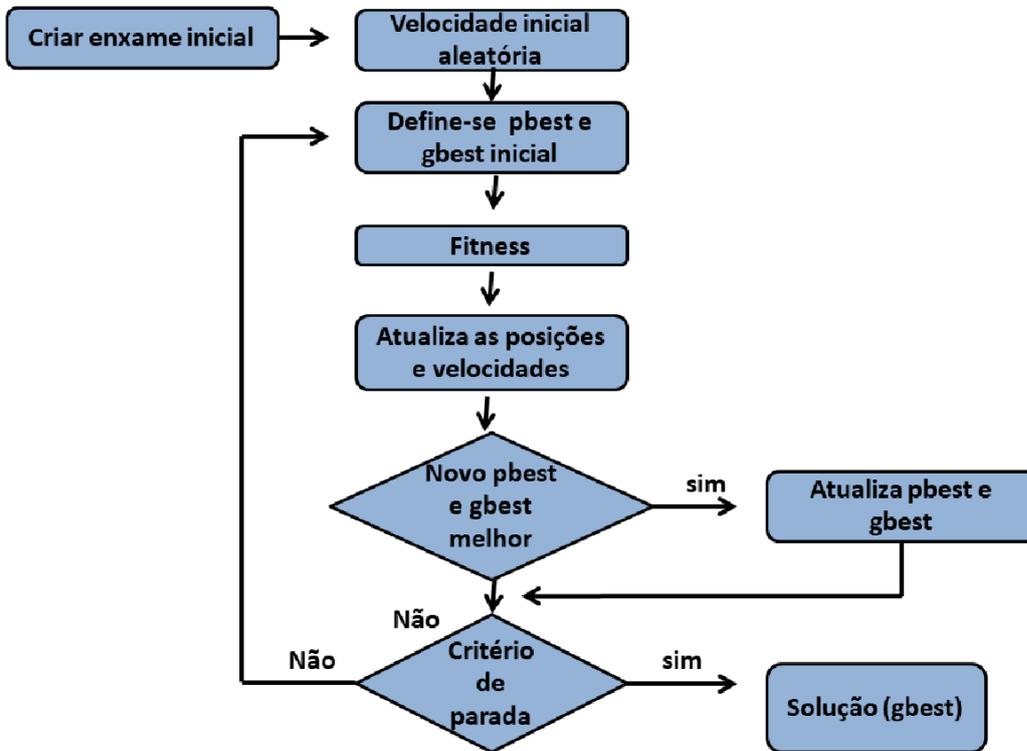


Figura 13 – Processo de execução do PSO, adaptado de [KHAN et al., 2015]

É apresentado como resultado, as melhores partículas, que são representadas por um vetor $\{S_1^{i+1}, S_2^{i+1}, \dots, S_N^{i+1}\}$, onde cada partícula caracteriza uma combinação possível de RRHs distribuídas nos setores de BBUs.

A posição de cada partícula é determinada pelo vetor e seu movimento pela velocidade da partícula. há três componentes principais para a atualização da velocidade:

- Por vezes referido como “inércia” e/ou “*momentum*”, modela a tendência da partícula de continuar na direção que tem viajado;
- O *pbest* (melhor local) de cada partícula indica o quão próximo que os dados já chegaram ao destino;
- E o *gbest* (melhor global), que só muda quando o valor do *pbest* de qualquer partícula se aproxima da meta, que é o *gbest*. A cada iteração do algoritmo, o *gbest* se aproxima do alvo, até que uma das partículas atinja o objetivo final.

O algoritmo PSO possui diversos parâmetros básicos listados a seguir:

- **Tamanho do Enxame:** Parâmetro que determina o número total de partículas que compõem o enxame. Quanto maior o enxame, maior a diversidade inicial, porém, o custo computacional também sobe [ENGELBRECHT, 2007]. Bergh e Engelbrecht (2001) demonstram empiricamente que um enxame com entre 10 a 20 partículas é eficiente para a maioria dos problemas.
- **Tamanho da vizinhança:** O tamanho da vizinhança vai depender do tipo de PSO utilizado. Se utilizada a versão PSO *gbest*, a vizinhança é o enxame inteiro. Se utilizado o PSO *lbest*, o tamanho da vizinhança deve ser estabelecido. Vizinhanças menores são mais eficientes na busca de soluções ótimas, mas possuem convergência mais lenta. Por outro lado, vizinhanças maiores promovem uma busca mais ampla, mas são mais susceptíveis a se prenderem a mínimos locais. Uma estratégia é usar tamanho incremental, iniciando com vizinhança pequena e incrementando-a proporcionalmente durante as iterações.
- **Condição de parada:** Essa condição é totalmente dependente do problema em questão e deve ser ajustada de acordo com o conhecimento prévio sobre os resultados que se desejam encontrar.
- **Coefficientes da velocidade:** Os coeficientes c_1 e c_2 que aparecem nas equações de atualização de velocidade do PSO controlam a influência que os componentes cognitivo e social exercem sobre a velocidade da partícula. Há vários esquemas utilizados, dentre pré-fixação dos valores a ajuste adaptativo [ENGELBRECHT, 2007].

3.2.2 Método da Otimização Discreta por Nuvem de Partículas

O Algoritmo *Discrete Particle Swarm Optimization* (DPSO), ou Otimização Discreta por Nuvem de Partículas, é uma variante do PSO e mantém suas características, isto é, não há diferenciação nas equações (1) e (2), propostas por [KENNEDY et al., 1995], é alto adaptativo, pois a dependência externa para a inicialização dos parâmetros é eliminada. Contudo, é importante ressaltar que, para o PSO, as variáveis devem ser contínuas; caso sejam discretas é preciso usar uma outra representação que, muitas vezes, depende do problema a ser estudado. Kennedy e Eberhart (1997), Maurice (2004) e Strasser et al. (2016) são alguns dos trabalhos que propuseram versões discretas do PSO e foge ao escopo deste trabalho representar as várias formulações propostas.

O Algoritmo do DPSO, inicia após a criação da solução inicial, ou nuvem de partículas, e enquanto o critério de parada não for atendido, isto é, enquanto o número de iterações corrente for menor que a quantidade máxima de iterações, cada partícula passa por uma avaliação. A melhor posição de uma partícula é armazenada na posição *pbest*, enquanto a melhor das posições já encontradas por todas as partículas do enxame é armazenada em *gbest*. Em seguida é realizada uma busca local na melhor solução até então encontrada, *gbest*. Por fim, atualizam-se os vetores velocidade e posição de cada partícula de acordo com o critério estabelecido pelos operadores no espaço de busca.

3.2.3 Função Aptidão

Uma função aptidão ou função *fitness* deve ser capaz de representar os requisitos necessários a que uma população deve se adaptar a fim de avançar para a geração seguinte. É importante que a função aptidão seja representativa e possa diferenciar com precisão os indivíduos (soluções) bons dos ruins. Uma função aptidão não ajustada na avaliação de indivíduos pode acabar descartando um indivíduo promissor, que poderia ajudar a encontrar soluções melhores para o problema, além do fato de consumir recursos em indivíduos que agregam pouco no desenvolvimento do algoritmo [EIBEN et al., 2003].

3.2.4 Otimização Bio-Inspirada

A computação bio-inspirada pode ser definida como sendo a convergência da Ciência da Computação, Biologia e Matemática para representar fenômenos estocásticos observados na natureza [DE CASTRO et al., 2005]. Neste contexto, este trabalho explora a computação bio-inspirada através do emprego do DPSO para otimização do balanceamento de cargas em C-RAN.

3.3 Considerações finais do capítulo

Neste capítulo foram apresentados conceitos importantes para o melhor entendimento do algoritmo PSO e sua derivação. Foram mostrados os principais passos e as equações necessárias à implementação do algoritmo, tornando esse capítulo um pequeno tutorial para a construção do algoritmo PSO, que será utilizado nas implementações desse trabalho, além de outros itens que corroboram com a proposta. No capítulo 4, será apresentado os trabalhos correlatos, dando ênfase aos trabalhos que serviram como referência para a proposta aqui apresentada.

4 TRABALHOS RELACIONADOS

Neste capítulo serão apresentadas soluções e propostas correlacionadas a esse trabalho, na área dessa pesquisa e suas contribuições. Serão mostradas diferentes abordagens sobre otimização do mapeamento BBU-RRH, através de diversas técnicas, para alcançar o balanceamento dos recursos de BBU, maximizando assim a qualidade do serviço. Para cada abordagem são feitas uma pequena descrição e uma avaliação sobre as estratégias adotadas. Também é discutida a contribuição deste trabalho em relação aos demais.

4.1 Pesquisas que abordam soluções para a otimização do balanceamento de carga em BBUs

Um algoritmo para diferenciação e integração de células (*CDI – Cell Differentiation and Integration*) e um mapeamento dinâmico de BBU-RRH, para maximizar o QoS, é proposto em [KHAN et al., 2018]. O método de previsão de carga nas RRHs proposto, possibilita um mapeamento de recursos de modo proativo, tornando a rede mais responsiva e adaptativa.

O desempenho do DPSO é testado com base em dois cenários e comparado ao algoritmo genético (GA) e ao algoritmo de busca exaustiva (ES). É possível observar que o DPSO oferece um desempenho ótimo para redes de pequena escala e desempenho próximo do ótimo para redes de grande escala, sendo muito mais rápido que os algoritmos GA e ES. Os resultados computacionais mostram um aumento significativo na taxa de transferência média da rede e uma diminuição notável na média de usuários bloqueados e na escassez de recursos médios (PRBs).

Em [YAO et al.; 2018], os autores otimizam o mapeamento BBU-RRH e a associação de usuários com o objetivo de minimizar o custo incorrido pela conta de energia das RRHs e aluguéis das BBUs sob a restrição de manter os padrões de QoS dos usuários, que é formulado através de programação linear (ILP). Os resultados ótimos são comparados com um algoritmo proposto pelos autores, que é baseado em um método de relaxamento do ILP. Os resultados se aproximam da solução ótima encontrada pelo ILP, reduzindo os custos de operação do sistema, como os autores demonstram ao variar a latência mínima aceitável para os clientes da rede e a taxa de chegadas de clientes.

Os autores, então, formulam a virtualização como um problema de empacotamento, onde cada BBU é um pacote com recursos computacionais finitos e em seguida, desenvolvem uma otimização utilizando a heurística do recozimento simulado. Este trabalho evidencia a aplicabilidade de técnicas de Inteligência Computacional para otimizar o balanceamento de carga na BBU Pool, bem como a importância de gerar dinamicamente novas alocações para

adequar a rede às variações nas cargas ao longo do dia (efeito de maré). No entanto, este não leva em consideração prover QoS mínimo a usuários individualmente e não trata seus resultados em métricas mais comparáveis à realidade, limitando-se a capacidade computacional da BBU em termos de milhões de operações por intervalo de tempo (*million operations per timeslot - MOPTS*).

Um esquema de mapeamento dinâmico BBU-RRH, usando uma abordagem de empréstimo, para alocar dinamicamente os recursos a RRHs de maneira inteligente e eficiente, é proposto em [CHEN et al., 2018]. Com a centralização do gerenciamento dos recursos da rede, o grupo menos ocupado empresta os recursos para o grupo mais ocupado, que toma emprestado esses recursos. Os resultados mostram que o mecanismo proposto reduz significativamente o desperdício do uso de recursos e melhora o rendimento da rede.

[DA PAIXÃO et al., 2018] mostra em seu artigo o mapeamento BBU-RRH, que é otimizado visando o balanceamento da carga e a minimização do número de chamadas bloqueadas, através de um algoritmo PSO. Dessa forma, o *Key Performance Indicator* (KPI) de chamadas bloqueadas foi utilizado para mensurar as métricas de QoS, maximizando-o. Os resultados obtidos indicam redução de até 100% das chamadas bloqueadas e uma distribuição de carga mais igualitária entre as BBUs.

Em [DUAN et al., 2016], a fim de aumentar a utilização dos recursos e melhorar o desempenho da rede, é proposto um modelo geral onde há um escalonamento dos usuários até a borda da rede de uma região metropolitana, para reduzir a sobrecarga desnecessária de roteamento inter-BBU. Baseado nas informações dos serviços utilizados pelos usuários na rede, um controlador é configurado para implementar o gerenciamento desses recursos ou serviços inter-BBU. O uso do controlador visa resolver o problema de balanceamento, dependendo da demanda ou carga. Os resultados da simulação avaliaram a viabilidade deste mecanismo, que se mostrou promissor, em comparação aos mecanismos tradicionais utilizados para este fim, reduzindo o número de bloqueios de serviços, melhorando significativamente a utilização dos recursos e o desempenho da rede.

Já em [BOULOS et al., 2015], os autores formulam o mapeamento BBU-RRH, a fim de melhorar a utilização dos recursos de transmissão e consumo de energia, através de soluções heurísticas, sem comprometer a qualidade de serviço do usuário. Inspirado nas estratégias *Best Fit Decreasing* (BFD) e *Worst Fit Decreasing* (WFD), é proposto um algoritmo heurístico que tenta preencher todas as lacunas não utilizadas, de forma a aproveitar ao máximo as RBs não usadas. A heurística proposta fornece desempenho próximo à solução ótima, em condições de baixa e média carga, com menor complexidade computacional. Os autores conseguem fornecer

o mesmo nível de qualidade de serviço. Os resultados da simulação ilustram os benefícios do agrupamento de RRHs, trazendo uma melhoria significativa na eficiência energética e na eficiência de utilização, em condições de baixa e média carga, sem comprometer o QoS do usuário.

O re-mapeamento dinâmico das conexões entre RRHs e BBUs é fortemente abordado em [KHAN et al., 2015], aqui estratégias são utilizadas para minimizar o número de usuários bloqueados e reduzir o *handover*, então, é formulado um problema de otimização para avaliar o QoS da C-RAN, são propostas soluções utilizando DPSO e AG, para resolver o problema de alocação das RRHs, e seus desempenhos comparados. O artigo explora o roteamento dos dados em uma arquitetura C-RAN para permitir o balanceamento de carga na rede. Khan, Alhumaima e Al-Raweshidy [2016] propuseram uma arquitetura que se auto organiza, automaticamente, para equilibrar o tráfego dos dados na rede ao reduzir o número de chamadas bloqueadas e como consequência melhorando a qualidade de entrega do serviço. Na arquitetura todos os BBUs são conectados a um mesmo gerenciador de *hosts*. Esse gerenciador monitora a carga de cada BBU e é responsável por realizar a conexão entre BBU e RRH mais adequada, através do uso de um algoritmo genético. A métrica de carga em cada BBU para a tomada de decisões é a quantidade total de usuários ativos, simultaneamente conectados, em cada um dos BBUs.

Através de testes, constatou-se que o DPSO forneceu uma convergência perceptivelmente mais rápida e melhor em comparação com o GA. Tanto o GA quanto o DPSO forneceram uma solução quase ótima para redes maiores. No entanto, o DPSO superou o GA em todos os cenários de rede, convergindo para uma solução ótima, com uma configuração adequada de RRHs e BBUs, reduzindo as chamadas bloqueadas de 80 para 0. Ainda, os *links* da rede recebem o tráfego de dados de modo equilibrado, impactando na melhora da qualidade na entrega do serviço, corroborando a abordagem proposta aqui neste trabalho, cuja técnica se assemelha ao proposto em [KHAN et al., 2017], no entanto, o problema de otimização foi formulado de modo diferente ao proposto neste trabalho. No trabalho é evidenciado que o DPSO entrega soluções igualmente adequadas mais rapidamente que o AG em qualquer cenário.

Em [NAMBA et al., 2012], é proposto um esquema de comutação semi-estático e adaptativo para C-RAN, para ajustar a configuração da BBU-RRH, com base nas cargas de tráfego no horário de pico, dentro de um determinado intervalo de tempo, sua eficácia foi avaliada por meio de simulações. Nesse esquema um RRH é comparável a uma célula, e a distribuição de tráfego é feita considerando o tráfego típico em uma área comercial. Foi possível verificar que os sistemas propostos reduziram o número de BBUs em 26% e 47%,

respectivamente, para cargas de tráfego de escritório em comparação com a implantação de células convencionais.

O trabalho de [SUNDARESAN et al., 2016], propõe uma estrutura leve e escalável que utiliza estratégias de transmissão ótimas, via reconfiguração BBU-RRH, para atender perfis dinâmicos de tráfego de usuários. No trabalho de [RAO; SRINIVAS, 2022] é proposto uma otimização para reduzir o bloqueio de chamadas e também equilibrar a carga de BBUs aplicando algoritmos *Swarm Intelligence* (SI). Os resultados da simulação comprovam que o algoritmo de computação inspirado na natureza reduziu as chamadas bloqueadas e maximizou o equilíbrio da carga de processamento das BBUs.

O trabalho proposto em [MOUAWAD et al, 2022] apresenta uma seleção de pares RRH-Sector para novas conexões e estrutura de balanceamento de carga de rede C-RAN, que otimiza o desempenho da rede e a recompensa do operador sem afetar o QoS dos usuários. Na abordagem proposta, é selecionado o melhor par RRH-Sector para cada nova demanda de conexão considerando os objetivos do usuário e do operador em conjunto. Esta decisão é baseada em um algoritmo derivado do processo de decisão de *Markov* (MDP). A seguir, o problema de balanceamento de carga é abordado por meio da otimização do mapeamento dinâmico RRH-Sector-BBU formulado como um problema de otimização restrito linear baseado em números inteiros. As soluções para este problema são obtidas através da comparação com diversos algoritmos evolutivos. Os resultados da simulação mostram que o esquema de seleção do setor RRH proposto proporciona, na maioria dos casos, ganhos significativos, atingindo a solução ótima para o número de usuários bloqueados, número de *handovers* e o consumo de energia da BBU, quando comparado com o método tradicional baseado na intensidade do sinal recebido (RSS).

Por fim, o trabalho de [JIBRIL; HASSEN; TADESSE, 2021], destaca a importância dos algoritmos de Inteligência de Enxame para resolver problemas computacionais com ótimos resultados e propõem o mapeamento da alocação de recursos de BBU-RRH em C-RAN, aplicando algoritmos de inteligência de enxame, algoritmo proposto é o PSO. Utilizando como índice de desempenho da rede (KPI) a redução de chamadas bloqueadas em todos os setores, para ganho de Qualidade de Serviço (QoS). Ao equilibrar os setores no BBU-RRH, os resultados mostraram um desempenho melhor do que o existente anteriormente, antes do mapeamento. E como trabalhos futuros, os autores planejam otimizar o balanceamento de carga de setores para BBU-RRH, aplicando algoritmos de computação inspirados na natureza híbrida para obter melhores resultados.

Em nenhum dos trabalhos citados acima foi investigada uma proposta de um modelo que busca otimização do balanceamento de carga da rede nas BBUs, nos recursos de processamento, com o mapeamento BBU-RRH, utilizando o DPSO em conjunto com uma função objetivo, deixando o uso dos recursos equilibrados nas BBUs. Este, por sua vez, podem ser considerados uma das soluções para estes problemas.

Objetivando um melhor esclarecimento das diferentes vertentes dos trabalhos citados acima, a Tabela 1 mostra de maneira clara e concisa os principais trabalhos relacionados a esta pesquisa, ressaltando o diferencial do referido trabalho proposto nesta dissertação.

Tabela 1 – Síntese dos Principais Trabalhos Relacionados

Publicação	Objetivos	Principais Características	Desafios
[KHAN et. al., 2018]	Enfatiza a técnica de auto-otimização no C-RAN sobre a melhoria do desempenho da rede.	Algoritmo CDI fixo seleciona BBUs e RRHs para ativação e desativação, e a parte dinâmica executa o mapeamento de carga da rede visando o máximo QoS com o mínimo de handovers.	Utiliza arquitetura CRAN que requerem altos custos de CAPEX/OPEX e causam altos níveis de interferência na rede.
[YAO et al; 2018]	Minimizar o custo incorrido pela conta de energia das RRHs e aluguéis das BBUs sob a restrição de manter os padrões de QoS dos usuários.	Otimização do mapeamento BBU-RRH e a associação de usuários, formulada através de programação linear (ILP).	Reduzir os custos de operação do sistema ao variar a latência mínima aceitável para os clientes da rede e a taxa de chegadas de clientes.
[CHEN et al, 2018]	Esquema para introduzir NFV e SDN, para diminuir os custos com energia	Análise através de uma alocação de energia e controle de custos.	Foi realizado um mapeamento, visando diminuir os custos da infraestrutura e organizar a carga de dados.
[PAIXÃO et. al., 2018]	Balanceamento otimizado dos recursos entre os setores de mais de uma BBU, em conjunto com o balanceamento de carga da rede.	Reformula os KPIs, otimizando o balanceamento entre os setores.	Utiliza a arquitetura CRAN, que ocasiona altas interferências na rede.
[DUAN et al, 2016]	Aumentar a utilização dos recursos e melhorar o desempenho da rede	O uso de um controlador para resolver o problema de balanceamento e estabelecer um link de acordo com o atendimento ou não dos usuários.	Escalonamento dos usuários até a borda da rede de uma região metropolitana, para reduzir a sobrecarga desnecessária de roteamento inter-BBU.
[BOULOS et al, 2015]	Melhorar a utilização dos recursos de transmissão e	Formulação do mapeamento BBU-RRH inspirado nas estratégias <i>Best Fit Decreasing</i>	Preencher todas as lacunas não utilizadas, de forma a aproveitar ao

	consumo de energia, através de soluções heurísticas, sem comprometer a qualidade de serviço do usuário.	(BFD) e <i>Worst Fit Decreasing</i> (WFD).	máximo as RBs não usadas, em condições de baixa e média carga, com menor complexidade computacional.
[KHAN et. al., 2015]	Mapeamento otimizado para RRH-BBU.	Análise através dos KPIs, para balanceamento no C-RAN	Foi realizado um balanceamento sem se preocupar com o balanceamento entre os setores, ocasionando um desequilíbrio na rede
[NAMBA et al, 2012]	Avaliar a eficácia de um esquema para ajustar a configuração da BBU-RRH, com base nas cargas de tráfego no horário de pico, dentro de um determinado intervalo de tempo.	Proposto um esquema de comutação semi-estático e adaptativo para C-RAN.	reduziram o número de BBUs para cargas de tráfego de escritório em comparação com a implantação de células convencionais.
[SUNDARESAN et al, 2016]	Reconfiguração BBU-RRH.	Proposta estrutura leve e escalável que utiliza estratégias de transmissão ótimas, via reconfiguração BBU-RRH.	Atender perfis dinâmicos de tráfego de usuários.
[RAO; SRINIVAS, 2022]	Otimizar e maximizar o equilíbrio da carga de processamento das BBUs	Um algoritmo de computação inspirado na natureza, <i>Swarm Intelligence</i> (SI), é aplicado para reduzir as chamadas bloqueadas e maximizará o equilíbrio da carga de processamento das BBUs.	arquitetura C-RAN é um Mapeamento BBU-RRHs, numa arquitetura C-RAN, para evitar o bloqueio de chamadas e menor qualidade das conexões de rede.
[MOUAWAD et al, 2022]	Seleção de pares RRH-Sector para novas conexões e estrutura de balanceamento de carga de rede C-RAN	seleção do melhor par RRH-Sector para cada nova demanda de conexão considerando os objetivos do usuário e do operador em conjunto. Decisão é baseada em um algoritmo derivado do processo de decisão de Markov (MDP).	Atingir a solução ótima para o número de usuários bloqueados, o número de <i>handovers</i> e o consumo de energia da BBU.
[JIBRIL; HASSEN; TADESSE, 2021]	mapeamento da alocação de recursos de BBU-RRH.	algoritmos de inteligência de enxame (PSO).	alocação de recursos para BBU-RRH, redução das chamadas bloqueadas em todos os setores. Para ganho de Qualidade de Serviço (QoS), alcançado equilibrando igualmente nos setores BBU-RRH.
[PROPOSTA DA DISSERTAÇÃO]	Mapeamento BBU-RRH em redes móveis H-CRAN, considerando aspectos de rádio e BBU, aliado ao balanceamento de carga da rede.	Otimização da função objetivo, utilizando metaheurística, em conjunto com balanceamento de carga da rede e o mapeamento dinâmico entre BBU-RRH, buscando a maximização de carga.	Mapeamento dinâmico BBU-RRH, em redes móveis C-RAN, aliado ao balanceamento de carga da rede, buscando maximização de carga.

4.2 Considerações finais do capítulo

Nesta seção foram apresentados alguns trabalhos relacionados encontrados na literatura, que buscam a otimização do mapeamento BBU-RRH através de diversas técnicas e com diferentes objetivos. Este trabalho propõe um esquema de mapeamento BBU-RRH dinâmico em C-RAN, utilizando o DPSO, para alcançar o balanceamento dos recursos de BBU, maximizando assim o QoS.

5. OTIMIZAÇÃO BIO-INSPIRADA PARA MAPEAMENTO BBU-RRH E BALANCEAMENTO DE CARGAS NAS BBUs

Será descrita, ao longo deste capítulo, a modelagem do Sistema juntamente com os principais parâmetros utilizados na técnica de otimização por Enxame de Partículas, com o DSOP, para equilibrar a carga da rede nos recursos de processamento e mapeamento BBU-RRH, e a apresentação da Função Objetivo (FO) proposta neste trabalho.

5.1 Arquitetura do sistema

Impulsionadas pela demanda insaciável por altas taxas de dados, entretenimento em tempo real, conteúdo multimídia e proliferação de dispositivos móveis inteligentes, já foi visto no capítulo 1, que as MNOs estão experimentando um enorme aumento no tráfego de dados em suas redes com a chegada do 5G. E, por ter a capacidade de ajustar dinamicamente as conexões lógicas entre as BBUs e RRHs de acordo com as condições de tráfego na rede, a C-RAN é uma importante solução para habilitar essas redes de comunicação móvel.

Na arquitetura *Cloud Radio Access Network* ou C-RAN, o *Host Manager* monitora a carga em cada BBU e é responsável por escolher a configuração adequada do BBU-RRH. Cada BBU pode lidar com vários setores ao mesmo tempo. Cada setor contém vários RRHs e pode utilizar recursos BBU inteiros. Cada RRH pertence a apenas um setor em um determinado período de tempo [KHAN at al. 2015].

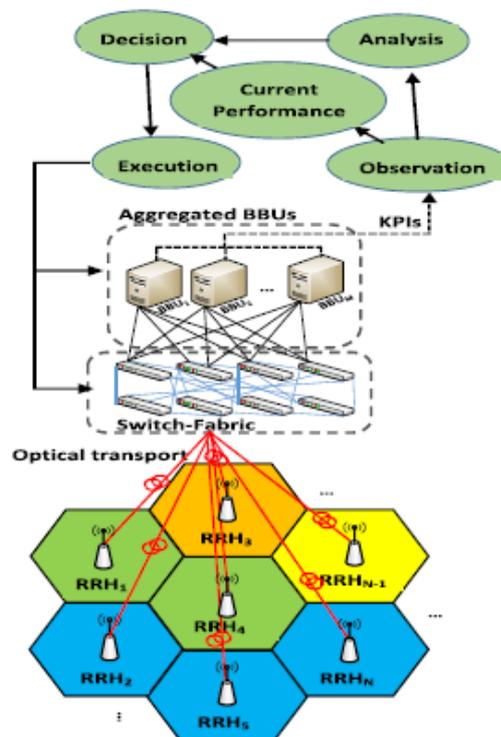


Figura 14 – Conceito genérico de uma auto-organização em C-RAN, [KHAN at al., 2017]

A Figura 14 mostra um conceito genérico de uma auto-organização em CRAN. Os RRHs são conectados ao pool de BBUs, agregado via *fronthaul*, que pode consistir em um *switch fabric* (incluindo uma rede de *switches*, divisores ópticos, multiplexadores) [SUNDARESAN, 2016], fibra de baixa latência e alta largura de banda, e ligações ópticas.

A carga em cada BBU é definida como o número total de usuários ativos, conectados simultaneamente aos setores das BBUs. No entanto, há uma limitação de hardware ou software no número de usuários ativos em cada setor. O balanceamento de carga e a escalabilidade de recursos tem sido bastante explorados na literatura, por se tratarem de fatores significativos para a redução de custos e o aumento do QoS. Isso pode ser facilmente visto em [KHAN et al., 2015; YAO et al., 2018; DA PAIXÃO et al., 2018; BOULOS et al., 2015]. O desempenho da rede impacta no QoS oferecido para os usuários, sendo assim as novas gerações de redes móveis buscam sanar os problemas que enfrentam para maximizar o QoS.

Neste trabalho, é apresentada um modelo de alocação de recursos, com o mapeamento BBU-RRH, através do DPSO que otimiza a função objetivo proposta., a fim de balancear a carga nas BBUs. Para isso é definido um novo Indicador Chave de Desempenho (KPI), que visa equilibrar o tráfego da rede reduzindo o número de chamadas bloqueadas e melhorando seu QoS. Com base nesse KPI, o gerenciador de host identifica a configuração adequada do BBU-RRH e equilibra a carga na rede.

5.2 Modelagem BBU-RRH

O desempenho de uma rede reflete seu QoS, que pode ser representado por KPIs, então se torna necessário que uma rede identifique padrões de tráfego de forma proativa, a fim de evitar situações que ameacem o desempenho e a disponibilidade da infraestrutura e serviços de rede. Com base no KPI, um servidor chamado “gerenciador central”, identifica a configuração BBU-RRH para equilibrar a carga na rede oferecendo, deste modo, um serviço de melhor qualidade (ou utilizando os recursos de forma mais eficiente). Assim, a função objetivo (FO) apresentada nesta dissertação, busca maximizar o balanceamento da rede.

5.2.1 KPI

Segundo [KHAN et al., 2017], pode-se avaliar o índice de desempenho da rede, através de vários KPIs, que podem ser considerados para medir o QoS da rede. O *framework* é modelado como um problema geral de otimização multiobjetivo incluindo vários critérios como se concentrar na avaliação de QoS com base em KPIs relevantes para o balanceamento de carga acionado por probabilidade de bloqueio.

O KPI de chamadas bloqueadas, visa avaliar a capacidade das BBUs e de seus setores (HCs). Quando tal valor é excedido, os usuários são considerados bloqueados [KHAN et. al., 2017]. O número de usuários bloqueados (BU) na rede no tempo $t + 1$, é dado por:

$$BU = \sum_s \max \left[\left(\left(\sum U_i R_{iS}^{t+1} \right) - HC_s \right), 0 \right] \quad (5)$$

Onde $i = 1, 2, \dots, M$, e $s = 1, 2, \dots, S$, em seguida, o KPI para usuários bloqueados [KPI_{BU}] pode ser apresentado como:

$$KPI_{BU} = \{1 \text{ If } BU = 0 \frac{1}{BU} \text{ otherwise} \quad (6)$$

Com o propósito de balanceamento de carga entre as BBUs, e aumento de seu desempenho, dessa forma, o número de chamadas bloqueadas, atendendo o QoS requerido por cada chamada.

No que se refere a avaliação e implantação de arquitetura C-RAN em cenários de balanceamento de carga, poucos artigos são encontrados na literatura. Sendo assim, essa dissertação adotou o mesmo KPI utilizado em [KHAN et. al., 2015], definido como o número de chamadas bloqueadas, seguindo os mesmos critérios. Baseado neste KPI, o *host manager* identifica as configurações adequadas para o conjunto BBU-RRH e faz o balanceamento dos recursos das BBUs.

5.2.2 Função Objetivo

Utilizando o DPSO para o mapeamento BBU-RRH, a função objetivo (FO) proposta por este trabalho, dada pela equação 7, busca o balanceamento de carga, minimizando o somatório do quadrado da diferença que, quanto mais desbalanceada a BBU estiver, ou seja, mais longe de X, maior será o quadrado da diferença.

O mapeamento BBU-RRH correspondente ao período t é representado como: $S = \{Rh_1, Rh_2, \dots, Rh_i\}$, no qual $Rh_i \forall i \in R$, representa a BBU que está alocada para a RRH i . Conhecendo o vetor S, é possível encontrar a carga total sendo processada no instante t e desta forma, o novo mapeamento no instante $t + 1$ é encontrado na FO dada a seguir:

$$\text{MinF1} = \sum_{j=1}^B [\text{CB}_j - \text{HC}]^2 \quad (7)$$

Onde CB_j é a carga sendo processada na BBU_i , e é dada por:

$$\text{CB}_j = \sum_{i=1}^R C_i * R h_i^t, \forall j \in B \quad (8)$$

E entregue pela FO o melhor balanceamento da rede, entretanto, algumas restrições para o problema são apresentadas, limitando o campo de soluções viáveis, sendo atribuição da função objetivo selecionar a melhor entre as soluções viáveis. Tais restrições são apresentadas a seguir:

$$\sum_{i=1}^B C_{r,i} = 1, \forall r \in B \quad (9)$$

$$\sum_{i=1}^R A_{r,i} * C_i \leq \text{HC}, \forall j \in B \quad (10)$$

A equação 9 garante que cada RRH vai ser atendida por uma única BBU. E a equação 10 garante que a carga da BBU não vai ser maior que sua capacidade máxima, definida por HC . Desta forma, as soluções apresentadas estarão dentro do conjunto que respeita as restrições do problema. Para buscar a melhor solução, foi utilizado o DPSO como uma meta-heurística.

5.2.3 Benchmark e Desvio Padrão

As equações 11, 12 e 13, dadas a seguir, são utilizadas como *benchmark* neste trabalho e apresentadas em [KHAN et al. 2018; DA PAIXÃO et al. 2018; KHAN et al. 2015], respectivamente, e buscam maximizar o balanceamento da rede, garantindo que a carga na rede seja equilibrada de forma a alcançar os melhores índices de balanceamento, onde CB_j é dada pela equação 8:

$$\text{MaxF2} = \frac{\sum_{j=1}^B [\text{CB}_j]^2}{B * \sum_{j=1}^B \text{CB}_j^2} \quad (11)$$

$$\mathit{MinF3} = \sum_{j=1}^B [X - CB_j] \quad (12)$$

$$\mathit{MinF4} = \sum_{j=1}^B [HC - CB_j] \quad (13)$$

A equação 11 utiliza o índice de justiça proposto por Jains [KHAN et al., 2018], que determina o nível de balanceamento de carga na rede em um determinado momento e é avaliado usando a distribuição de carga em todas as células. Este índice é um critério de desempenho utilizado em todos os esquemas de alocação de recursos, incluindo os sistemas de computadores distribuídos. O índice de justiça não possui dimensão, e pode assumir valores entre $[0 - 1]$.

As equações 12 e 13 são similares e buscam o balanceamento através da diferença da carga da BBU em relação a uma variável [DA PAIXÃO et al., 2018] e a uma constante [KHAN et al., 2015]. Para comparação entre as FOs é utilizado o desvio padrão (DP), que é uma medida de dispersão em torno da média amostral, sendo assim, é considerado que quanto mais próximo de zero, melhor balanceada a rede está [KHAN et al., 2018]. O desvio padrão é calculado como:

$$DP = \sqrt{\frac{\sum_{j=1}^B [CB_i - X]^2}{B}} \quad (14)$$

Sendo a configuração da BBU-RRH no momento t conhecida, é possível identificar a nova configuração no momento $t + 1$, sendo este o objetivo final buscado pelas FOs. O modelo aqui proposto consiste em 5 parâmetros e 1 variável, que definem o cenário, descritos a seguir:

- Parâmetros

- R é o total de RRHs no cenário
- B é o total de BBUs no cenário
- HC_j representa a capacidade total que a BBU_j consegue processar
- X representa a média da carga total no instante t presente na rede
- C_i representa a carga da RRC_i no instante t

- Variável

A_{ij} é uma variável binária, que quando igual a 1, representa que a RRH_i está sendo processada na BBU_j .

5.2 Simulação do modelo

A simulação desempenha um papel fundamental em termos de escalabilidade, reprodutibilidade, rápida prototipagem e educação. Nesta dissertação utilizou-se a simulação computacional que, segundo [SILVA, 2018], tem menor custo e complexidade, o que torna mais prática a modelagem e estudo de um sistema celular móvel.

Para a elaboração do cenário de simulação do modelo aqui proposto, foi escolhido a ferramenta NS-3, por melhor se adequar às necessidades deste trabalho, por sua natureza *open-source*, que contribui para melhores documentações e comunidade mais ativa, e por ser extensível, é possível adaptá-lo para cenários que não foram desenvolvidos previamente [BILALB et al., 2013]. Para a elaboração dos gráficos foi usado o MATLAB® (do inglês *MATrix LABORatory*), que é uma poderosa ferramenta matemática e uma IDE de desenvolvimento [HIGHAM; HIGHAM, 2016].

O cenário adotado para a simulação do modelo apresentado nessa dissertação, foi o mesmo utilizado em [KHAN et al., 2015], por ser um cenário bastante utilizado na literatura, o que traz mais confiabilidade aos resultados. Assumiu-se uma área geográfica de 1000m x 1000m, onde um conjunto de 19 RRHs foram dispostas, uma BBU Pool contendo duas BBUs gerenciando três setores cada, com as RRHs distribuídas entre 6 setores, em uma faixa de potência padrão. A BBU Pool é representada por um conjunto de nós, sendo esses switches do módulo OFSwitch13, de modo a formar as estruturas de tomada de decisão e controle de fluxo inerentes ao modelo de uma BBU Pool, como pode ser visto na Figura 16 abaixo. O tráfego gerado pelos clientes nas RRHs é um fluxo UDP contínuo.

A configuração inicial do BBU-RRH gera 80 chamadas bloqueadas em toda a rede com um valor de avaliação de QoS de 0,0125. Após a execução do algoritmo, observa-se que a rede é balanceada reconfigurando as conexões BBURRH, encontrando a solução ótima, reduzindo as chamadas bloqueadas a zero e melhorando a QoS da rede, ou seja, QoS=1.e o KPI de chamadas bloqueadas é traduzido para pacotes perdidos na simulação, onde, após excedido o limite de conexões no setor da BBU, os pacotes dos clientes não alocados são perdidos na rede.

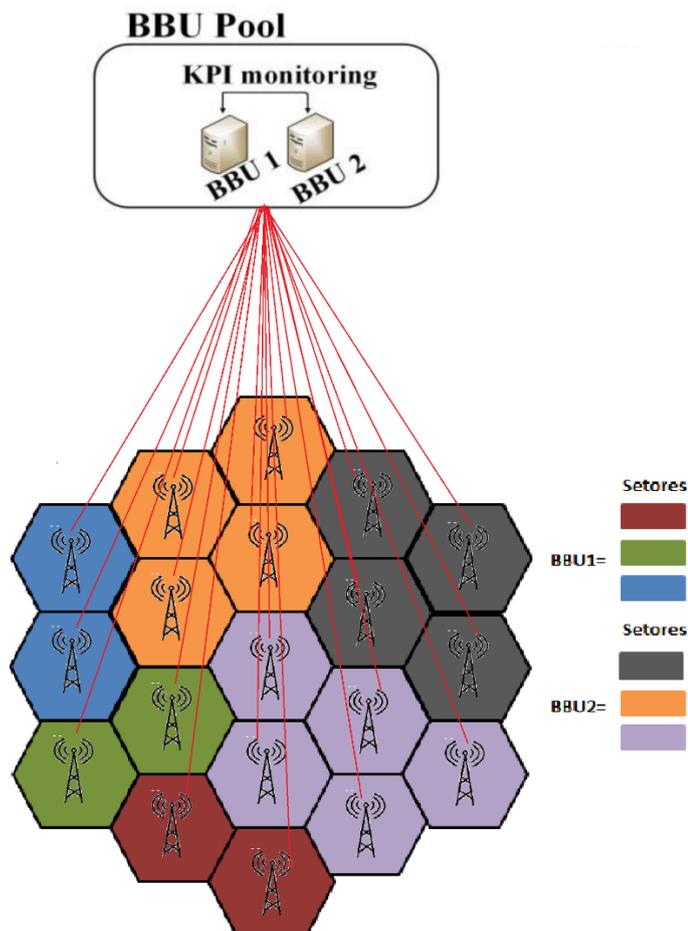


Figura 15 – Demonstração do cenário proposto.

Para a avaliação de desempenho, essa dissertação utilizou três problemas de benchmark, apresentados na seção 5.2, item 5.2.3, além dos 6 cenários propostos na coleta de resultados. Equações 11, 12 e 13, a fim de analisar e verificar o desempenho do algoritmo DSOP para o mapeamento BBU-RRH, a função objetivo (FO) proposta por este trabalho, dada pela equação 7.

Para a coleta de resultados foram gerados 6 cenários, contendo 300, 600, 900, 1200, 1500 e 1800 usuários, respectivamente, cobrindo, desta forma, um conjunto maior de possíveis cenários de alocação, em uma implantação real de C-RAN, com o intuito de obter o desempenho da rede, pois com a diversidade no número de usuários é possível simular representações variadas no tráfego da rede ao longo do dia, para regiões densamente povoadas a regiões esparsamente povoadas, em regiões com vastas áreas ou pequenas áreas.

Os cenários mencionados anteriormente, foram baseados em [KHAN et al., 2018] e [GONÇALVES, 2018], por terem obtido um modelo de balanceamento de carga ótimo para sistemas celulares C-RAN, usando um método de otimização de dimensão infinita para obter

uma melhor distribuição de recursos. Os resultados mostraram que esta técnica recomendada alocou recursos de forma eficaz sem agregar mais custos ao setor operacional. Destas referências foram extraídos parâmetros relacionados à caracterização do tráfego gerado por um usuário individual, ao qual é atribuído 400 *Kbps* (*Kilo bit* por segundo), e aos números de BBUs e RRHs ativas, presentes na proposta.

O valor das constantes de aceleração foi definido como $C1 = C2 = 2$, valor adotado em KHAN et al., 2015]. O tamanho do enxame foi definido com base na análise realizada por [PIOTROWSKI et al., 2020], que observou que os criadores do PSO, autores das variações do algoritmo e pesquisadores utilizaram tamanho do enxame de 30 em seus experimentos de sucesso.

5.4 Algoritmo DPSO

O algoritmo DPSO, é proposto como algoritmo evolucionário para resolver o problema de otimização de alocação do setor RRH, buscando maximizar os fatores de QoS do sistema, através do balanceamento de carga. Para a quantificação do QoS, são levados em consideração o KPI mostrado anteriormente. É auto-otimizado, executado pelo Host Manager para identificar a configuração adequada do RRH na BBU, com base nas informações coletadas de todos os RRHs na rede.

O DPSO começa com a geração aleatória de partículas, com uma distribuição linear, onde cada partícula é uma solução possível para o problema e representa o conjunto de RRHs, e seus respectivos setores, nos quais eles estão conectados. Na sequência, cada partícula é avaliada pela função objetivo, na qual retorna à qualidade da solução. Assim, os valores das partículas são atualizados, buscando o melhor *Gbest*, depois disso, ele está em um *loop* até atingir os critérios de parada.

O critério de parada do DPSO é baseado na variação do valor da função objetivo após um número sucessivo de iterações ou até atingir uma quantidade máxima de iterações previamente estabelecida. A convergência do processo é verificada pelo atendimento de todas as restrições impostas e se os valores discretos respeitam certa tolerância, para essa dissertação foram definidos 100 iterações como critério de parada, onde a cada iteração atualiza suas partículas e calcula sua adequação.

Quando a solução atual é melhor que a anterior, os valores de partículas são atualizados (*PBest*), se não, a solução é descartada e a solução anterior é mantida. Para cada iteração,

verifica-se se a melhor solução encontrada permanece a mesma, caso contrário, ela é atualizada pelos melhores (GBest) encontrados.

A tomada de decisão é explicada na figura 16, onde as análises de KPIs e QoS (função aptidão) são pré-determinadas. Uma das regras de restrição na alocação é que uma RRH não pode ter mais usuários conectados nela do que a capacidade máxima de um setor de BBU (HC). Ou seja, ela não pode estar ocupando sozinha um setor e ter usuários bloqueados. Nesse caso, todas as RRHs são selecionadas para escolher o usuário que será atendido em sua demanda. Normalmente, o DPSO é isento de função de penalidade para discretização, as posições das partículas são arredondadas para o valor inteiro mais próximo a cada iteração e as regiões infactíveis são evitadas.

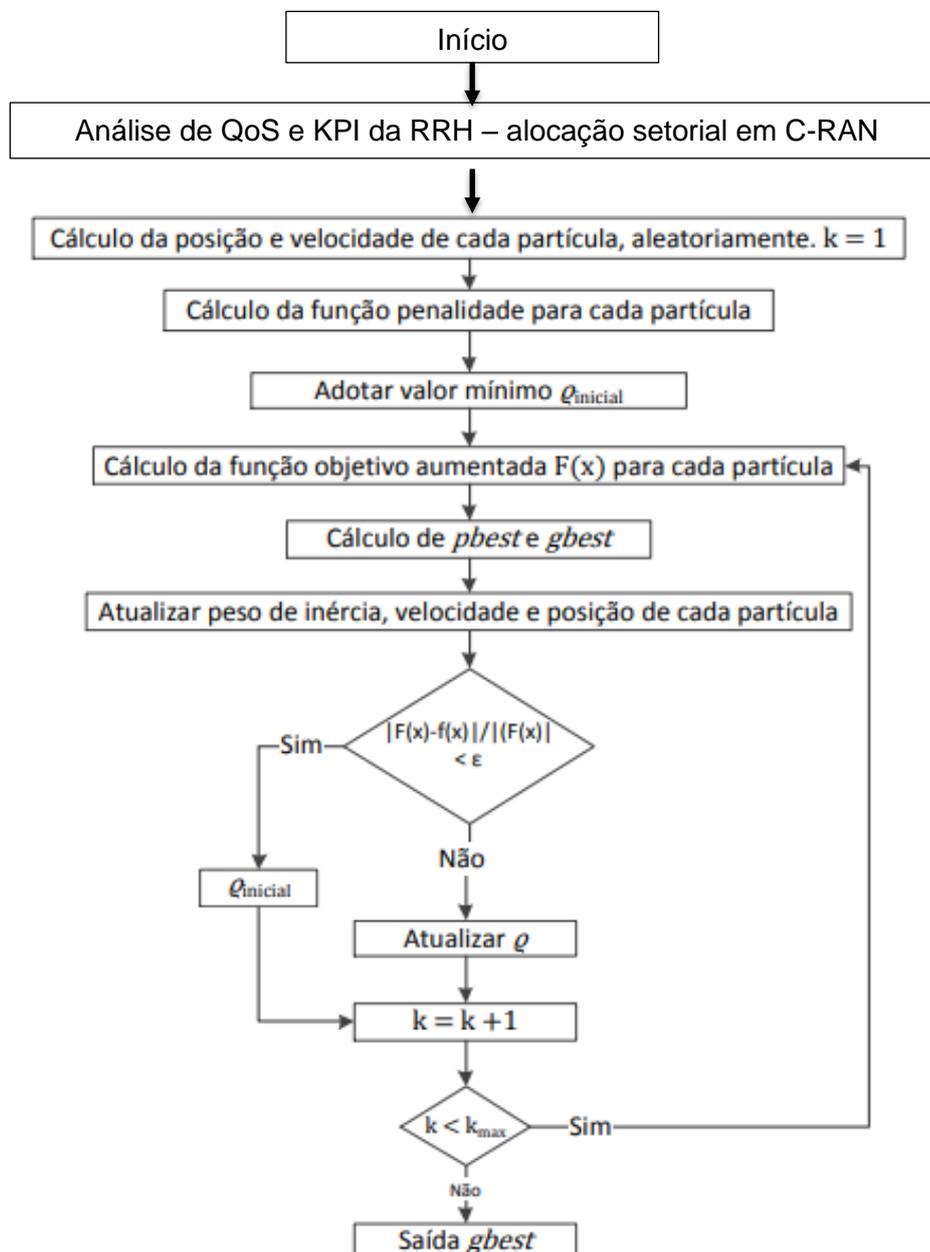


Figura 16 – Fluxograma da implementação do DPSO usando função objetivo.

Porém, quando o DPSO é executado, a organização da rede resulta ainda com usuários bloqueados, o sistema então seleciona os setores das BBUs que tiveram seu HC superado, escolhendo, dentre as antenas presentes nesses setores, tal usuário. Os parâmetros, do DPSO, aplicados na simulação são mostrados na Tabela 2.

Tabela 2 – Parâmetros do DPSO

Parâmetros	Valores
Número de partículas	30
Número de interações como Critério de Parada	100
Constante de aceleração Global	1.8
Constante de aceleração Pessoal	1.8
C1	2
C2	2

5.5 Considerações finais do capítulo

Neste capítulo, foram apresentados a formulação do problema, bem como a otimização utilizada nesta dissertação. A técnica otimização discreta por nuvem de partículas e modelagem foram descritas, visando um melhor esclarecimento acerca das formulações. Com base nos KPIs escolhidos, um algoritmo decide a melhor configuração do sistema e, por fim, a nova topologia (configuração BBU-RRH) é aplicada na fase de execução. No capítulo posterior, são apresentados os resultados obtidos, e a discussão dos mesmos.

6. ANÁLISE DOS RESULTADOS

Nesse capítulo são apresentados os resultados pertinentes ao modelo de alocação de recursos visando a otimização do balanceamento de carga nas BBUs e respectivas análises em sistemas testes usando o algoritmo DPSO, objetivando explicar que a proposta desta dissertação se mostrou eficaz. O desempenho da abordagem proposta é avaliado comparando-se os resultados obtidos com os resultados publicados na literatura

6.1 Resultados

Uma característica importante para a análise estatística correta dos resultados da simulação são os intervalos de confiança. Neste trabalho, foi adotado o intervalo de confiança de 95%, no qual todos os cenários foram simulados 30 vezes, como e os resultados apresentados são referentes à média destas.

Para os 6 cenários gerados para coleta de resultados, contendo a distribuição de usuários dentro de cada célula é uniforme onde 300 e 1800 usuários são considerados para células não densas e de alta densidade, respectivamente. Um quinto das células suportadas por RRH são escolhidos aleatoriamente para ter usuários de alta densidade.

Para as medidas de desempenho, foram adotadas o número de conexões por tipo de RAN (nas RRHs), a média de número de usuários conectados, desconectados, e a Taxa de Bits Máxima na rede. E para a extração dos parâmetros de caracterização tráfego, foi considerado que cada usuário gerou um tráfego de 400 Kbps.

O resultado obtido para a simulação do DPSO com 300 usuários, pode ser observado na Tabela 3, assim como a média de usuários alocados em cada setor após as 30 repetições. Além disso, é apresentado o desvio padrão desta carga. Por se tratar de um problema de poucas dimensões, já se esperava que os resultados fossem muito próximos do ótimo global. Sendo assim todos apresentam a função objetivo proposta próximo a 100%, apenas como objetivo demonstrar a convergência e funcionamento dos algoritmos.

Tabela 3 – Número médio de usuários (300 usuários)

Algoritmo	Média	Desvio padrão
BBU Eq 11	707,80	4,27
BBU Eq 12	708,70	4,02
BBU Eq 13	699,52	6,98
BBU (Proposta)	702,70	4,18

As performances para a simulação do DPSO para 600 usuários, pode ser observada na Tabela 4. Nesse caso, o modelo proposto obteve vantagens sobre os demais métodos, porém, pode-se observar que a maior *média* encontrada pelo algoritmo BBU Eq 11 é maior que a média do modelo proposto, mostrando que algumas melhorias no seu desenvolvimento podem impactar no seu desempenho, tornando-o equiparável ao modelo proposto.

Tabela 4 – Número médio de usuários (600 usuários)

Algoritmo	Média	Desvio padrão
BBU Eq 11	9409,28	164,95
BBU Eq 12	9347,82	161,11
BBU Eq 13	8837,36	287,89
BBU (Proposta)	9897,30	215,32

O resultado obtido para a simulação do DPSO para 900 usuários, é apresentado na Tabela 5. Aqui o modelo proposto, que ainda se mantém com um bom resultado, entretanto um detalhe bem visível é a dispersão do BBU Eq 13, que apresentou o valor acima dos demais.

Tabela 5 – Número médio de usuários (900 usuários)

Algoritmo	Média	Desvio padrão
BBU Eq 11	6989,54	212,39
BBU Eq 12	6990,14	147,21
BBU Eq 13	7041,26	131,31
BBU (Proposta)	6895,98	192,63

O resultado obtido para a simulação do DPSO, para 1200 usuários, é apresentado na Tabela 6. Pode-se ver claramente que todos os algoritmos obtiveram ótimos resultados, mostrando que com o crescer do número de usuários, tendem a apresentar melhores resultados, mesmo que ainda atrás do modelo proposto, que se mantém o melhor.

Tabela 6 – Número médio de usuários (1200 usuários)

Algoritmo	Média	Desvio padrão
BBU Eq 11	7504,86	67,27
BBU Eq 12	7519,54	75,96
BBU Eq 13	7264,78	126,96
BBU (Proposta)	7631,28	104,16

O resultado obtido para a simulação do DPSO, para 1500 usuários, é apresentado na Tabela 7. Aqui os modelos BBU Eq 11, BBU Eq 12 e BBU Eq 13 alcançaram desempenhos próximos.

Tabela 7 – Número médio de usuários (1500 usuários)

Algoritmo	Média	Desvio padrão
BBU Eq 11	39695,02	421,54
BBU Eq 12	39438,10	387,14
BBU Eq 13	37672,40	713,40
BBU (Proposta)	40941,76	608,61

O resultado obtido para a simulação do DPSO, para 1800 usuários, é apresentado na Tabela 8.

Tabela 8 – Número médio de usuários (1800 usuários)

Algoritmo	Média	Desvio padrão
BBU Eq 11	36500,90	1033,28
BBU Eq 12	36658,50	914,28
BBU Eq 13	39719,04	514,29
BBU (Proposta)	39843,28	770,74

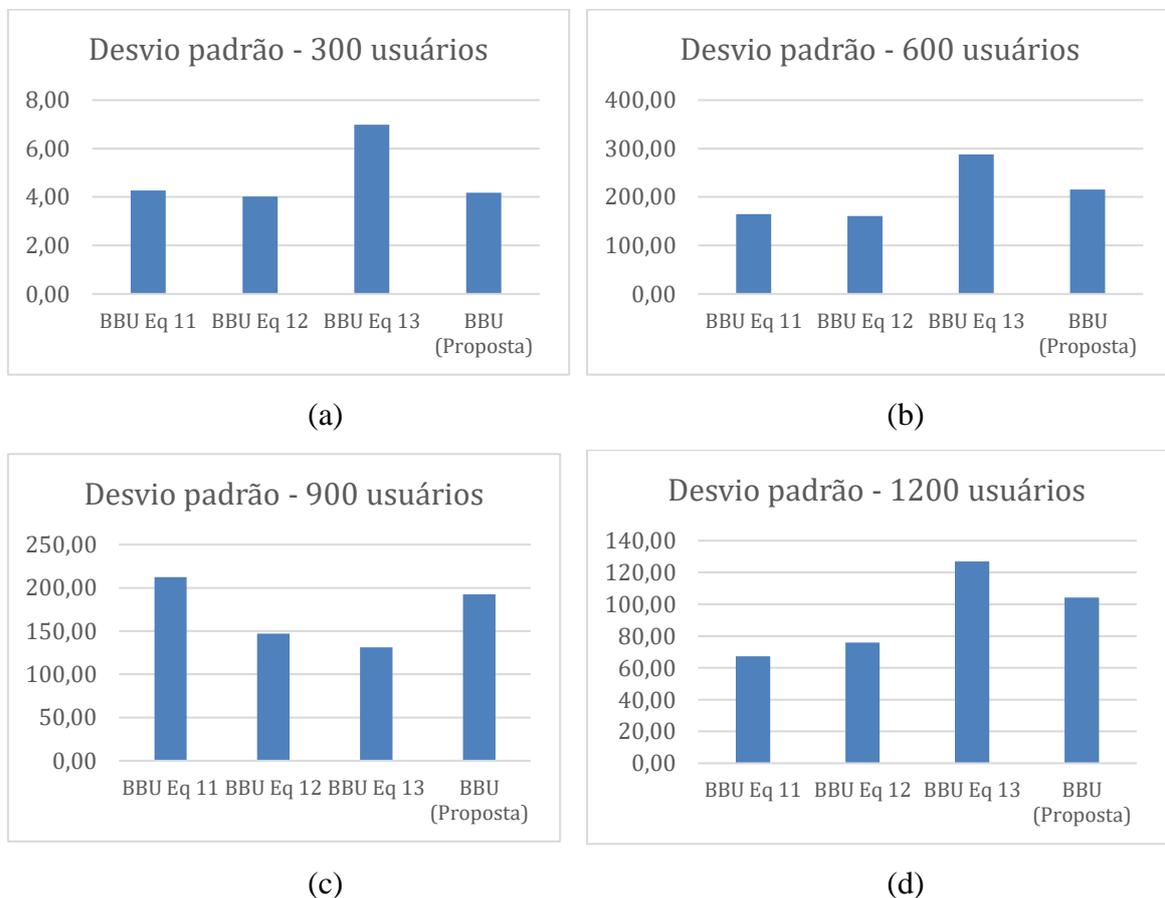
Por fim, a maior dimensão testada apresentou resultados peculiares. O modelo proposto manteve sua liderança, mostrando-se o modelo mais eficaz dentre os avaliados, observando-se as experimentações dos artigos que foram utilizados como referência para este trabalho. Contudo, mesmo os algoritmos apresentando resultados muito próximos ao modelo proposto, ainda assim a necessidade de serem reavaliados, podendo esses resultados serem frutos de um possível vício na série de seleções. Além disso, se faz necessário também testes em casos reais, para que seja comprovada a funcionalidade vista nos testes.

Os resultados gerais mostram que, além do modelo aqui apresentado, a proposta considerada clássica – BBU Eq 11 - foi também a que alcançou melhores desempenhos gerais para as tarefas propostas. Todos os modelos alcançaram desempenhos geralmente próximos, mostrando que, mesmo em diferentes cenários e metodologia, os métodos de enxame conseguiram, no caso geral, chegar a resultados muito próximos do ótimo, porém em nenhum dos cenários, o modelo proposto por essa dissertação, foi superado, quando comparado aos seus benchmarks. Em todos os casos, o modelo proposto nesta dissertação, consegue, através de sua melhor distribuição dos usuários, possibilitando, assim, que mais usuários sejam atendidos e,

além disso, não gerando estresses desnecessários no hardware destes equipamentos e, com isso, aumentando seu tempo de vida e otimizando o mapeamento.

O DPSO é utilizado por ser capaz de resultar em maiores níveis de QoS com menos iterações como é comprovado em [KHAN et al., 2015], [DA PAIXÃO et al., 2018] e [KHAN et al., 2018]. O fato de convergir mais rápido e alcançar resultados ótimos é um fator extremamente importante em redes que se propõe em oferecer serviços de alta qualidade e em tempo real. Nesse caso, pode-se observar que os valores ótimos mostrados na figura 9[a], é alcançado pela busca exaustiva de todas as soluções possíveis, o que ajuda a demonstrar a melhoria em cada iteração do algoritmo DPSO.

Assim, a fim de comparar os benchmarks com a função objetivo proposta por essa dissertação, em equidade, é utilizado o desvio padrão da equação 14, para medir o quão balanceadas as BBU's ficam. Através da Figura 17, é possível avaliar os resultados para cada cenário de coleta de resultados, aqui proposto. Observa-se que a proposta apresentada neste trabalho entrega resultados satisfatórios e consegue atender a todas as restrições, mesmo tendo que lidar com cenários mais robustos. O balanceamento de setores obtido retorna um mapeamento aceitável perante as FOs comparadas.



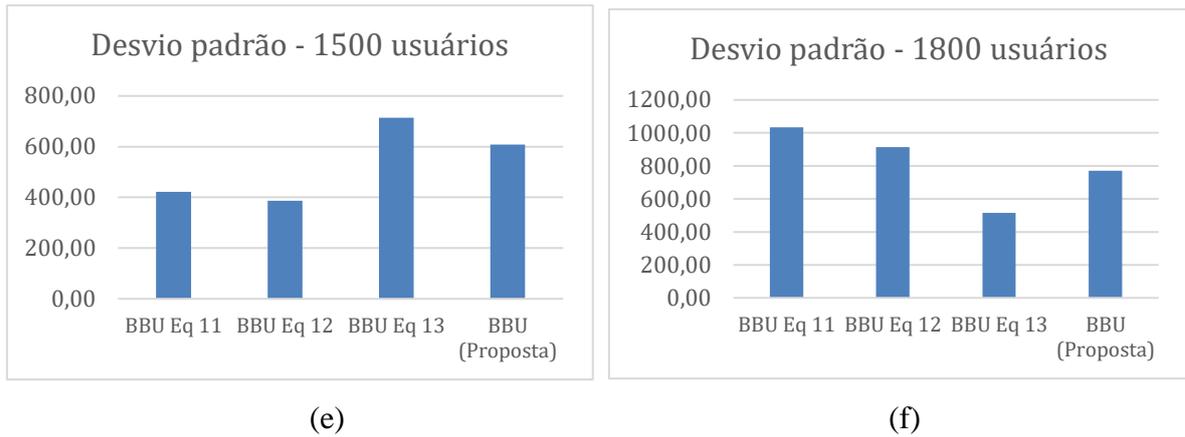


Figura 17. Nível de Balanceamento - Medida de Dispersão

Pode-se observar que a função objetivo proposta consegue balancear os setores e respeitar a restrição do limite de capacidade imposto. Os resultados apontaram para o desempenho próximo do ótimo para o modelo apresentado nesta dissertação, tanto em cenários de alta, como em baixa densidade de tráfego.

6.2 Considerações finais do capítulo

Neste capítulo foi mostrado os resultados do desempenho do DPSO, em 6 cenários, variando o número de usuários na rede, e comparado com FOs de referência, usadas como *benchmark*, com a proposta neste trabalho, foi possível verificar melhor desempenho para essa função. No capítulo seguinte serão feitas as considerações finais a respeito dessa dissertação.

7 CONCLUSÃO

Para atender a crescente demanda de dispositivos móveis conectados as redes sem fio de longa distância e o rápido crescimento no número de aplicações e serviços, iniciativas científicas e indústria voltaram esforços para o desenvolvimento de uma nova geração de redes móveis, a rede 5G. O que está forçando as operadoras de redes a investirem em infraestruturas para aumentar a capacidade de atendimento, gerando mais tráfego e requisitos de QoS mais restritos, logo, tecnologias que se propõem atender os usuários conseguindo, reduzindo os custos operacionais são muito atrativas, como a C-RAN, todavia, trazem novos desafios para que ofereçam mais benefícios.

A arquitetura C-RAN tem por escopo viabilizar e atender as especificações da nova geração de redes móveis, permitir que os operadores respondam as demandas através do uso de técnicas de virtualização e recursos de processamento de dados centralizado. Considerando os inúmeros desafios que precisam ser enfrentados em relação a orquestração de recursos, as métricas para tomada de decisão de ações e a distribuição.

Neste trabalho foi apresentado um modelo para o mapeamento BBU-RRH em C-RAN para as redes 5G, busca-se equilibrar a carga da rede nos recursos de processamento através do algoritmo para otimização por enxame de partículas, DPSO, visando a otimização do balanceamento da carga nas BBUs, com o objetivo de garantir que a eficiência energética e o QoS sejam maximizados. Considerando que o gerenciamento de redes está associado ao controle das atividades e ao monitoramento de recursos e um planejamento de operação é de extrema importância.

Foram apresentados cenários com variações no número de usuários. Como forma de comparação, e utilizadas três FOs, encontradas na literatura, como benchmarks, buscando equilibrar a carga da rede nas BBUs e deixando o uso dos recursos equilibrados.

Os resultados computacionais deixaram evidente que o modelo proposto se mostrou superior aos demais métodos e ao mesmo tempo, pode-se dizer que os algoritmos de inteligência de enxame, para o caso de mapeamento BBU-RRH e balanceamento das BBUs, se mostraram bastante eficientes, entretanto faz-se necessário avaliar os métodos em problemas reais de mapeamento e balanceamento de cargas em redes centralizadas.

Resumidamente, é possível obter informações precisas através dos dados e aplicá-las as soluções, tendo assim, uma resposta eficiente, afetando de maneira benéfica a relação custo-benefício da rede. A função proposta obteve melhores resultados no balanceamento da rede quando se tem alta variabilidade no tráfego comparado as FOs *benchmark*. O modelo proposto

se mostrou, dessa forma, eficaz para o mapeamento BBU-RRH, trazendo benefícios agregados a rede.

7.1 Contribuições

As principais contribuições do trabalho podem ser resumidas da seguinte forma:

- O modelo proposto nesta dissertação, conseguiu, através de sua melhor distribuição dos usuários, possibilitar, assim, o melhor desempenho da rede, atingindo o seu principal objetivo, que mais usuários sejam atendidos, através do mapeamento BB-RRU e o balanceamento de carga na rede, servindo dessa forma como uma solução para melhorar o desempenho da C-RAN.
- Levantamento bibliográfico minucioso acerca dos temas envolvidos nesta proposta de dissertação, tais como 5G, mapeamento BBU-RRH, PSO e avaliação de desempenho.
- As elaborações de relatórios técnicos para auxiliar o desenvolvimento de pesquisas de iniciação científica;

Além disto, a divulgação do trabalho por meio de publicação de artigos em conferências Nacionais, onde são apresentadas a proposta e os resultados do estudo de caso desta dissertação.

Os artigos são listados abaixo:

- CUNHA, R., VIEIRA, R., PEREIRA, P., FALCÃO, I., SOUZA, D., & CARDOSO, D. [2019]. Meta-heurística para Alocação de Recursos Aplicada a Redes de Acesso Centralizado. In *Anais do XLVI Seminário Integrado de Software e Hardware*, [pp. 149-160]. Porto Alegre: SBC. doi:10.5753/semish.2019.6575
- CUNHA, R. de C. P., Vieira, R. F., Pereira, P. H. A., Falcão, I. W. S., Souza, D. da S., & Cardoso, D. L. [2019]. Alocação de recursos de Hardware em arquitetura C-RAN utilizando DPSO / Hardware resource Allocation in C-RAN Architecture Using DPSO. *Brazilian Journal of Development*, 5[12], 29940–29955. <https://doi.org/10.34117/bjdv5n12-132>

7.2 Trabalhos futuros

Sob a perspectiva acadêmica alguns pontos, destacados a seguir, são vislumbrados como possibilidades de trabalhos futuros:

- 1) A realização de uma simulação e o aumento da complexidade do modelo são previstos.

- 2) A realização de experimentos em ambientes reais para que sejam analisados os diversos recursos avaliados nesse trabalho, buscando assim ratificar as informações fornecidas.
- 3) Embora o processo de implementação de redes 5G ainda esteja em fase inicial, pesquisas sobre a geração de redes móveis que sucederá o 5G, denominada por 6G já estão em pauta, e são interessantes para futuras discussões e estudos.
- 4) Explorar a utilização de virtualização com base em containers ao invés da utilização das tradicionais máquinas virtuais.

8 REFERÊNCIAS

- AHMAD, I. et al. Towards software defined cognitive networking. In: 2015 7th International Conference on New Technologies, Mobility and Security (NTMS). [S.l.: s.n.], 2015. p. 1–5. ISSN 2157-4952. 15
- AKYILDIZ, I. F.; WANG, P.; LIN, S.-C. Softair: A software defined networking architecture for 5g wireless systems. *Computer Networks*, v. 85, p. 1 – 18, 2015. ISSN 1389-1286. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1389128615001632>. 15
- AL-FALAHY, N. and Alani, O. Y. (2017). Technologies for 5G Networks: Challenges and Opportunities. *IT Professional*, 19(1):12–20.
- BABA, Majid Irfan et al. Evolution of Mobile Wireless Communication Systems from 1G to 5G: A Comparative Analysis. 2018.
- BECCENERI, José & Ramos, Fernando & Campos Velho, Haroldo & Silva, José & Lorena, Luiz & Vijaykumar, Nandamudi & Santos, Rafael & Rosa, Reinaldo & Travelho, Jerônimo. (2008). *Meta-Heurísticas e Otimização Combinatória: Aplicações em Problemas Ambientais*.
- BENZEKKI, K.; FERGOUGUI, A. E.; ELALAOUI, A. E. Software-defined networking (sdn): a survey. *Security and communication networks*, Wiley Online Library, v. 9, n. 18, p. 5803–5833, 2016. 14, 15
- BERGH, F. V. den; ENGELBRECHT, A. Effects of swarm size on cooperative particle swarm optimisers. *Proceedings of the Genetic and Evolutionary Computation Conference*, p. 892–899, 2001. Citado na página 46.
- Bhaumik, S., Chandrabose, S. P., Jataprolu, M. K., Kumar, G., Muralidhar, A., Polakos, P., Srinivasan, V., and Woo, T. (2012). CloudIQ: A Framework for Processing Base Stations in a Data Center. *Proceedings of the 18th annual international conference on Mobile computing and networking*, pages 125–136.
- BILALB, Sardar M. et al. A performance comparison of network simulators for wireless networks. arXiv preprint arXiv:1307.4129, 2013.

BOULOS, K., El Helou, M., and Lahoud, S. (2015). RRH Clustering in Cloud Radio Access Networks. In Applied Research in Computer Science and Engineering (ICAR), 2015 International Conference on, pages 1–6. IEEE.

CHANCLOU, P. et al. Optical fiber solution for mobile fronthaul to achieve cloud radio access network. In: FUTURE NETWORK AND MOBILE SUMMIT (FUTURENETWORKSUMMIT), 2013, 2013. Anais. . . 2013. p. 1–11.

CHAVES, L. J.; GARCIA, I. C.; MADEIRA, E. R. M. Openflow-based mechanisms for qos in lte backhaul networks. In: 2016 IEEE Symposium on Computers and Communication (ISCC). [S.l.: s.n.], 2016. p. 1233–1238. 13, 43

CHECKO, A. et al. Cloud ran for mobile networks—a technology overview. IEEE Communications Surveys Tutorials, v. 17, n. 1, 2015. p. 405–426, 2015. 6, 7, 8, 10, 11.

CHEN, Y.-S., Chiang, W.-L., and Shih, M.-C. (2018). A Dynamic BBU-RRH Mapping Scheme using Borrow-and-lend approach in Cloud Radio Access Networks. IEEE Systems Journal, 12(2):1632–1643.

CHEN, T. et al. Software defined mobile networks: concept, survey, and research directions. IEEE Communications Magazine, v. 53, n. 11, p. 126–133, November 2015. ISSN 0163-6804. 13

CHINA, M. C-ran: the road towards green ran. White paper, v. 2, p. 1–10, 2011. 1, 3, 6, 7, 9, 10, 11, 12.

CISCO, C. Global mobile data traffic forecast update, 2017–2022. *white paper*, 2019. 1

CISCO, S. Cisco visual networking index: Global mobile data traffic forecast update, 2017–2022. Cisco Public Information: San Jose, CA, USA, 2017.

CMRI. C-ran: the road towards green ran. C-RAN International Workshop, Beijing, apr 2010.

DAHLMAN, E.; PARKVALL, S.; SKOLD, J. 5G NR: The next generation wireless access technology. [S.l.]: Academic Press, 2018.

- DA PAIXÃO, E. A. R., Vieira, R. F., Araújo, W. V., and Cardoso, D. L. (2018). Optimized Load Balancing by Dynamic BBU-RRH Mapping in C-RAN Architecture. In Fog and Mobile Edge Computing (FMEC), 2018 Third International Conference on, pages 100–104. IEEE.
- DE CASTRO, L. N.; VON ZUBEN, F. J. Recent Developments in Biologically inspired Computing. Igi Global, 2005.
- DUAN, T. et al. Inter-bbu control mechanism for load balancing in c-ran-based bbu pool. Computer and Communications (ICCC), Chengdu, may 2017.
- DUCH, W. What Is Computational Intelligence and What Could It Become? In: DUCH, W; MANDZIUK, J. (Eds.) Challenges for Computational Intelligence. Springer, 2007. Disponível em: <cogprints.org/5358/01/06-CIdef.pdf>. Acesso em 14 Jun. 2022.
- EBERHART, R. C. and KENNEDY, J. A new optimizer using particle swarm theory. In Proceedings of the sixth international symposium on micro machine and human science, volume 1, pages 39–43. New York, NY. 1995
- EBERHART R. C.; SHI Y. Particle Swarm Optimization: Developments, Applications and Resources. Congress on Evolutionary Computation, 2001, vol. 1, pp. 81-86, 2001.
- EIBEN, A. E.; SMITH, J. E. Introduction to evolutionary computing. [S.l.]: Springer, 2003. v.53.
- ENGELBRECHT, A. P. Computational Intelligence: An Introduction. 2. ed. [S.l.]: John Wiley and Sons, 2007. Citado 12 vezes nas páginas 40, 42, 43, 44, 46, 47, 49, 50, 51, 52, 55 e 68.
- ERISSON. "Ericsson mobility report 2018."white paper, 2018.
- FOULDS, L. R. Combinatorial Optimization for Undergraduates. Springer-Verlag, New York, 1984, p. 114.
- GONÇALVES, M. D. P. D. S. e. a. (2018). Abordagem de Offloading para a alocação de dois níveis em uma Arquitetura H-CRAN. VIII Conferência Nacional em Comunicações, Redes e Segurança da Informação, pages 79–80.
- GUPTA, A. and Jha, R. K. (2015). A Survey of 5G Network: Architecture and Emerging Technologies. IEEE access, 3:1206–1232.

HOLLAND, J. H. Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control, and artificial intelligence. [S.l.]: U Michigan Press, 1975. Citado na página 25.

HOSSAIN, M. F. et al. Recent research in cloud radio access network (c-ran) for 5g cellular systems - a survey. *Journal of Network and Computer Applications*, v. 139, p. 31 – 48, 2019. ISSN 1084-8045. Disponível em: <[http://www.sciencedirect.com/science- /article/pii/S1084804519301432](http://www.sciencedirect.com/science/article/pii/S1084804519301432)>. 7, 10, 11.

KARAKUS, M.; DURRESI, A. Quality of service (qos) in software defined networking (sdn): A survey. *Journal of Network and Computer Applications*, v. 80, p. 200 – 218, 2017. ISSN 1084-8045. Disponível em: . 14, 15, 16, 19.

KARAKUS, M.; DURRESI, A. A survey: Control plane scalability issues and approaches in software-defined networking (sdn). *Computer Networks*, v. 112, p. 279 – 293, 2017. ISSN 1389-1286. Disponível em: . 13, 14, 15, 16.

KENNEDY, J.; EBERHART, R. Particle swarm optimization. In: *Proceedings IEEE International Conference on Neural Networks*. [S.l.: s.n.], 1995. p. 1942–1948. Citado 3 vezes nas páginas 22, 24 e 37.

KENNEDY, J.; EBERHART, R. C. A discrete binary version of the particle swarm algorithm. In: *IEEE International Conference on Systems, Man, and Cybernetics. Computational Cybernetics and Simulation*. [S.l.: s.n.], 1997. p. 4104–4108. Citado na página 24.

KENNEDY, J. Particle swarm optimization. In: *Encyclopedia of machine learning*. [S.l.]: Springer, 2011. p.760–766.

KHAN, M., Fakhri, Z. H., and Al-Raweshidy, H. S. (2018). Semistatic Cell Differentiation and Integration with Dynamic BBU-RRH Mapping in Cloud Radio Access Network. *IEEE Transactions on Network and Service Management*, 15(1):289–303.

KHAN, M., Alhumaima, R. S., and Al-Raweshidy, H. S. (2017). Qos-Aware Dynamic RRH Allocation in a Self-Optimized Cloud Radio Access Network with RRH proximity Constraint. *IEEE Transactions on Network and Service Management*, 14(3):730–744.

KHAN, M.; ALHUMAIMA, R. S.; AL-RAWESHIDY, H. S. Quality of service aware dynamic bbu-rrh mapping in cloud radio access network. In: 2015 International Conference on Emerging Technologies (ICET). [S.l.: s.n.], 2015. p. 1–5. 1, 3, 30, 32, 37.

KITINDI, E. J. et al. Wireless network virtualization with sdn and c-ran for 5g networks: Requirements, opportunities, and challenges. *IEEE Access*, v. 5, p.19099–19115, 2017.

KOTECHA, K.; SANGHANI, G.; GAMBHAVA, N. Genetic algorithm for airline crew scheduling problem using cost-based uniform crossover. In: Applied Computing. [S.l.]: Springer, 2014. p. 84–91. Citado na página 25.

LACERDA, E. G. M, Carvalho, A. C. P. L. (1999). Introdução aos algoritmos genéticos. In: Sistemas inteligentes: aplicações a recursos hídricos e ciências ambientais [Edited by Galvão, C. O., Valença, M. J. S.], Porto Alegre, Ed. Universidade/UFRGS: Associação Brasileira de Recursos Hídricos, 99-150.

LIU, M.; HAGHANI, A.; TOOBAIE, S. Genetic algorithm-based column generation approach to passenger rail crew scheduling. *Transportation Research Record: Journal of the Transportation Research Board*, Trans Res Board, v. 2159, n. 1, p. 36–43, 2010. Citado na página 25.

MAROTTA, M. A. et al. Resource sharing in heterogeneous cloud radio access networks. *IEEE Wireless Communications*, v. 22, n. 3, p. 74–82, 2015.

MAURICE, C. Discrete particle swarm optimization, illustrated by the traveling salesman problem. In: . *New Optimization Techniques in Engineering*. [S.l.]: Springer Berlin Heidelberg, 2004. p. 219–239, 24.

HIGHAM DJ, Higham NJ. *MATLAB guide*. Vol. 150. Siam; 2016.

JIBRIL, Mustefa; HASSEN, Nurys; TADESSE, Mesay. Load Balancing in 5G C-RAN using Swarm Intelligence (SI) Algorithms. *International Journal of Communications*, v. 6, 2021.

LIU, Qiang & Wang, Chengen. (2011). A discrete particle swarm optimization algorithm for rectilinear branch pipe routing. *Assembly Automation*. 31. 363-368. 10.1108/01445151111172952.

- MOUAWAD, Mostafa; MAH, Firmin; DZIONG, Zbigniew. RRH-Sector selection and load balancing based on MDP and dynamic RRH-Sector-BBU mapping in C-RAN. *Computer Networks*, v. 215, p. 109192, 2022.
- NAMAL, S. et al. Implementation of openflow based cognitive radio network architecture: Sdn&r. *Wireless Networks*, v. 22, n. 2, p. 663–677, Feb 2016. ISSN 1572-8196.
- NGUYEN, V.-G.; DO, T.-X.; KIM, Y. Sdn and virtualization-based lte mobile network architectures: A comprehensive survey. *Wireless Personal Communications*, Springer, v. 86, n. 3, p. 1401–1438, 2016. 12, 14
- NIKAEIN, N. et al. Demo – closer to cloud-ran: ran as a service. *ACM SIGCOMM Computer Communication Review*, Paris, p. 193–195, sep 2015.
- PAN, C., Elkashlan, M., Wang, J., Yuan, J., and Hanzo, L. (2018). User-Centric C-RAN Architecture for Ultra-Dense 5G Networks: Challenges and Methodologies. *IEEE Communications Magazine*, 56(6):14–20.
- PIOTROWSKI, Adam P.; NAPIORKOWSKI, Jaroslaw J.; PIOTROWSKA, Agnieszka E. Population size particle swarm optimization. *Swarm and Evolutionary Computation*, v. 58, p. 100718, 2020.
- POMPILI, D.; HAJISAMI, A.; VISWANATHAN, H. Dynamic provisioning and allocation in cloud radio access networks (c-rans). *Ad Hoc Networks*, v. 30, 03 2015. 2, 7, 8, 9, 10, 31, 32.
- RANA, D. S.; DHONDIYAL, S. A.; CHAMOLI, S. K. Software Defined Networking (SDN) Challenges. *issues and Solution*, 2019.
- RAO, Voore & Srinivas, K.. (2022). Optimization of BBU-RRH Mapping for Load Balancing in 5G C-RAN Using Swarm Intelligence (SI) Algorithms. 10.1007/978-981-19-0105-8_35.
- RAZA, M. R. et al. Demonstration of dynamic resource sharing benefits in an optical c-ran. *IEEE/OSA Journal of Optical Communications and Networking*, v. 8, n. 8, aug 2016.
- SALMAN, T. Cloud ran: Basics, advances, and challenges, a survey of c-ran basics, visualization. *Resource Allocation, and Challenges*, 2016. 8, 9, 12.

SDN vs. NFV: What's the difference?. Disponível em: <https://www.cisco.com/c/en/us/solutions/software-defined-networking/sdn-vs-nfv.html>. Acesso em: 14 nov. 2022.

SERAPIAO, A. B. S. Fundamentos de otimização por inteligência de enxames: uma visão geral. Sba: Controle & Automação - Sociedade Brasileira de Automática, Scielo, v. 20, p. 271 – 304, set. 2009. ISSN 0103-1759.

SIGWELE, T.; PILLAI, P.; HU, Y. F. itree: intelligent traffic and resource elastic energy scheme for cloud-ran. Future Internet of Things and Cloud (FiCloud), Rome, oct 2015

SILVA, Ketyllen da Costa. Estratégia de planejamento e otimização do handover em redes móveis densificadas. 88 f. Tese (Doutorado) - Universidade Federal do Pará, Instituto de Tecnologia, Belém, 2018. Programa de Pós-Graduação em Engenharia Elétrica. Disponível em:<<http://repositorio.ufpa.br/jspui/handle/2011/10249>>. Acesso em: Julho de 2022.

SUNDARESAN, K. et al. Fluidnet: A flexible cloud-based radio access network for small cells. IEEE/ACM Transactions on Networking, v. 24, n. 2, p. 915–928, 2016. 7.

STRASSER, S. et al. A new discrete particle swarm optimization algorithm. In: Proceedings of the Genetic and Evolutionary Computation Conference 2016. [S.l.: s.n.], 2016. (GECCO '16), p. 53–60, 24.

VANECK, N.J. *ET AL*. Visualizing the Computational Intelligence Field. IEEE Computational Intelligence Magazine, p. 6-10, nov. 2006.

WANG, S.-C. Artificial neural network. In: Interdisciplinary computing in java programming. [S.l.]: Springer, 2003. p.81–100.

WANG, R.; HU, H.; YANG, X. Potentials and challenges of c-ran supporting multi-rats toward 5g mobile networks. IEEE Access, v. 2, p. 1187–1195, 2014. ISSN 2169-3536. 7, 8.

WEI, L. et al. Key Elements to Enable Millimeter Wave Communications for 5G Wireless Systems. IEEE Wireless Communications, IEEE, v. 21, n. 6, p. 136–143, 2014.

WU, J. et al. Cloud radio access network (c-ran): a primer. *IEEE Network*, v. 29, n. 1, p. 35–41, Jan 2015. ISSN 0890-8044. 3, 7.

YAN, B., Zhao, Y., Yu, X., Wang, W., Wu, Y., Wang, Y., and Zhang, J. (2018). Tidaltraffic-aware Routing and Spectrum Allocation in Elastic Optical Networks. *Journal of Optical Communications and Networking*, 10(11):832–842.

YASEEN, F. A.; AL-KHALIDI, N. A.; AL-RAWESHIDY, H. S. Smart virtual enb (svenb) for 5g mobile communication. *Fog and Mobile Edge Computing (FMEC)*, Valencia, jun 2017.

YAO, J. and Ansari, N. (2018). Qos-aware Joint BBU-RRH Mapping and User Association in Cloud-RANs. *IEEE Transactions on Green Communications and Networking*.

ZHANG, Y. et al. Dynamic allocation of processing resources in cloud-ran for a virtualised 5g mobile network. In: *IEEE. 2018 26th European Signal Processing Conference (EUSIPCO)*. [S.l.], 2018. p. 782–786. 3.

ZHANG, Y. *Network Function Virtualization Concepts and Applicability in 5G Networks*, John Wiley & Sons, Inc., 2018, pp.37–50.

Considerações Prof Marcelino

1. Introdução
 - Objetivos colocados de forma mais simples e direta, conforme observado.
 - Organização do texto reformulada.
2. Capítulo 2, onde continha fundamentação teórica, foi dividido em dois tópicos, conforme sugestão:
 - Capítulo 2: Redes móveis
 - Capítulo 3: Inteligência computacional
3. Cenário
Inserido na página 42.
4. Fundamentação Teórica
 - O capítulo foi dividido em dois tópicos, entretanto todas as correções foram realizadas. Conforme orientação.
 - O item 4.1.3 foi retirado do capítulo 4 (agora capítulo 5), passando a ser o item
 - Pag. 4, 2.1, parágrafo 1: alterado o texto para melhorar a coesão, coerência e boa compreensão;
5. Colocar o cenário de forma mais clara levar para o capítulo 4, agora 5.
 - Foi totalmente atualizado
6. Apresentação do modelo
 - Capítulo 5
7. Como pode ser implementado o DPSO
 - Capítulo 3 mostra o funcionamento do algoritmo e capítulo 5 mostra o fluxograma de funcionamento
8. Contextualizar o PSO
 - Foi criado um capítulo exclusivo para o pso
9. Onde aplica o kpi
 - Mostrado no fluxograma, pagina 44
10. Funções objetivas – mostrar de forma mais clara a proposta
 - Capítulo 5 é mostrada
11. Modelagem toda explicar a forma proposta
 - Capítulo 5 descreve a modelagem utilizada no trabalho
12. Organizar e dar destaque ao que é novo e o a impotância
 - Objetivos e motivação refeitos

Considerações Prof Jorge Antonio

1. Resumo (pg. xv)
 - Todas as alterações realizadas e o texto submetido novamente a correção.
2. Lista de siglas atualizada
3. Referências atualizadas
4. Introdução: Atualização do tempo verbal e Link entre os parágrafos.
 - Texto totalmente reformulado no intuito de atualizar o tempo verbal, colocar os dados de forma mais atualizadas e criar uma melhor harmonia entre os parágrafos, dando uma continuidade e uma coesão no texto.
 - Motivação colocada de forma mais clara e objetiva.
5. Fundamentação Teórica
 - Todas as demais orientações de correções gramaticais realizadas no capítulo.
6. Capítulo 4, Trabalhos relacionados, inseridas novas referencias a partir da página 35.
7. Voltar a falar no 5g parâmetros como se aplica
 - Sempre mencionado agora, durante todo o texto
8. Resultados
 - Refeitos
9. Teste
 - Capítulo 6, resultados
10. Descrever melhor o cenário se é 5g
 - Capítulo 5
11. Resultados figura 9 colocar todos na mesma escala
 - Foi alterada
12. Conclusões: refazer todos os parágrafos
 - Refeita
13. Tentar colocar o que o trabalho deixou de conhecimento novo, se for aplicar ao 5g nos dois modelos.
 - Página 51
14. O que foi observado a execução do dpso
 - Método eficiente para o balanceamento de cargas
15. DPSO é dinâmico?
 - Sim, assim como o PSO
16. Colocar a relevancia do trabalho
 - Considerações finais
17. A quantidade de dado gerado.
 - Foram refeitas as amostras
18. Avaliação de desempenho
 - Capítulo 6
19. Avaliação e seleção do algoritmo
 - Capítulo 6