

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ  
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA  
E INFORMÁTICA INDUSTRIAL

ANA FLÁVIA DOS REIS

**EFICIÊNCIA ENERGÉTICA NO POSICIONAMENTO DE  
ESTAÇÕES RÁDIO-BASE MÓVEIS BASEADAS EM VANTS  
APLICANDO APRENDIZAGEM POR REFORÇO**

DISSERTAÇÃO

CURITIBA

2020

ANA FLÁVIA DOS REIS

**EFICIÊNCIA ENERGÉTICA NO POSICIONAMENTO DE  
ESTAÇÕES RÁDIO-BASE MÓVEIS BASEADAS EM VANTS  
APLICANDO APRENDIZAGEM POR REFORÇO**

Dissertação apresentada ao Programa de Pós-graduação em Engenharia Elétrica e Informática Industrial da Universidade Tecnológica Federal do Paraná como requisito parcial para obtenção do título de “Mestre em Engenharia Elétrica” – Área de Concentração: Telecomunicações e Redes.

Orientador: Prof. Glauber Gomes de Oliveira Brante

Coorientador: Prof. Richard Demo Souza

CURITIBA  
2020

### Dados Internacionais de Catalogação na Publicação

Reis, Ana Flávia dos

Eficiência energética no posicionamento de estações rádio-base móveis baseadas em VANTs aplicando aprendizagem por reforço [recurso eletrônico] / Ana Flávia dos Reis. -- 2020.

1 arquivo eletrônico (46 f.) : PDF ; 849 KB.

Modo de acesso: World Wide Web.

Texto em português com resumo em inglês.

Dissertação (Mestrado) - Universidade Tecnológica Federal do Paraná. Programa de Pós-graduação em Engenharia Elétrica e Informática Industrial. Área de Concentração: Telecomunicações e Redes. Linha de Pesquisa: Comunicações e Processamento de Dados, Curitiba, 2020.

Bibliografia: f. 43-46.

1. Engenharia elétrica - Dissertações. 2. Energia - Consumo. 3. Sistemas de telefonia celular. 4. Sistemas de informação geográfica móvel. 5. Drone. 6. Aprendizado do computador. 7. Aprendizagem por reforço. 8. Sistemas de comunicação sem fio. 9. Sistemas de transmissão de dados. 10. Processamento eletrônico de dados em tempo real. 11. Sistemas auto-organizáveis. 12. Desempenho - Avaliação. 13. Métodos de simulação. I. Brante, Glauber Gomes de Oliveira, orient. II. Souza, Richard Demo, coorient. III. Universidade Tecnológica Federal do Paraná. Programa de Pós-graduação em Engenharia Elétrica e Informática Industrial. IV. Título.

CDD: Ed. 23 -- 621.3

## TERMO DE APROVAÇÃO DE DISSERTAÇÃO

A Dissertação de Mestrado intitulada **“Eficiência Energética no Posicionamento de Estações Rádio-Base Móveis Baseadas em VANTs Aplicando Aprendizagem por Reforço”** defendida em sessão pública pelo(a) candidato(a) **Ana Flávia dos Reis**, no dia **21 de fevereiro de 2020**, foi julgada para a obtenção do título de Mestre em Ciências, Área de Concentração: **Telecomunicações e Redes**, Linha de Pesquisa: **Comunicações e Processamento de Dados**, e aprovada em sua forma final, pelo Programa de Pós-Graduação em Engenharia Elétrica e Informática Industrial - CPGEI.

### BANCA EXAMINADORA:

Prof. Dr. Glauber Gomes de Oliveira Brante – Presidente – UTFPR

Prof. Dr. Marcelo Eduardo Pellenz – PUC/PR

Prof. Dr. Jamil de Araújo Farhat – Banco do Brasil

A via original deste documento encontra-se arquivada na Secretaria do Programa, contendo a assinatura da Coordenação após a entrega da versão corrigida do trabalho.

Curitiba, 21 de fevereiro de 2020.

---

Carimbo e Assinatura do(a) Coordenador(a) do Programa

Aos meus pais, meus eternos amigos e incentivadores.

## AGRADECIMENTOS

A Deus, pelo dom vida.

Aos meus pais, Décio e Deusana, por todas as renúncias feitas em nome da minha formação.

Ao meu namorado, Gabriel, pelo incentivo ao longo dessa caminhada.

Ao meu orientador, Prof. Glauber Brante, por toda a generosidade, paciência e dedicação.

Ao coorientador, Prof. Richard Demo Souza, por suas valiosas sugestões.

Aos demais professores integrantes da banca examinadora, por aceitarem o convite.

Educai as crianças e não será necessário punir os adultos. *Pitágoras*

## RESUMO

DOS REIS, Ana Flávia. Eficiência Energética no Posicionamento de Estações Rádio-Base Móveis Baseadas em VANTs Aplicando Aprendizagem por Reforço. 47 f. Dissertação – Programa de Pós-graduação em Engenharia Elétrica e Informática Industrial, Universidade Tecnológica Federal do Paraná. Curitiba, 2020.

Para suprir as exigências de seus usuários, as redes celulares 5G devem ser altamente eficientes, tornando a eficiência energética um dos desafios e objetivos primários dos cenários de comunicações modernos. Além disso, o aumento no número de dispositivos e requisitos dos usuários dessas redes podem tornar as soluções analíticas aplicadas no contexto da eficiência energética e otimização de parâmetros da comunicação não factíveis, dada a complexidade atrelada à essas soluções. Para superar esse problema, soluções mais robustas e flexíveis, que exploram os dados gerados pela rede e tornam decisões em tempo real são as exigidas por essas redes futuras. Nesse contexto, a utilização de técnicas de Aprendizado de Máquina, capazes de analisar uma enorme quantidade de dados e aprender com eles, é uma possível solução, podendo ser aplicadas, por exemplo, em diferentes contextos de redes auto-organizáveis.

O objetivo da presente dissertação é analisar uma aplicação de otimização do posicionamento de VANTs, Veículos Aéreos Não Tripulados, em um cenário de eventos temporários baseando-se em um algoritmo *Q-learning*, uma técnica de aprendizado por reforço, na qual os agentes das soluções propostas aprendem a agir para maximizar uma determinada recompensa. Duas diferentes recompensas são propostas e comparadas, com a primeira visando apenas minimizar o número de usuários em indisponibilidade e a segunda buscando maximizar a eficiência energética, considerando a quantidade de energia usada para a movimentação VANTs. A solução *Q-learning* é responsável por determinar o melhor posicionamento no espaço tridimensional e a potência de transmissão ideal, dados alguns níveis de potência possíveis.

Os resultados obtidos mostram que a otimização da potência de transmissão implementada reduziu a probabilidade de indisponibilidade dos usuários da rede, uma vez que a interferência entre os VANTs foi minimizada a partir da alocação da potência adequada e, além disso, é mostrado que a introdução de um fator de desconto na recompensa do algoritmo relacionado ao gasto de energia para as movimentações VANTs foi capaz de aumentar consideravelmente a eficiência energética da rede, com uma penalidade no desempenho em relação à cobertura de usuários. Finalmente, para as duas soluções consideradas e de acordo os seus respectivos objetivos, a técnica *Q-learning*, a partir da sua capacidade de aprender com a iteração com um ambiente e de experiências anteriores, mostrou-se capaz de resolver o problema de auto-organização da rede baseada em VANTs de uma maneira eficiente, apresentando-se como uma solução promissora para as futuras redes celulares.

**Palavras-chave:** Eficiência Energética, VANTs, Aprendizado por Reforço.



## ABSTRACT

DOS REIS, Ana Flávia. ENERGY EFFICIENCY IN POSITIONING UAV-BASED MOBILE BASE STATIONS APPLYING REINFORCEMENT LEARNING. 47 f. Dissertação – Programa de Pós-graduação em Engenharia Elétrica e Informática Industrial, Universidade Tecnológica Federal do Paraná. Curitiba, 2020.

To meet users demands, 5G cellular networks require high operational efficiency, making energy efficiency one of the primary challenges and objectives of modern communication systems. In addition, the increase in the number of devices and user requirements in these networks may turn analytical solutions applied in the context of energy efficiency and optimization of communication parameters not feasible, given the complexity associated to these solutions. To overcome this problem, more robust and flexible solutions that exploit the data generated by the network and make real-time decisions are required by these future networks. In this context, the use of Machine Learning techniques, capable of analyzing a big amount of data and learning from it, is a possible solution and can be applied, for example, in different self-organizing network contexts.

The aim of this work is to analyze a positioning optimization application of UAVs, Unmanned Aerial Vehicles, in a scenario of temporary events based on a Q-learning algorithm, a reinforcement learning technique in which the agents learn to act to maximize a given reward. Two different rewards are proposed and compared, with the first aiming only to minimize the number of users in outage and the second provided to maximize energy efficiency by considering the amount of energy used for UAVs movements. The Q-learning solution is responsible for determining the best three-dimensional positioning and optimal transmit power given some possible power levels.

The results show that the optimization of the transmission power reduced the percentage of users in outage, since interference between UAVs was minimized by allocating adequate power. Furthermore, it is shown that the introduction of a discount factor in the algorithm reward for UAV movements has been able to considerably increase the energy efficiency of the network, with a performance penalty over the user coverage. Finally, for the two solutions considered and in accordance with their respective objectives, the Q-learning technique, based on its ability to learn from iteration with an environment and from previous experiences, proved to be able to solve the problem of self-organization of the UAV-based network in an efficient manner, presenting itself as a promising solution for future cellular networks.

**Keywords:** Energy Efficiency, UAVs, Reinforcement Learning.

## LISTA DE FIGURAS

Figura 1	–	Interação entre agente e ambiente no aprendizado por reforço. ....	20
Figura 2	–	Ramificações das redes auto-organizáveis. ....	22
Figura 3	–	Modelo urbano de distribuição de edifícios. ....	26
Figura 4	–	Área de cobertura do VANT $j$ à altura $h_j$ com ângulo de abertura da antena $\theta$ . ....	27
Figura 5	–	Porcentagem de usuários indisponíveis em função do número de episódios executados em um cenário com $D = 7$ VANTs. ....	34
Figura 6	–	Usuários indisponíveis em função do número de VANTs ( $D$ ). ....	35
Figura 7	–	Potência consumida pelos VANTs. ....	36
Figura 8	–	EE para diferentes pesos $w$ empregados à recompensa $r_2$ . ....	37
Figura 9	–	Consumo de potência devido às movimentações dos VANTs para a recompensas $r_1$ e $r_2$ em função do número de VANTs ( $D$ ). ....	38
Figura 10	–	Número médio de movimentações dos VANTs para as recompensas propostas. ....	38
Figura 11	–	Porcentagem de usuários em indisponibilidade em função de $D$ para as recompensas propostas. ....	39
Figura 12	–	Trajetória feita pelo 7º VANT durante um episódio específico com as recompensas $r_1$ e $r_2$ . ....	40
Figura 13	–	EE para as recompensas propostas. ....	41
Figura 14	–	Tempo de vida da rede em função do número de VANTs. ....	41

## LISTA DE TABELAS

Tabela 1 – Parâmetros para o ambiente urbano. ....	26
Tabela 2 – Parâmetros de simulação .....	33

## LISTA DE SIGLAS

VANTs	Veículos Aéreos Não Tripulados
EE	Eficiência Energética
ERBs	Estações Rádio Base
WSNs	<i>Wireless Sensor Networks</i>
RF	Rádio-frequência
RNs	<i>Redes Neurais</i>
ILP	<i>Integer Linear Programming</i>
ML	<i>Machine Learning</i>
RL	<i>Reinforcement Learning</i>
IA	Inteligência Artificial
UL	<i>Unsupervised Learning</i>
SINR	<i>Signal to Interference Plus Noise Ratio</i>
SARSA	<i>State-Action-Reward-State-Action</i>
VFA	<i>Value Function Approximation</i>

## LISTA DE SÍMBOLOS

$\alpha$	Razão entre a área construída e a área total
$\beta$	Número médio de edifícios por km <sup>2</sup>
$\gamma$	Parâmetro de escala das alturas dos edifícios
$D$	Numero de VANTs
$\theta$	Ângulo de abertura das antenas
$h_j$	Altura de cada VANT
$\kappa_{ij}$	Perda de percuso entre VANT e usuário
$d_{ij}$	Distância entre o VANT e o usuário
$f_c$	Frequência da portadora
$c$	Velocidade da luz
$\xi$	Perdas adicionais
$B$	Largura de banda
$N_0$	Densidade espectral de ruído
$P_{r_{ij}}$	Potência de referência recebida
$P_{t_j}$	Potência de transmissão
$P(V)$	Potência consumida para a movimentação do VANT
$\Omega$	Velocidade angular das lâminas
$\psi$	Raio do rotor
$v_0$	Velocidade média induzida pelo rotor
$d_0$	Taxa de arrasto da fuselagem
$\rho$	Densidade do ar
$s$	Solidez do rotor
$A$	Área do rotor
$P_0$	Potência do perfil da lâmina
$\delta$	Coefficiente de arrasto do perfil
$P_i$	Potência induzida
$\zeta$	Fator de correção da potência induzida
$\chi$	Peso do VANT
$U_{out}$	Porcentagem de usuários em indisponibilidade
$N_{out}$	Número de usuários indisponíveis
$U_j$	Número de usuários alocados ao VANT $j$
$R_b$	Taxa de bit
$E_{total}$	Energia total gasta para movimentar VANTs
$\lambda$	Taxa de aprendizagem do algoritmo
$\phi$	Fator de desconto do algoritmo

## SUMÁRIO

<b>1 INTRODUÇÃO</b> .....	<b>14</b>
1.1 OBJETIVOS .....	16
1.1.1 Objetivo Geral .....	16
1.1.2 Objetivos Específicos .....	16
1.2 ESTRUTURA DO TRABALHO .....	17
<b>2 REFERENCIAL TEÓRICO</b> .....	<b>18</b>
2.1 APRENDIZADO DE MÁQUINA PARA A OTIMIZAÇÃO DA EFICIÊNCIA ENERGÉTICA EM REDES SEM FIO .....	18
2.1.1 Aprendizado por Reforço .....	19
2.2 REDES SEM FIO AUTO-ORGANIZÁVEIS .....	21
<b>3 DESENVOLVIMENTO</b> .....	<b>25</b>
3.1 MODELO URBANO .....	25
3.2 MODELO DE COMUNICAÇÃO .....	26
3.3 MODELO DE CONSUMO DE ENERGIA .....	28
3.4 MÉTRICAS .....	29
3.5 ALGORITMO PROPOSTO .....	30
3.5.1 Recompensa dedicada à cobertura da rede .....	32
3.5.2 Recompensa dedicada à cobertura da rede e consumo energético .....	32
<b>4 RESULTADOS E DISCUSSÕES</b> .....	<b>33</b>
4.1 MAXIMIZAÇÃO DE USUÁRIOS COBERTOS PELA REDE SEM FIO .....	34
4.2 MAXIMIZAÇÃO DE USUÁRIOS COBERTOS PELA REDE COM RESTRICÇÕES EM CONSUMO DE ENERGIA .....	37
<b>5 CONCLUSÕES</b> .....	<b>42</b>
<b>REFERÊNCIAS</b> .....	<b>44</b>

## 1 INTRODUÇÃO

Conforme descrito por (CISCO, 2017; HUAWEI, 2013), é esperado que as próximas gerações de comunicações atinjam novos padrões fornecimento de cobertura e também melhorem as experiências dos usuários da rede em relação aos seus requisitos. Para isso, as redes 5G precisarão ter desempenho suficiente para promover conectividade à uma grande quantidade desses dispositivos. Além disso, os sistemas 5G também deverão promover cobertura e capacidade em qualquer localização, enquanto aprimorando as experiências dos usuários e melhorando as taxas de dados (HUAWEI, 2013; KLAINÉ et al., 2017a). Dessa forma, a 5ª geração de redes sem fio não deverá apenas superar as limitações das redes de celular atuais, mas também introduzir novos usos e cenários, possibilitando uma nova gama de aplicações. Entretanto, uma grande preocupação cresce dada todas essas expectativas, uma vez que para que o alcance de todos esses requisitos seja possível, as redes sem fio 5G também tendem a consumir mais energia que os sistemas atuais (BUZZI et al., 2016). Para atingir essas expectativas de maneira sustentável, os estudos que buscam viabilizar as redes 5G com o objetivo de torná-las energeticamente eficientes são um dos grandes focos de pesquisas de grupos vigentes (IMRAN et al., 2019).

Em (BUZZI et al., 2016), os autores descrevem os aspectos econômicos, operacionais e ambientais como as principais preocupações na última década dos sistemas de comunicação sem fio, enquanto a eficiência energética (EE) surge como uma nova figura de mérito proeminente para as redes celulares. Para superar os problemas inerentes à sua implementação, as redes 5G deverão se tornar energeticamente eficientes, com o desenvolvimento sustentável não apenas a respeito de suas estações rádio base (ERBs), mas também dos dispositivos móveis (TULLBERG et al., 2016). Dessa forma, a EE está estabelecida como um dos objetivos primários dos sistemas de comunicação modernos, tornando-se o foco de diversos grupos acadêmicos e industriais de pesquisa.

Na literatura, diferentes métricas vem sendo utilizadas para definir a EE. O método mais comum para defini-la é como a razão entre o *throughput* do sistema e a energia consumida, sendo expressa em [bits/Joules] (FENG et al., 2013). Baseados nessa definição, muitos estudos na literatura tentam maximizar a EE dependendo de uma aplicação ou ambiente em particular. Por exemplo, em redes de celular, EE tem sido alvo de interesse devido ao potencial em reduzir custos operacionais (FENG et al., 2013). Além disso, o crescimento exponencial do número de dispositivos em redes de comunicações gera sérias preocupações em relação ao crescimento sustentável, devido à emissões de carbono

chegando a taxas alarmantes. Portanto, muitos estudos recentes atentam às redes de celular mais sustentáveis, mostrando uma melhora significativa na EE em ordem a suprir as demandas dos usuários e minimizar esses efeitos adversos (GANDOTRA et al., 2017). Nesses cenários de larga escala, as ERBs representam a maior parte da energia consumida, que pode chegar a 80% do total devido às fontes de alimentação, resfriamento e conexão à rede elétrica (AUER et al., 2011).

Por exemplo, o trabalho descrito em (RICHTER et al., 2009) investiga os impactos de diferentes configurações em *layouts* das ERBs no consumo energético de redes de celulares, mostrando que a economia de energia obtida pela implantação de microssites, ERBs auxiliares e de menor potência, é moderada em cenários de carga total e depende muito do consumo de energia compensado entre os macro e microssites. A extensão desse trabalho é apresentada em (FEHSKE et al., 2009) e emprega simulações para avaliar o consumo de energia para a aplicação desses microssites adicionais às ERBs dos macrossites, em que é mostrado que essa estratégia permite uma redução significativa no consumo de potência da rede. A EE também é estudada em outros tipos de rede sem fio, como as redes de sensores, do inglês *wireless sensor networks* (WSNs). Em (CUI et al., 2004, 2005; PEREIRA et al., 2018), por exemplo, os autores mostraram que quando a distância de transmissão é grande ( $> 100$  m) a potência de transmissão domina o consumo total de potência, enquanto para aplicações de pequeno alcance ( $< 10$  m) a potência usada pelo circuito de rádio-frequência (RF) pode dominar a potência total consumida. Portanto, um modelo apropriado para o consumo de potência de WSNs é particular a cada cenário, assim como o método utilizado para otimizar a EE pode ser consideravelmente diferente.

Tradicionalmente técnicas como otimização convexa (XU; QIU, 2013), programação fracional (ZAPPONE et al., 2015) e teoria de jogos (ALSKAIF et al., 2015) tem sido aplicadas para maximizar a EE em redes sem fio. Esses métodos tentam encontrar soluções analíticas, otimizando muitos parâmetros de comunicação como, por exemplo, a potência de transmissão, ordem de modulação ou modo operacional. Entretanto, muitas simplificações são geralmente requeridas para obter as expressões dessas soluções, de modo que apenas cenários ponto a ponto são frequentemente considerados. Além disso, várias premissas irreais são assumidas, o que também é fator limitante dessas soluções, como muitas delas requerem informações que não são oportunas ou realisticamente disponíveis em uma rede móvel, como o número de usuários conectados em um horário específico ou suas posições exatas. Além disso, espera-se que as redes celulares futuras se tornem ainda mais complexas devido ao crescimento do número



de dispositivos conectados, da densidade do tráfego e de seus requisitos heterogêneos (KLAINÉ et al., 2017a) e, como resultado dessa complexidade, a aplicação de soluções de otimização clássica (ZAPPONE et al., 2015; ALSKAIF et al., 2015; XU; QIU, 2013) pode não ser viável em cenários de rede futuros destacados.

Para superar esses problemas, fica claro que as soluções analíticas não serão ideais, pois as redes se tornam cada vez mais complexas. Além disso, seriam preferidas soluções mais robustas e flexíveis, que explorem os dados gerados pela rede e tomem decisões em tempo real. Dessa forma, técnicas capazes de analisar uma enorme quantidade de dados e aprender com eles, como algoritmos de aprendizado de máquina, podem ser uma solução viável para redes futuras (KLAINÉ et al., 2017a).

Nesse sentido, o presente trabalho apresenta uma aplicação de aprendizagem de máquina com a finalidade de posicionar Veículos Aéreos Não Tripulados (VANTs) formadores de uma rede sem fio capaz de prover conectividade à usuários, restringindo-se a um cenário de eventos temporários onde toda a conectividade é promovida pelos VANTs otimizando, e configurando a rede de forma autônoma e energeticamente eficiente.

## 1.1 OBJETIVOS

### 1.1.1 Objetivo Geral

Desenvolver uma solução baseada em *Q-learning*, uma técnica de aprendizado por reforço, para promover o posicionamento de múltiplos VANTs e maximizar a eficiência energética de uma rede sem fio em um cenário de eventos temporários.

### 1.1.2 Objetivos Específicos

- Promover uma revisão bibliográfica do uso de técnicas de aprendizagem de máquina para a otimização da eficiência energética em redes sem fio;
- Analisar o impacto da aplicação de soluções de aprendizado por reforço no consumo energético de uma rede formada por múltiplos VANTs;
- Desenvolver um método de aplicação de aprendizado por reforço para a maximização da eficiência energética em uma rede sem fio formada por múltiplos VANTs;
- Selecionar um modelo de consumo de energia representativo à VANTs com hélice, aplicados a esse trabalho;

- Fornecer conclusões referentes à aplicação de aprendizado por reforço em redes auto-organizáveis e estudar tendências futuras para a obtenção da eficiência energética aplicadas a esse contexto.

## 1.2 ESTRUTURA DO TRABALHO

No presente trabalho o desempenho de uma rede sem fio baseada em VANTs é analisado em termos de cobertura da rede e eficiência energética em seu posicionamento. A solução proposta é baseada na técnica *Q-learning* de aprendizado por reforço, onde a recompensa pode ser balanceada entre os objetivos pretendidos. Dessa forma, a solução busca adaptar o posicionamento dos VANTs de forma autônoma para promover cobertura para os usuários de uma rede energeticamente eficiente. Além disso, a potência de transmissão também é regulada com o objetivo de reduzir a interferência da rede.

Os resultados obtidos mostram que a introdução de restrições ao consumo de energia implicam em uma menor movimentação dos VANTs, maximizando a eficiência energética da rede proposta e apresentando ganhos em termos de autonomia e tempo de vida para diferentes quantidades de VANTs considerados. Para tanto, esse estudo é organizado de forma a apresentar o referencial teórico, que contempla o estudo de redes de celulares 5G, de técnicas de aprendizado de máquina para a otimização de redes sem fio em termos de eficiência energética e da auto-organização de redes sem fio no Capítulo 2. Na sequência, o Capítulo 3 é iniciado pela apresentação do modelo do sistema, incluindo o cenário urbano, modelo de comunicação e modelo de consumo dos VANTs, seguido pela apresentação da solução *Q-learning* proposta. Ainda este capítulo apresenta os resultados obtidos e suas respectivas discussões. Finalmente, o Capítulo 5 reflete as considerações finais do trabalho e propõe trabalhos para o aprofundamento desse estudo.

## 2 REFERENCIAL TEÓRICO

### 2.1 APRENDIZADO DE MÁQUINA PARA A OTIMIZAÇÃO DA EFICIÊNCIA ENERGÉTICA EM REDES SEM FIO

Problemas ligados à complexidade das redes 5G alertam à necessidade de soluções mais robustas, flexíveis e capazes de analisar uma maior quantidade de dados. Alguns exemplos incluem redes neurais (RNs) combinadas com programação linear inteira, do inglês *integer linear programming* (ILP), como em (PELEKANOU et al., 2018), como um meio de otimizar a infraestrutura de um cenário de comunicação 5G, com o objetivo de minimizar o consumo geral de energia da rede. Os resultados indicam que o uso de uma RN pode reduzir consideravelmente a complexidade computacional quando comparado ao ILP sozinho, enquanto fornece desempenho semelhante. A eficiência da técnica proposta também é demonstrada em um fornecimento genérico de serviços de *backhaul* e *fronthaul*, em que o consumo total de energia da infraestrutura 5G foi minimizado.

Em (ZHENG et al., 2016), várias técnicas de Aprendizado de Máquina, do inglês *Machine Learning* (ML), são usadas para explorar *big data*, ou seja, dados que geralmente são coletados e armazenados por operadoras móveis para fins de otimização. Como os autores concluem, apesar dos desafios em termos de coleta de dados, sobrecarga de comunicação e latência, combinar ferramentas de ML com análise de *big data* é uma solução promissora para otimizar constantemente o desempenho da rede 5G. Além disso, (KIRAN et al., 2016) também explora *big data* de redes móveis para alocação de recursos de rádio, implementando um controlador difuso para alocar com eficiência a largura de banda entre os usuários. Como resultado, o método proposto é mostrado como redutor da latência de processamento, o que permite uma menor complexidade para soluções distribuídas. Comum aos trabalhos em (ZHENG et al., 2016; KIRAN et al., 2016) está a noção de que as abordagens baseadas em ML são capazes de acompanhar as rápidas condições de mudança do ambiente sem fio, especialmente em cenários densos representativos dos sistemas de comunicação 5G, sendo ferramentas importantes para melhorar continuamente o desempenho da rede e a eficiência operacional.

Recentemente, o ML foi empregado em (ZAPPONE et al., 2018) para permitir a alocação de energia *online*, visando maximizar a EE em redes de sem fio. Como mostram seus resultados, o modelo baseado em aprendizado profundo com RNs é capaz de alcançar o desempenho do método analítico, exigindo menor complexidade computacional.

Em (ORO et al., 2018), os autores maximizam a EE, dada a alocação de potência de transmissão e a atribuição de subportadoras. Uma vez que a otimização é um problema fracionário não-convexo desafiador, os autores empregam uma combinação de programação fracionária, ML usando uma abordagem de aprendizado estocástico e teoria dos jogos. Como resultado, é demonstrado que a solução ML proposta tem desempenho semelhante a outros algoritmos da literatura, mas o método possui uma complexidade linear no número de usuários e subportadoras, enquanto outras soluções disponíveis podem garantir apenas uma complexidade polinomial no número de usuários e subportadoras. Em resumo, soluções como as propostas de (ZAPPONE et al., 2018; ORO et al., 2018) mostram que o ML pode fornecer soluções robustas e quase ideais às necessidades intrínsecas às redes celulares 5G.

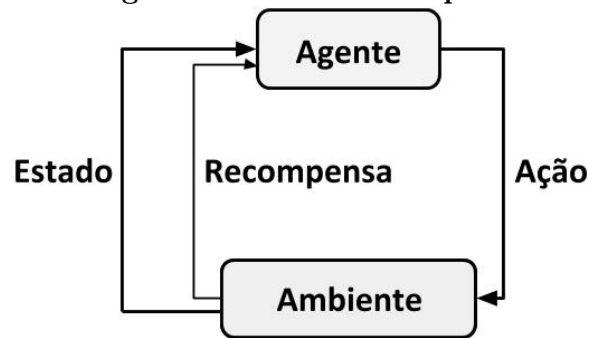
### 2.1.1 Aprendizado por Reforço

Devido ao crescimento da densidade e dos requisitos de tráfego nas redes de comunicação sem fio modernas, a aplicação de soluções analíticas pode ser bastante imprecisa ou até inviável. Nesse sentido, soluções mais flexíveis que exploram os dados gerados pela rede e tomam decisões em tempo real são importantes para implementações práticas. Dessa forma, o aprendizado de máquina é uma ferramenta poderosa para permitir a auto-organização da rede de comunicação usando VANTs (KLAINÉ et al., 2017b).

O aprendizado por reforço, do inglês *Reinforcement Learning* (RL), por sua vez, é uma técnica de aprendizado de máquina onde os agentes da solução aprendem a partir do resultado de suas ações e, dessa forma, apresenta-se como uma solução relevante para problemas onde há necessidade de adequação da aprendizagem de acordo com as iterações com o ambiente. Para essas soluções iterativas, muitas vezes aplicadas à ambientes desconhecidos, o RL permite que o agente aprenda a partir de suas experiências, de acordo com a atualização da recompensa proposta. Como apresentado pela Figura 1, na proposta do método de RL, o agente da solução interage com o ambiente continuamente a partir da seleção de ações e o ambiente, por sua vez, atualiza o conhecimento desse agente com a apresentação dos estados e da recompensa proposta. Dessa forma, o objetivo do agente é maximizar a recompensa obtida a partir das suas ações que são escolhidas de acordo com a definição de uma política (SUTTON et al., 1998).

Por exemplo, o trabalho desenvolvido em (KLAINÉ et al., 2018) aplicou uma solução *Q-learning* para otimizar a posição dos VANTs em um cenário de rede de

**Figura 1 – Interação entre agente e ambiente no aprendizado por reforço.**



**Fonte:** Adaptado de (SUTTON et al., 1998).

comunicação de emergência. O principal objetivo da solução é maximizar o número de usuários cobertos pela rede de comunicação, a partir do melhor posicionamento de um número determinado de VANTs, com intuito de maximizar a cobertura da rede. Quando comparada a diferentes estratégias de posicionamento, como posicionar os VANTs de forma aleatória, em torno de um círculo no centro do cenário ou nos pontos da rede destruída anteriormente, a solução de aprendizado por reforço apresentou desempenho aprimorado, minimizando o número de usuários indisponíveis e convergindo mais rapidamente. No entanto, a eficiência energética não é investigada em por (KLAINÉ et al., 2018), embora seja uma questão crucial em vista da quantidade limitada de energia e da necessidade de recarga constante dos VANTs. Outro trabalho, proposto por (BAYERLEIN et al., 2018), também aplicou uma técnica de aprendizado por reforço para otimizar a trajetória dos VANTs para atender usuários em uma rede sem fio mas nesse trabalho o objetivo da solução proposta é aprender a trajetória que maximiza o parâmetro de taxa de transferência da rede.

Além disso, outra solução de aprendizado por reforço foi aplicada por (WANG et al., 2019) para rastrear os movimentos dos VANTs, a fim de maximizar a transmissão de dados de *downlink* em uma rede sem fio. Como resultado, o algoritmo proposto fornece uma capacidade de rede aprimorada quando comparada à implantação do VANTs em posições estáticas aleatórias. O aprendizado por reforço profundo foi aplicado por (CHALLITA et al., 2019), o primeiro estudo a usar essa técnica para gerenciar interferências em uma rede baseada em VANTs. A solução proposta reduziu em até 62% a latência da comunicação por VANT, aumentando em 14% a eficiência energética quando comparada a uma solução heurística que define as trajetórias dos VANTs pela distância mais curta em direção ao destino. Além disso, a minimização do consumo de energia para redes baseadas em VANT também foi considerada por (ZHAN; LAI, 2019), estudo no

qual o objetivo é projetar as trajetórias ideais de VANTs de hélice e em que a otimização alternada e a técnica sucessiva de aproximação convexa foram combinadas para minimizar o consumo de energia.

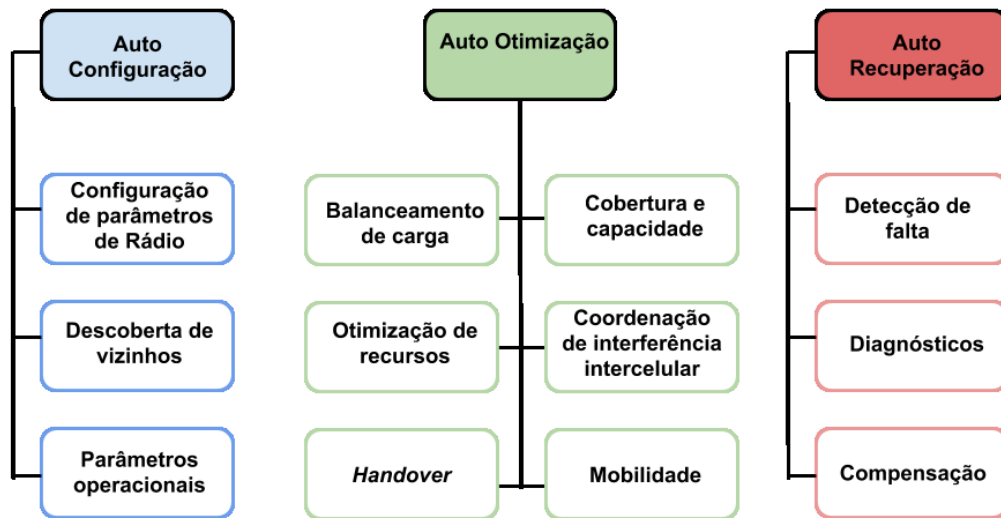
## 2.2 REDES SEM FIO AUTO-ORGANIZÁVEIS

Como definido em (KLAINÉ et al., 2017a; ALIU et al., 2013), redes sem fio auto-organizáveis são redes que pode ser adaptáveis, escaláveis, estáveis, ágeis e autônomas para gerenciar e controlar determinados objetivos da rede. Como tal, essas redes são capazes de tomar decisões autonomamente em tempo real, bem como aprender e melhorar seu desempenho com base em dados históricos e anteriores. Devido a essa automação, a auto-organização pode fornecer inteligência dentro da rede, a fim de facilitar o trabalho das operadoras e reduzir a complexidade geral, proporcionando economia significativa de custo e energia para as operadoras móveis.

Essas redes podem ser divididas em três principais ramos (KLAINÉ et al., 2017a; ALIU et al., 2013): auto-configuração, auto-otimização e auto-recuperação, que são, juntas, denominadas como *funções auto-x*, do inglês *self-x functions*. Nesse contexto, as técnicas de ML certamente podem melhorar a auto-organização das redes, permitindo que essa rede se adapte observando seu status atual e use sua experiência prévia para ajustar parâmetros em ações futuras. Nesse sentido, melhorar a EE pode ser uma das principais diretrizes para o ajuste de parâmetros. Em (IMRAN et al., 2019), foi realizada uma revisão bibliográfica das aplicações de inteligência artificial (IA) para aprimorar a auto-organização das redes e promover a melhoria da eficiência energética a partir da adaptação da rede conforme sua experiência. A Figura 2 ilustra algumas aplicações da auto-organização da rede, destacando as principais ramificações das *funções auto-x* e os casos mais comuns de seu uso.

Outro elemento capaz de aumentar a EE da rede, assim como a capacidade e cobertura, é o ajuste do ângulo de abertura das antenas. Dependendo da distribuição de usuários nas ERBs, a configuração do ângulo de abertura das antenas pode melhorar a recepção do sinal e, além disso, pode reduzir a interferência entre as ERBs vizinhas, melhorando o desempenho da rede em relação à EE (DANDANOV et al., 2017). Conforme apresentado em (YILMAZ et al., 2009), a mudança do ângulo das antenas é possível mecanicamente e eletricamente e, uma vez que a otimização da inclinação da antena depende muito da carga da célula, as técnicas ML se tornam úteis nesse contexto. Por exemplo, os padrões de mobilidade e atividade humana podem ser usados para prever

Figura 2 – Ramificações das redes auto-organizáveis.



Fonte: Autoria própria, adaptado de (IMRAN et al., 2019).

as demandas da rede, permitindo que o ângulo das antenas componentes da rede seja ajustado de maneira inteligente e proativa, melhorando o desempenho. Uma das técnicas mais promissoras nesse contexto preditivo é a aprendizagem por reforço, que permite otimização adaptativa constante da rede, promovendo ações autônomas baseadas nas observações do algoritmo, a fim de melhorar o rendimento da rede ou EE.

As técnicas de RL para otimização de inclinação de antena foram estudadas, por exemplo, em (DANDANOV et al., 2017). Diferentemente de outras pesquisas que otimizam a inclinação da antena apenas em ambientes de rede estáticos, esse trabalho considera a rede móvel um ambiente dinâmico, adaptável às distribuições de usuários atuais. Os resultados da simulação mostram que o algoritmo melhora a taxa geral de dados da rede, em comparação com um cenário sem a otimização da inclinação da antena.

Além do ajuste de inclinação de antenas, outra área de aplicação de algoritmos de RL voltados a redes auto-organizáveis está na suspensão de ERBs. Os autores de (MIOZZO et al., 2017), por exemplo, conceberam uma abordagem de RL baseada em *Q-learning* onde múltiplos agentes distribuídos para projetar políticas liga/desliga da rede, em que as ERBs devem ser ativadas ou desativadas. Além disso, a solução proposta também é distribuída e tem baixa complexidade, permitindo que seja executada *online*. Os resultados mostram que a EE da rede é aprimorada em até 15% quando comparado a um esquema em que os agentes da rede seguem uma política gananciosa ( $\epsilon$ -greedy), em

que estados aleatórios são visitados pelos agentes de acordo com uma probabilidade  $\epsilon$ . Os autores, em (SINCLAIR et al., 2013) apresentam outro ramo do ML útil em redes sem fio auto-organizáveis, desta vez na área de aprendizado não-supervisionado, do inglês *Unsupervised Learning* (UL). Neste artigo, um novo algoritmo de mapa de Kernel auto-organizável é proposto para minimizar transferências de dados desnecessárias e, como consequência, aumentar a capacidade da rede e melhorar sua EE. Simulações demonstram que o algoritmo UL proposto reduz em até 70% das transferências desnecessárias na rede.

Conforme destacado por (ZENG et al., 2016; GUPTA et al., 2016), outro aspecto importante dos sistemas baseados em VANTs é o planejamento de trajetória. Nesse sentido, as oportunidades e os desafios da aplicação dos VANTs podem ser potencializados pelo planejamento do caminho apropriado, considerando a auto-organização da rede. Os autores, em (MERWADAY; GUVENC, 2015), citam o fornecimento de conectividade rápida ao usuário nos horários e locais mais solicitados, por meio de uma infraestrutura de custo mais baixo quando comparada com redes não organizadas, como características habilitadas pela auto-organização de redes. Os autores de (ZENG et al., 2016) também apresentam restrições de energia como um aspecto importante para as comunicações sem fio com a implantação do VANTs, destacando a economia no consumo de energia como uma grande preocupação.

Nesse contexto, os autores em (ZENG; ZHANG, 2017) aplicaram técnicas de otimização convexa para resolver problemas de otimização de trajetória. Outro trabalho, apresentado em (ALZENAD et al., 2017), propôs um posicionamento de VANTs com eficiência energética, baseado na maximização da cobertura da rede, considerando a energia transmitida mínima exigida. Por outro lado, o trabalho em (PATRA; MUTHUCHIDAMBARANATHAN, 2018) empregou uma otimização para aumentar a eficiência do espectro e a eficiência energética de uma rede baseada em VANTs. Uma observação importante é que os resultados obtidos em todos esses estudos mencionados são baseados em VANTs de asa fixa, mais facilmente encontrados na literatura. Pelo contrário, um modelo VANT de hélice para comunicação eficiente em termos de energia com otimização de trajetória é considerado em (ZENG; ZHANG, 2017; ZENG et al., 2019), onde várias vantagens desse tipo de drone são mostradas em comparação com os VANTs de asa fixa, como a capacidade de decolar e pousar verticalmente, tornando-o mais popular e atraente para esse mercado.

Diferente dos estudos apresentados, o presente trabalho aborda o desempenho em cobertura e a eficiência energética do posicionamento de VANTs de hélice, de



modo que o modelo de consumo de energia proposto por (ZENG et al., 2019) é empregado. A solução é baseada na técnica *Q-learning*, onde a recompensa pode ser adaptada para equilibrar as metas e, portanto, a abordagem consiste em adaptar o posicionamento dos VANTs de maneira autônoma, a fim de fornecer cobertura e eficiência energética. Além disso, diferente de (KLAINÉ et al., 2018), a potência de transmissão dos VANTs é adaptada para reduzir a interferência, estendendo os resultados anteriormente apresentados por (PARISOTTO et al., 2019) e é mostrado que a adição de restrições de consumo de energia implica em menores movimentações dos VANTs, maximizando a eficiência energética da rede proposta.

### 3 DESENVOLVIMENTO

O presente trabalho propõe uma solução para a maximização da cobertura da rede e da eficiência energética a partir da implementação de um algoritmo capaz de otimizar o posicionamento de múltiplos VANTs formadores de uma rede sem fio provedora de conexão para usuários em uma situação de eventos temporários. A solução baseada em *Q-learning*, uma técnica de aprendizado por reforço, é aplicada para determinar o melhor posicionamento em um espaço tridimensional, bem como a potência de transmissão ótima dados diferentes níveis possíveis. Parte desses resultados e suas respectivas discussões são apresentados em (REIS et al., 2019) onde, para o cenário proposto, mostra-se que a possibilidade de alocar diferentes níveis de potência de transmissão aos VANTs formadores de uma rede de celular maximizou a cobertura e minimizou o consumo energético.

#### 3.1 MODELO URBANO

É considerado um cenário de evento temporário, em que o número de usuários a serem atendidos pela rede de celular é desconhecido e muito maior do que a implantação usual da região em questão. Portanto, como pior hipótese possível, considera-se que a conectividade é fornecida apenas pelos VANTs. Além disso, o cenário urbano considerado segue o modelo de distribuição de edifícios e usuários definido pela União Internacional de Telecomunicações (ITU-R). De acordo com (ITU-R, 2003), três são os parâmetros necessários para descrever qualquer ambiente urbano, estes são:

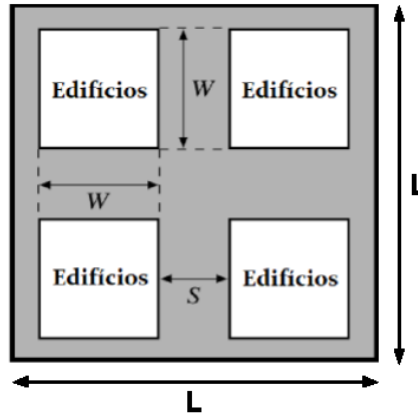
- $\alpha$ : a razão entre a área construída e a área total;
- $\beta$ : o número médio de edifícios por  $\text{km}^2$ ;
- $\gamma$ : o parâmetro de escala das alturas dos edifícios, de acordo com a função densidade de probabilidade de Rayleigh (PAPOULIS; PILLAI, 2002),

$$p(x) = \frac{x}{\gamma^2} \cdot \exp\left(\frac{-x^2}{2\gamma^2}\right). \quad (1)$$

Para um cenário urbano, de acordo com (AL-HOURANI et al., 2014), assume-se que a largura dos edifícios é dada por

$$W = 1000 \cdot \sqrt{\frac{\alpha}{\beta}}, \quad (2)$$

Figura 3 – Modelo urbano de distribuição de edifícios.



Fonte: Adaptado de (AL-HOURANI et al., 2014).

enquanto o espaçamento entre os edifícios é dado por

$$S = \frac{1000}{\beta} - W. \quad (3)$$

Conforme a classificação feita em (AL-HOURANI et al., 2014), um ambiente urbano apresenta os valores dos parâmetros  $\alpha$ ,  $\beta$  e  $\gamma$  descritos na Tabela 1.

Tabela 1 – Parâmetros para o ambiente urbano.

$\alpha$	$\beta$	$\gamma$
0,3	500	15

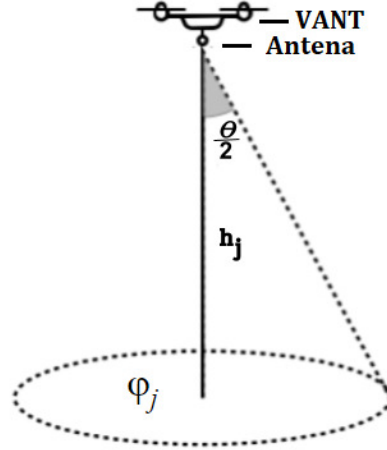
Fonte: Adaptado de (AL-HOURANI et al., 2014).

São assumidos edifícios quadrados e variáveis  $W$  e  $S$  dimensionados em metros. A Figura 3 apresenta o cenário proposto, onde são considerados  $K$  usuários aleatoriamente distribuídos na área  $L \times L$  representada, onde  $\mathcal{K} = \{1, 2, \dots, K\}$  denota o conjunto de usuários ativos.

### 3.2 MODELO DE COMUNICAÇÃO

É considerado um conjunto de  $\mathcal{D} = \{1, 2, \dots, D\}$  VANTs, em que  $D$  denota o número de VANTs disponíveis e trabalhando como ERBs móveis. Além disso, cada VANT possui uma antena com um ângulo de abertura denotado por  $\theta$ . Dessa forma, como

Figura 4 – Área de cobertura do VANT  $j$  à altura  $h_j$  com ângulo de abertura da antena  $\theta$ .



Fonte: Adaptado de (KLAINÉ et al., 2018).

ilustrado pela Figura 4, a área de cobertura de cada VANT  $j \in \mathcal{D}$ , é descrita por

$$\varphi_j = h_j \cdot \tan \frac{\theta}{2}, \quad (4)$$

em que  $h_j$  é a altura de cada VANT  $j$ .

A perda de percurso  $\kappa_{ij}$ , em dB, entre um VANT  $j \in \mathcal{D}$  e um usuário da rede sem fio  $i \in \mathcal{K}$  é dada por

$$\kappa_{ij} = 20 \cdot \log_{10} \left( \frac{4\pi f_c d_{ij}}{c} \right) + \xi, \quad (5)$$

onde  $d_{ij}$  é a distância entre o VANT e o usuário,  $f_c$  é a frequência da portadora,  $c$  é a velocidade da luz no vácuo e  $\xi$  representa as perdas adicionais, que assumem diferentes valores de acordo com a existência ou não de linha de visada entre o usuário e o VANT (MOZAFFARI et al., 2015). Vale ressaltar que o cálculo da presença ou não de linha de visada depende da realização instantânea do modelo urbano, levando em consideração a posição tridimensional do VANT, a localização e a altura dos edifícios, assim como a posição dos usuários da rede sem fio.

A razão sinal-ruído mais interferência, do inglês *Signal to Interference Plus Noise Ratio* (SINR), para um usuário  $i$  em referência a um VANT  $j$  é descrita por

$$\varrho_{ij} = \frac{P_{r_{ij}}}{B N_0 + \sum_{k=1, k \neq j}^D P_{r_{ik}}}, \quad (6)$$

onde  $B$  é a largura de banda do sistema, em Hz,  $N_0$  é a densidade espectral de ruído, em

W/Hz, e  $P_{r_{ij}} = P_{t_j} - \kappa_{ij}$  é a potência recebida pelo usuário  $i$  quando o VANT  $j$  emprega a potência de transmissão  $P_{t_j}$ . Além disso, o somatório apresentado no denominador da equação (6) apresenta a interferência causada pelos VANTs vizinhos que transmitem ao mesmo tempo e mesma banda de frequência.

A alocação dos usuários à cada VANT é feita de acordo com as suas respectivas SINRs. Se o usuário  $i$  apresenta uma SINR  $\rho_{ij}$  em relação ao VANT  $j$ , que está acima de dado limite, e o referido VANT possui recursos suficientes para alocar esse usuário, então o usuário é alocado. Entretanto, se o VANT em questão não possui recursos disponíveis ou se a SINR encontra-se abaixo do limite pretendido, o próximo VANT é considerado como uma alternativa para a alocação do usuário em questão. Finalmente, se todos os VANTs foram testados e o usuário foi incapaz de ser alocado a qualquer um deles, esse usuário é considerado indisponível para o intervalo de tempo corrente, ou seja, fora da área de cobertura da rede sem fio

### 3.3 MODELO DE CONSUMO DE ENERGIA

No presente trabalho são considerados VANTs de hélice que, de acordo com (ZENG et al., 2016), são equipamentos capazes de mover-se facilmente e em qualquer direção, bem como manter-se estacionários no ar. Portanto, esses equipamentos apresentam características interessantes para operar como ERBs em redes de celulares.

O consumo energético relacionado à movimentação de VANTs de hélice depende, em geral, da sua velocidade de voo e de sua aceleração. Ainda assim, conforme descrito por (ZENG et al., 2019), em cenários onde as manobras dos VANTs têm pequenas durações de tempo, é razoável ignorar o consumo energético causado pela sua aceleração. Dessa forma, a potência consumida para a movimentação de um VANT de hélice com velocidade  $V$ ,  $P(V)$ , é dada por (ZENG et al., 2019)

$$P(V) = P_0 \left( 1 + \frac{3V^2}{\Omega^2 \psi^2} \right) + P_i \left( \sqrt{1 + \frac{V^4}{4v_0^4} - \frac{V^2}{2v_0^2}} \right)^{\frac{1}{2}} + \frac{1}{2} d_0 \rho s A V^3, \quad (7)$$

onde  $\Omega$  é a velocidade angular das lâminas, em radianos/segundo,  $\psi$  é o raio do rotor, em metros,  $v_0$  é a velocidade média induzida pelo rotor,  $d_0$  é a taxa de arrasto da fuselagem,  $\rho$  é a densidade do ar,  $s$  é a solidez do rotor e  $A$  é a área do rotor, em m<sup>2</sup>. Além disso,  $P_0$  é a potência do perfil da lâmina, dada por

$$P_0 = \frac{\delta}{8} \rho s A \Omega^3 R^3, \quad (8)$$

em que  $\delta$  é o coeficiente de arrasto do perfil, ao passo que  $P_i$  é a potência induzida obtida pela expressão

$$P_i = (1 + \zeta) \frac{\chi^{\frac{3}{2}}}{\sqrt{2\rho A}}, \quad (9)$$

onde  $\zeta$  e  $\chi$  são o fator de correção da potência induzida e peso do VANT, respectivamente.

Por ser consideravelmente menor que a potência empregada para mover o VANT, a potência transmitida não foi considerada no modelo de consumo de energia descrita (ZENG et al., 2019).

### 3.4 MÉTRICAS

As métricas consideradas para avaliar o algoritmo proposto foram a porcentagem dos usuários em indisponibilidade,  $U_{\text{out}}$  e a eficiência energética dos VANTs.

A porcentagem de usuários em indisponibilidade é dada por

$$U_{\text{out}} = \frac{N_{\text{out}}}{K}, \quad (10)$$

onde  $N_{\text{out}}$  é o número total de usuários indisponíveis, definido como

$$N_{\text{out}} = K - \sum_{j=1}^D U_j, \quad (11)$$

onde  $U_j$  é o número de usuários alocados ao VANT  $j$ .

Por sua vez, a eficiência energética dos VANTs é definida como

$$\eta = \frac{R_b \cdot (1 - U_{\text{out}})}{E_{\text{total}}}, \quad (12)$$

onde  $R_b$  é a taxa de bit, em bps, e  $E_{\text{total}}$  é a energia total gasta para movimentar todos os VANTs componentes da rede, considerando a soma da energia gasta para cada iteração, ou seja, todos os instantes em que os usuários são considerados estáticos durante a execução do algoritmo. Denotando  $\mathcal{I}$  como o conjunto de iterações realizadas, o consumo total de energia é dado por

$$E_{\text{total}} = \sum_{l \in \mathcal{I}} \sum_{j \in \mathcal{D}} E_{j,l}, \quad (13)$$

onde o consumo de energia para cada  $l$ -ésima iteração para um dado VANT  $j$  é definido como

$$E_{j,l} = P_{j,l} \cdot \Delta t, \quad (14)$$

onde  $P_{j,l}$  é a potência consumida para mover o VANT  $j$  na iteração corrente  $l$ , calculado a partir de (7), e  $\Delta_t = \frac{d_{j,l}}{V}$  é o tempo decorrido para realizar essa movimentação, com  $d_{j,l}$  sendo a distância percorrida pelo VANT nessa iteração com velocidade  $V$ .

### 3.5 ALGORITMO PROPOSTO

No presente trabalho o algoritmo *Q-learning* é empregado para encontrar o melhor posicionamento e a melhor potência de transmissão a ser alocada em VANTs constituintes de uma rede sem fio. A solução proposta permite que os VANTs, agentes da solução, aprendam como agir de forma autônoma, utilizando o conhecimentos de suas ações prévias (WATKINS, 1992). Para esta proposta, são necessários:

- *Agentes*: Cada VANT é um agente independente da solução *Q-learning*, capaz de mover-se em um ambiente composto por edifícios e usuários;
- *Usuários*: Pedestres, tentando se conectar a algum dos VANTs que compõe a rede;
- *Estados*: Um conjunto de estados possíveis para os agentes consiste nas posições tridimensionais e os níveis de potência de transmissão a serem implementados;
- *Ações*: Cada VANT pode tomar uma entre nove ações possíveis por episódio: Mover-se para frente, para trás, para cima, para baixo, para esquerda, direita, aumentar ou diminuir a potência transmitida ou não fazer nada.

A proposta principal do algoritmo *Q-learning* é definir uma Tabela-*Q* para cada agente independente, conforme proposto por (WATKINS, 1992), composta por elementos  $Q(s_t, a_t)$  que representam o valor em estar em um *estado* específico  $s_t$ , em um instante  $t$ , enquanto realizando uma *ação* específica  $a_t$ . A Tabela-*Q* é atualizada de acordo com a expressão

$$Q(s_t, a_t) = Q(s_t, a_t) + \lambda \left[ r_{k_{t+1}} + \phi \max_a \{Q(s_{t+1}, a)\} - Q(s_t, a_t) \right], \quad (15)$$

onde  $\lambda$  é a taxa de aprendizagem,  $r_{k_{t+1}}$  é a recompensa esperada para o próximo instante, com  $k \in \{1, 2\}$  denotando uma dentre as duas recompensas propostas,  $\phi$  é o fator de desconto e  $\max_a \{Q(s_{t+1}, a)\}$  é a função que estima o valor ótimo para a ação do instante futuro. Então, obtém-se as seguintes definições

- *Política*: A política aplicada para definir o comportamento dos agentes é denominada  *$\epsilon$ -greedy*, em que o parâmetro  $\epsilon$  define quando o agente da solução

$Q$ -learning deve assumir uma atitude de *explore*, em que as ações aleatórias realizadas em diferentes estados são priorizadas, ou *exploit*, quando o agente utiliza o conhecimento adquirido por suas ações para buscar a melhor ação possível dentro estado corrente (SUTTON et al., 1998);

- *Inicialização*: Os VANTs são posicionados aleatoriamente e cada Tabela- $Q$  individual é inicializada em zero;
- *Episódios*: A otimização proposta é dividida em episódios e cada episódio é também dividido em iterações. Então, os VANTs são capazes de mover-se testando algumas das ações possíveis e observando a recompensa resultante;
- *Crítérios de parada*: As ações são realizadas até que algum dentre os critérios de parada seja atingido, esses são: (i.) o número de usuários cobertos pela rede sem fio não ter melhorado em um certo número de iterações ( $\iota_{\text{out,max}}$ ); (ii.) o VANT ter se movimentado por um número máximo de iterações ( $\iota_{\text{max}}$ ); (iii.) o VANT ter esgotado todos os recursos disponíveis. Quando o VANT atingir uma dessas três condições, ele se movimenta para o estado que resultou na maior recompensa, então todos os VANTs são otimizados de forma sequencial, até que o próximo episódio se inicie.

A solução proposta é sumarizada pelo Algoritmo 1.

---

**Algoritmo 1:** Algoritmo  $Q$ -learning proposto.

---

```

1 Inicializa os VANTs em posições aleatórias;
2 Inicializa tabelas- $Q$  com todos os elementos nulos;
3 para Todo episódio faça
4   enquanto Parâmetros de parada não são atingidos faça
5     VANTs selecionam ações com maior  $Q(s_t, a_t)$ , movendo para o
6       respectivo estado  $s_t$  e realizando a ação  $a_t$ ;
7     VANTs alocam os usuários e observam  $r_{k_{t+1}}$ ;
8     Cada tabela- $Q$  é atualizada de acordo com (15);
9   fim
10 fim

```

---

O objetivo dos agentes da solução nas possíveis ações é aumentar a recompensa da solução  $Q$ -learning. Nesse estudo, são propostas duas recompensas com diferentes objetivos, conforme a seguir.



### 3.5.1 Recompensa dedicada à cobertura da rede

Para obter a maximização do número de usuários cobertos pela rede sem fio proposta, a primeira recompensa foi definida como:

$$r_1 = \sum_{j=1}^D U_j, \quad (16)$$

recordando que  $U_j$  representa o número de usuários conectados a cada VANT  $j$ . Esse objetivo dedica-se apenas ao número de usuários cobertos pela rede e, portanto, leva a um algoritmo que visa minimizar  $U_{\text{out}}$  em (10). Vale ressaltar que  $r_1$  é o mesmo para todos os VANTs e é calculado localmente, considerando a conexão de *backhaul* entre os VANTs.

### 3.5.2 Recompensa dedicada à cobertura da rede e consumo energético

A segunda recompensa apresenta uma generalização de  $r_1$ , definida como

$$r_2 = \sum_{j=1}^D U_j - w \cdot E_{j,l}, \quad (17)$$

que também é calculado localmente, mas é independente entre um VANT para outro. Além disso,  $w$  em (17) é o peso dado para a energia consumida nessa recompensa. Vale ressaltar que essa recompensa apresenta um *trade off* entre a cobertura da rede sem fio e a energia consumida para movimentar os VANTs, agentes da solução. Portanto, a seleção apropriada do fator  $w$  é importante para minimizar  $U_{\text{out}}$  com uma penalidade adequada para cada movimentação dos VANTs.

## 4 RESULTADOS E DISCUSSÕES

O cenário simulado consiste em uma área urbana de  $250.000 \text{ m}^2$  e  $K = 200$  usuários. Na simulação realizada as posições desses usuários foram aleatoriamente iniciadas e 100 rodadas de 100 episódios foram executadas de maneira independente. Os parâmetros da simulação são apresentados na Tabela 2.

**Tabela 2 – Parâmetros de simulação**

Parâmetro	Valor
Área	500 m $\times$ 500 m
Número de usuários ( $K$ )	200 usuários
Capacidade dos VANTs	50 usuários
Passo do VANT (eixos X e Y)	50 m
Passo do VANT (eixo Z)	100 m
Altura do VANT	[100, 1000] m
$\alpha$	0.3 (ITU-R, 2003)
$\beta$	500 Edifícios/km <sup>2</sup> (ITU-R, 2003)
$\gamma$	15 m (ITU-R, 2003)
Perda Adicional ( $\xi$ ) LOS / NLOS	1 dB / 20 dB (AL-HOURANI et al., 2014)
Ângulo de abertura das antenas ( $\theta$ )	60°
Largura de banda ( $B$ )	180 kHz
Frequência portadora ( $f_c$ )	1 GHz
Taxa de Bit ( $R_b$ )	360 Kbps
Número máximo de iterações ( $\iota_{\max}$ )	3600
Máx. Iter. sem melhora na cobertura ( $\iota_{\text{out},\max}$ )	20
Taxa de aprendizagem ( $\lambda$ )	0,9
Fator de desconto ( $\phi$ )	0,9
Fator entre <i>exploit</i> e <i>explore</i> ( $\epsilon$ )	0,5
Velocidade do VANT ( $V$ )	5 m/s
Peso do VANT ( $\chi$ )	20 N
Densidade do Ar ( $\rho$ )	1,225 kg/m <sup>3</sup> (ZENG et al., 2019)
Raio do Rotor ( $R$ )	0,4 m (ZENG et al., 2019)
Área do disco do Rotor ( $A$ )	0,503 m <sup>2</sup> (ZENG et al., 2019)
Velocidade angular da lâmina ( $\Omega$ )	300 rad/s (ZENG et al., 2019)
Velocidade média induzida pelo rotor ( $v_0$ )	4,03 m/s (ZENG et al., 2019)
Taxa de arrasto da fuselagem ( $d_0$ )	0,6 (ZENG et al., 2019)
Solidez do rotor ( $s$ )	0,05 (ZENG et al., 2019)
Potência do perfil da lâmina ( $\delta$ )	0,012 (ZENG et al., 2019)
Fator de correção da potência induzida ( $k$ )	0,1 (ZENG et al., 2019)

**Fonte: Autoria própria.**

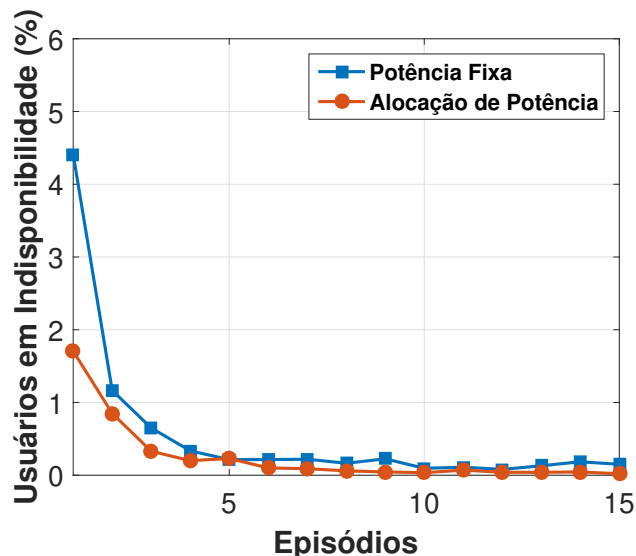
Estendendo (PARISOTTO et al., 2019), é considerado que os VANTs podem agir para escolher a potência transmitida dentre três possibilidades:  $-20 \text{ dB}$ ,  $-10 \text{ dB}$  e

0 dB, o que pode melhorar o número de usuários alocados aos VANTs devido a redução da interferência entre os VANTs quando comparada ao caso com potência de transmissão fixa. A seguir, a seção 4.1 compara os resultados implementando a recompensa apresentada por  $r_1$  na equação (16) com o algoritmo *Q-learning* proposto por (KLAINÉ et al., 2018), onde a potência transmitida é fixa em 0 dB. Na sequência, a seção 4.2 é focada na eficiência energética obtida pelas diferentes recompensas  $r_1$  e  $r_2$ , apresentada em (17), e seus respectivos impactos no número de usuários alocados à rede sem fio proposta.

#### 4.1 MAXIMIZAÇÃO DE USUÁRIOS COBERTOS PELA REDE SEM FIO

A evolução da porcentagem de usuários em situação de indisponibilidade de acordo com a passagem do número de episódios decorridos é apresentada na Figura 5 em um cenário com  $D = 7$  VANTs. A solução proposta por (KLAINÉ et al., 2018), onde uma potência de transmissão fixa de 0 dB é empregada por todos os VANTs é comparada com a metodologia onde há possibilidade dos agentes da solução assumirem diferentes níveis de potência de transmissão. Como pode ser observado, ambas soluções convergem rapidamente, e o método de alocação de potência proposto apresenta uma vantagem. A partir desse resultado, em ordem a demonstrar apenas o desempenho das soluções em estado estacionário, os resultados numéricos seguintes descartam os primeiros 10 episódios executados, apresentando os resultados como uma média iniciada ao 11º episódio.

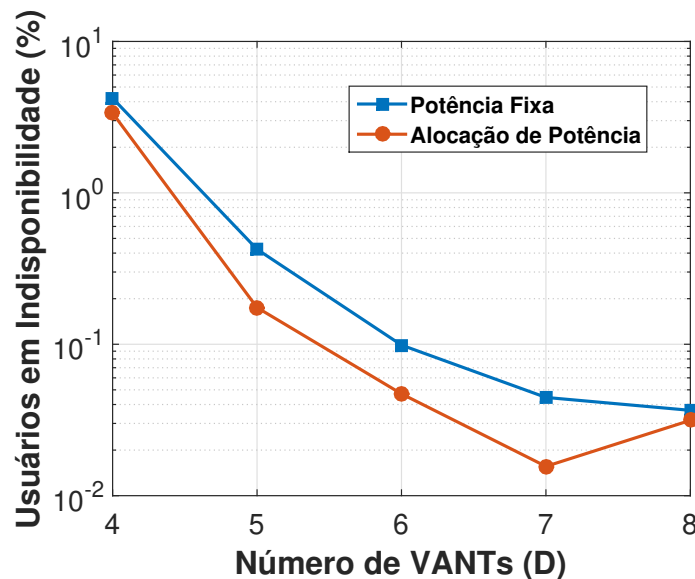
**Figura 5 – Porcentagem de usuários indisponíveis em função do número de episódios executados em um cenário com  $D = 7$  VANTs.**



Fonte: Autoria própria.

A Figura 6 apresenta a porcentagem de usuários em indisponibilidade em função do número de VANTs, quando o método proposto emprega  $r_1$  como recompensa. Os resultados são calculados como a média de 100 rodadas simuladas. Como pode ser observado, a possibilidade dos agentes alocarem diferentes níveis de potência de transmissão pelos agentes da solução melhora consideravelmente o seu desempenho. Esse decremento no número de usuários em situação de indisponibilidade é resultado da redução da interferência entre VANTs vizinhos, uma vez que a potência transmitida média é reduzida. Por exemplo, quando  $D = 7$  VANTs são empregados, a porcentagem total de usuários indisponíveis é reduzida em 3,3% entre o método com alocação de potência e potência fixa, atingindo um total de 98,5% de usuários cobertos pela rede para a aplicação com diferentes níveis de potência de transmissão.

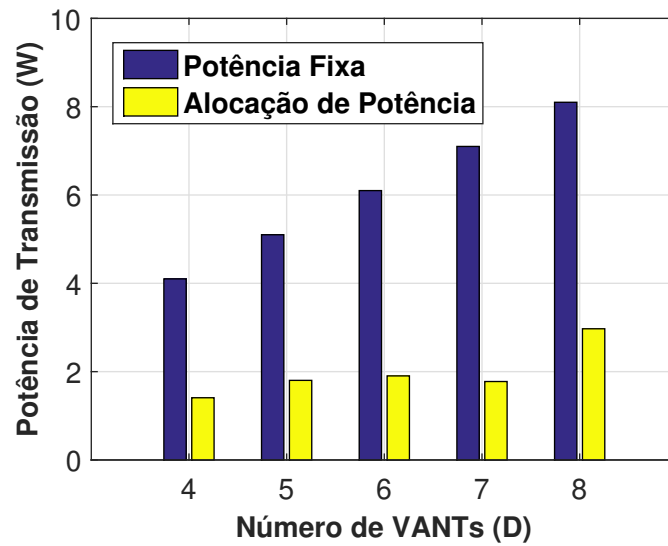
**Figura 6 – Usuários indisponíveis em função do número de VANTs ( $D$ ).**



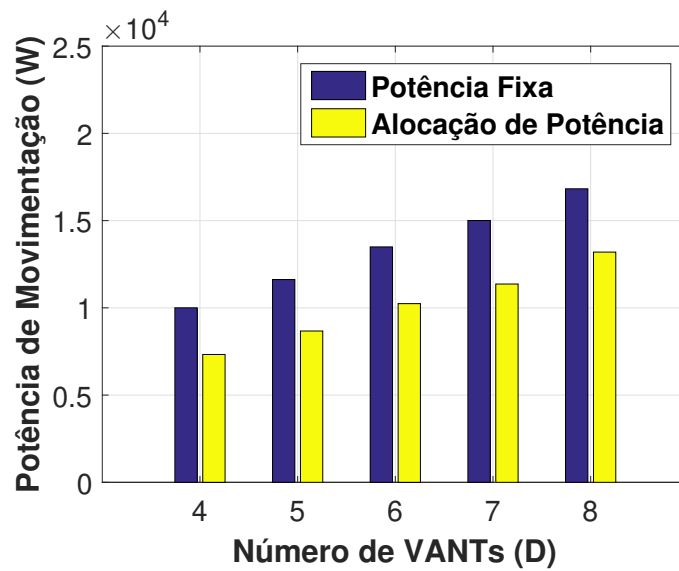
**Fonte:** Autoria própria.

A Figura 7 apresenta o consumo total de potência devido à transmissão de dados (Figura 7a) e devido à movimentação dos VANTs (Figura 7b), considerando a recompensa  $r_1$ . Como pode ser observado, a alocação de potência para os VANTs resulta em uma redução da potência média consumida para ambas as situações. Entretanto, comparando as duas subfiguras, também pode ser observado que o consumo de potência para a movimentação dos VANTs têm magnitude várias ordens de grandeza maior que o consumo devido à transmissão de potência, justificando o foco da recompensa  $r_2$  em encontrar a solução mais energeticamente eficiente, que será considerada na subseção seguinte.

Figura 7 – Potência consumida pelos VANTs.



(a) Potência de transmissão.



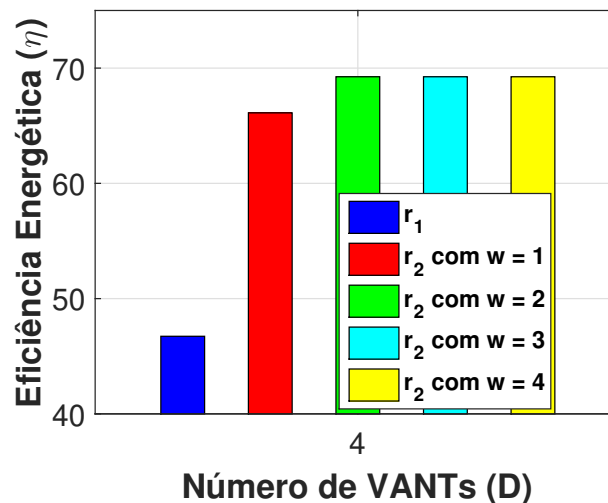
(b) Potência de movimentação.

Fonte: Autoria própria.

## 4.2 MAXIMIZAÇÃO DE USUÁRIOS COBERTOS PELA REDE COM RESTRIÇÕES EM CONSUMO DE ENERGIA

Nessa seção as recompensas  $r_1$  e  $r_2$  são comparadas a fim de observar os *trade-offs* em termos de porcentagem de usuários indisponíveis e eficiência energética. Primeiramente, a Figura 8 apresenta a eficiência em um cenário com  $D = 4$  VANTs, onde a recompensa  $r_2$  é aplicada com  $w \in \{1; 1,5; 2; 2,5; 3\}$ . Como pode ser observado,  $w = 2$  é o peso que alcança o melhor resultado em termos de eficiência energética e a EE é saturada a partir de  $w = 2$ , portanto, para os resultados as análises são feitas considerando  $w = 2$  é escolhido. Para  $w > 2$ , a eficiência energética satura ao mesmo valor, indicando que os VANTs estão viajando nas menores distâncias possíveis nessas situações. Essas tendências na análise do peso  $w$  são confirmadas para diferentes quantidades de VANTs  $D$  e, baseado nisso, para os resultados a seguir é considerado  $r_2$  com  $w = 2$  como o valor padrão para esse cenário.

**Figura 8 – EE para diferentes pesos  $w$  empregados à recompensa  $r_2$ .**

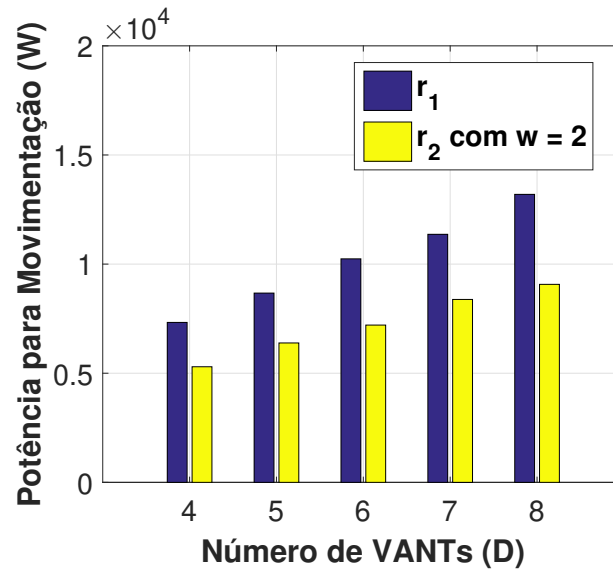


**Fonte: Autoria própria.**

A Figura 9 apresenta o consumo de potência devido às movimentações dos VANTs considerando as recompensas  $r_1$  e  $r_2$  como uma função do número de VANTs ( $D$ ). Como pode ser observado, a recompensa proposta consome menos potência para todos os valores de  $D$  considerados, alcançando uma economia de 31,25% de potência quando  $r_2$  com  $w = 2$  é comparado a  $r_1$  em uma rede sem fio formada por  $D = 8$  VANTs. De forma complementar, a Figura 10 ilustra o número médio de movimentações dos VANTs, que é

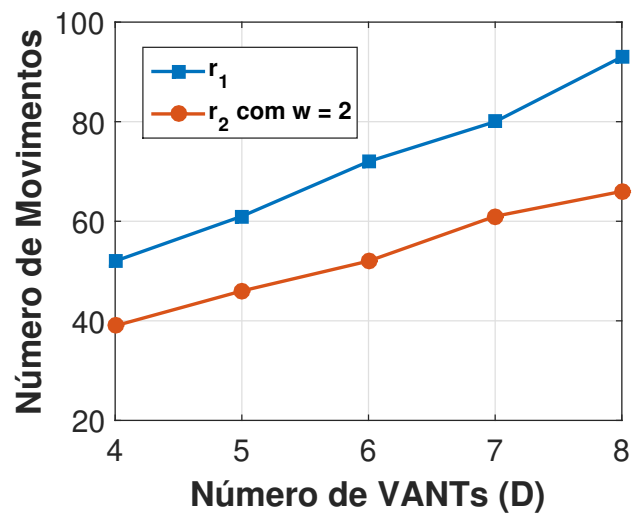
consideravelmente decrescido quando a recompensa  $r_2$  aplicado com  $w = 2$  é comparada a  $r_1$ .

Figura 9 – Consumo de potência devido às movimentações dos VANTs para a recompensas  $r_1$  e  $r_2$  em função do número de VANTs ( $D$ ).



Fonte: Autoria própria.

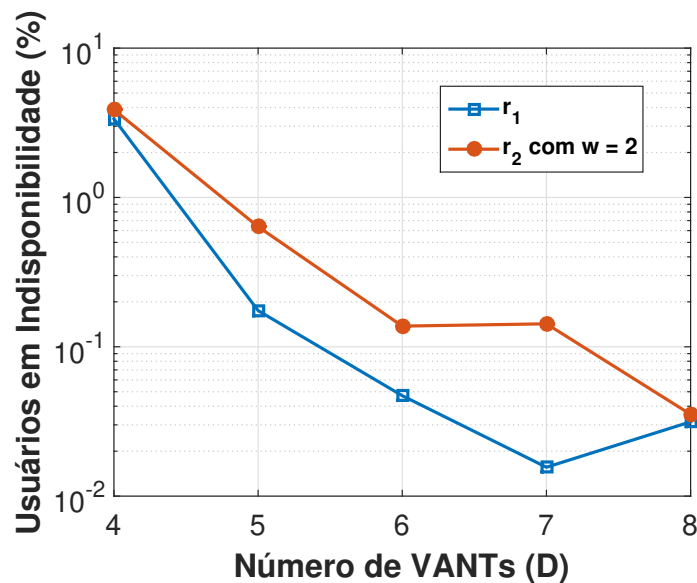
Figura 10 – Número médio de movimentações dos VANTs para as recompensas propostas.



Fonte: Autoria própria.

Por outro lado, a Figura 11 apresenta a média dos usuários em situação de indisponibilidade em função do número de VANTs. Como ilustrado, a penalidade trazida por  $w$  em termos da movimentação dos VANTs afeta a cobertura da rede sem fio, uma vez que a porcentagem de usuários em indisponibilidade aumenta. Comparando  $r_2$  com  $w = 2$  e  $r_1$ ,  $U_{\text{out}}$  é incrementado em 12,69% quando  $D = 7$  VANTs constituem a rede. Ainda assim, a porcentagem de usuários fora da cobertura da rede formada por  $D = 7$  VANTs para o  $Q$ -learning com  $r_2$  e  $w = 2$  alcança 0,14%, o que justifica a economia em termos das movimentações dos VANTs.

**Figura 11 – Porcentagem de usuários em indisponibilidade em função de  $D$  para as recompensas propostas.**

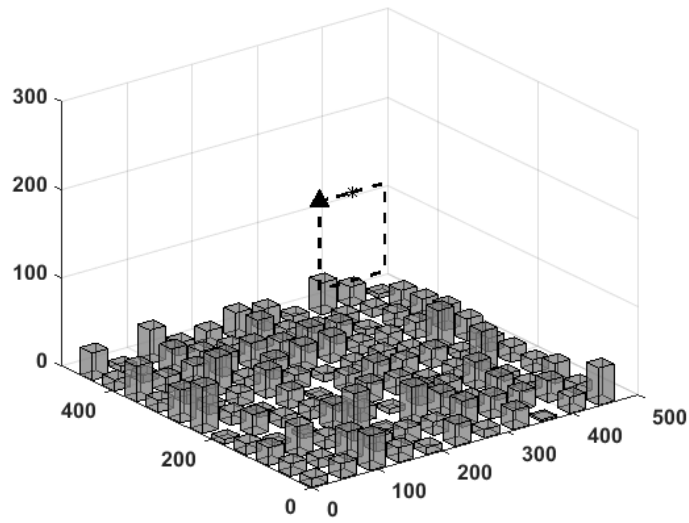


**Fonte: Autoria própria.**

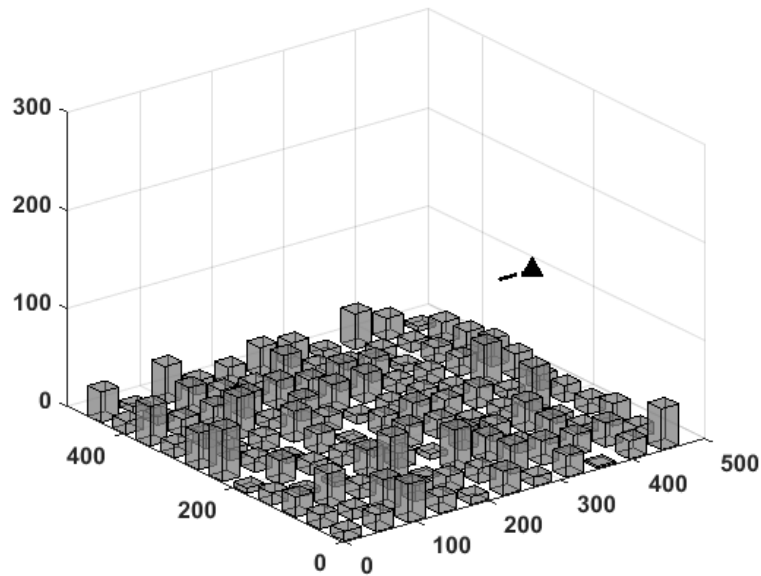
A trajetória do 7º VANT durante um episódio executado é apresentada na Figura 12, com cenários de aplicação da recompensa  $r_1$  (Figura 12a) e  $r_2$  com  $w = 2$  (Figura 12b), que ilustram como o algoritmo com restrições no consumo energético reduz consideravelmente a movimentação do VANT, demonstrando um aumento na eficiência energética da solução. Isso se reflete na Figura 13, a qual apresenta a eficiência energética considerando a recompensa  $r_1$  e  $r_2$  com  $w = 2$ . É observado que, por exemplo, uma rede com  $D = 8$  VANTs obtém uma melhora de desempenho de 36,91% comparando  $r_2$  com a recompensa  $r_1$ . Esse resultado evidencia os ganhos obtidos em função do parâmetro  $w$  em conjunto com a solução  $Q$ -learning para a economia de energia.



Figura 12 – Trajetória feita pelo 7º VANT durante um episódio específico com as recompensas  $r_1$  e  $r_2$ .



(a) Recompensa  $r_1$ .



(b) Recompensa  $r_2$  com  $w = 2$ .

Fonte: Autoria própria.

Por fim, o tempo de vida da rede é considerado assumindo o consumo de bateria dos VANTs durante a execução do algoritmo. Seguindo (GALKIN et al., 2019), a bateria inicial dos VANTs é considerada 100 Wh. Então, o consumo energético a cada movimentação dos VANTs é decrementado do valor inicial de bateria até que o critério de parada, nessa análise determinado como o primeiro VANT a atingir 10% da carga inicial da bateria. Nessa situação esse VANT deve retornar à base para recarregar sua bateria

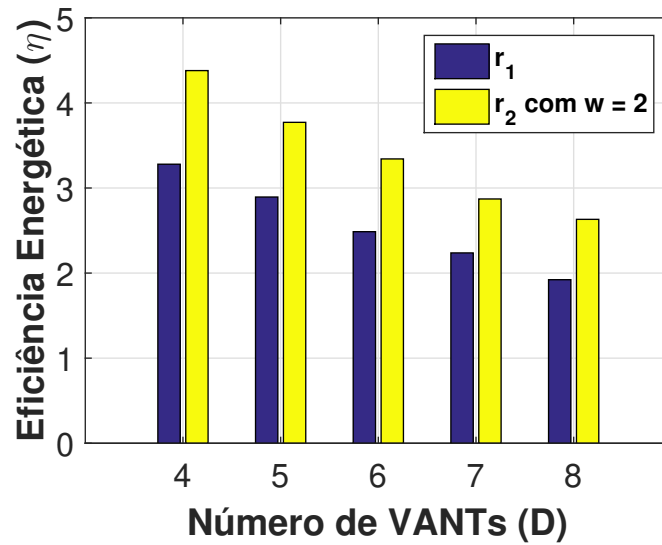
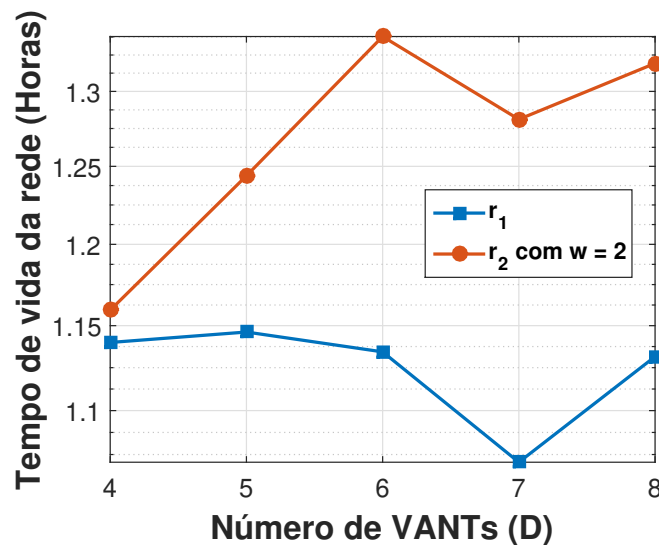


Figura 13 – EE para as recompensas propostas.

e o tempo de vida da rede é computado nesse instante de tempo. A Figura 14 ilustra essa análise considerando as recompensas  $r_1$  e  $r_2$  com  $w = 2$ . Como pode ser observado, a implementação da recompensa  $r_2$  aumenta em 19,66% o tempo de vida da rede sem fio quando  $D = 7$  VANTs são empregados, o que pode ser considerado um ganho considerável em termos de autonomia da rede.

Figura 14 – Tempo de vida da rede em função do número de VANTs.



Fonte: Autoria própria.

## 5 CONCLUSÕES

Técnicas para o posicionamento de VANTs em redes sem fio podem resultar no aumento da cobertura e da autonomia da rede em diversos cenários de aplicação. Nesse contexto, propõe-se uma solução baseada no método *Q-learning* de aprendizagem por reforço para capacitar VANTs, agentes da solução, a agir de forma autônoma a partir do conhecimento dos resultados de suas ações anteriores para encontrar o melhor posicionamento 3D e a melhor potência de transmissão em uma rede sem fio aplicada ao cenário de eventos temporários.

O cenário considerado contempla um modelo urbano de distribuição de usuários e edifícios e um modelo consumo de energia para VANTs de hélice para comparação de duas diferentes propostas da solução *Q-learning*: uma com recompensa dedicada exclusivamente à cobertura de usuários e outra com recompensa multi-objetivo, onde número de usuários em situação de indisponibilidade e consumo energético devido à movimentação dos VANTs são considerados.

Com uma avaliação de desempenho feita a partir das métricas de porcentagem de usuários em indisponibilidade e eficiência energética da rede, os resultados apresentados mostram que, alcançando a cobertura de 99,86% dos usuários, a proposta com penalidade atribuída pela restrição na movimentação dos VANTs resultou em uma melhora do desempenho em EE de até 36,91% e de uma maximização de até 19,66% no tempo de vida da rede se comparada à solução com recompensa exclusiva à cobertura da rede, resultados que evidenciam o potencial da técnica *Q-learning* para aplicações nos cenários de futuras gerações de redes celulares.

Para trabalhos futuros, sugere-se a extensão das requisições dos usuários da rede, considerando restrições em relação à velocidade e latência, por exemplo, de forma a tornar o cenário mais próximo de uma aplicação real. Além disso, a ampliação do número de ações a serem tomadas pela rede, como a mudança de abertura das antenas e da frequência de transmissão, pode ter um efeito na diminuição da interferência entre os agentes da rede, efeito capaz de diminuir a necessidade de movimentações dos VANTs em altura e, conseqüentemente, reduzir o consumo de energia. De forma complementar, cenários futuros também podem considerar a presença de múltiplas antenas, fato que, apesar de resultar em um aumento de consumo, poderá resultar na melhora na interferência e cobertura dos usuários.

Outra análise poderá considerar a implementação de diferentes técnicas de aprendizado de máquina, a exemplo do algoritmo SARSA, do inglês *state-action-reward-state-action* onde, diferentemente do *Q-learning*, utiliza-se a mesma política para gerar o comportamento dos agentes e para avaliar as ações tomadas (SUTTON et al., 1998). Juntamente à técnica VFA, do inglês *Value Function Approximation*, a técnica SARSA pode ter sua viabilidade estudada para a aplicação apresentada nesse trabalho, de forma a compará-la com a técnica *Q-learning*, uma vez que a técnica VFA apresenta-se como capaz de evitar a limitação de memória sistema (SUTTON et al., 1998).

## REFERÊNCIAS

- AL-HOURANI, A.; KANDEEPAN, S.; JAMALIPOUR, A. Modeling air-to-ground path loss for low altitude platforms in urban environments. In: **IEEE Global Communications Conference (GLOBECOM)**. Austin, USA, 2014. p. 2898–2904.
- ALIU, O. G.; IMRAN, A.; IMRAN, M. A.; EVANS, B. A survey of self organisation in future cellular networks. **IEEE Communications Surveys & Tutorials**, IEEE, v. 15, n. 1, p. 336–361, 2013.
- ALSKAIF, T.; ZAPATA, M. G.; BELLALTA, B. Game theory for energy efficiency in wireless sensor networks: Latest trends. **Journal of Network and Computer Applications**, Elsevier, v. 54, p. 33–61, 2015.
- ALZENAD, M.; EL-KEYI, A.; LAGUM, F.; YANIKOMEROGLU, H. 3-D placement of an unmanned aerial vehicle base station (UAV-BS) for energy-efficient maximal coverage. **IEEE Wireless Communications Letters**, v. 6, n. 4, p. 434–437, 2017.
- AUER, G.; GIANNINI, V.; DESSET, C.; GODOR, I.; SKILLERMARK, P.; OLSSON, M.; IMRAN, M.; SABELLA, D.; GONZALEZ, M.; BLUME, O.; FEHSKE, A. How much energy is needed to run a wireless network? **IEEE Wireless Communications**, v. 18, n. 5, p. 40–49, Oct. 2011. ISSN 1536-1284.
- BAYERLEIN, H.; KERRET, P. D.; GESBERT, D. Trajectory optimization for autonomous flying base station via reinforcement learning. In: **IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)**. Kalamata, Greece, 2018. p. 1–5.
- BUZZI, S.; I, C.; KLEIN, T. E.; POOR, H. V.; YANG, C.; ZAPPONE, A. A survey of energy-efficient techniques for 5G networks and challenges ahead. **IEEE Journal on Selected Areas in Communications**, v. 34, n. 4, p. 697–709, April 2016. ISSN 0733-8716.
- CHALLITA, U.; SAAD, W.; BETTSTETTER, C. Interference management for cellular-connected UAVs: A deep reinforcement learning approach. **IEEE Transactions on Wireless Communications**, v. 18, n. 4, p. 2125–2140, April 2019.
- CISCO. Cisco visual networking index: Forecast and methodology, 2016 - 2021. **White Paper**, 2017.
- CUI, S.; GOLDSMITH, A. J.; BAHAI, A. Energy-efficiency of MIMO and cooperative MIMO techniques in sensor networks. **IEEE Journal on selected areas in communications**, v. 22, n. 6, p. 1089–1098, 2004.
- CUI, S.; GOLDSMITH, A. J.; BAHAI, A. Energy-constrained modulation optimization. **IEEE Transactions on Wireless Communications**, IEEE, v. 4, n. 5, p. 2349–2360, 2005.

- DANDANOV, N.; AL-SHATRI, H.; KLEIN, A.; POULKOV, V. Dynamic self-optimization of the antenna tilt for best trade-off between coverage and capacity in mobile networks. **Wireless Personal Communications**, Springer, v. 92, n. 1, p. 251–278, 2017.
- FEHSKE, A. J.; RICHTER, F.; FETTWEIS, G. P. Energy efficiency improvements through micro sites in cellular mobile radio networks. In: **IEEE GLOBECOM Workshops**. 2009. p. 1–5.
- FENG, D.; JIANG, C.; LIM, G.; CIMINI, L. J.; FENG, G.; LI, G. Y. A survey of energy-efficient wireless communications. **IEEE Communications Surveys & Tutorials**, IEEE, v. 15, n. 1, p. 167–178, 2013.
- GALKIN, B.; KIBILDA, J.; SILVA, L. A. UAVs as mobile infrastructure: Addressing battery lifetime. **IEEE Communications Magazine**, v. 57, n. 6, p. 132–137, June 2019.
- GANDOTRA, P.; JHA, R. K.; JAIN, S. Green communication in next generation cellular networks: a survey. **IEEE Access**, IEEE, v. 5, p. 11727–11758, 2017.
- GUPTA, L.; JAIN, R.; VASZKUN, G. Survey of important issues in UAV communication networks. **IEEE Communications Surveys Tutorials**, v. 18, n. 2, p. 1123–1152, 2016.
- HUAWEI. 5G: A technology vision. **White Paper**, 2013.
- IMRAN, M. A.; REIS, A. F. dos; BRANTE, G.; KLAINÉ, P. V.; SOUZA, R. D. Machine learning in energy efficiency optimization. **Machine Learning for Future Wireless Communications**, Wiley Online Library, p. 105–117, 2019.
- ITU-R. Propagation Data and Prediction Methods for The Design of Terrestrial Broadband Millimetric Radio Access Systems. p. 1410–2, 2003.
- KIRAN, P.; JIBUKUMAR, M.; PREMKUMAR, C. Resource allocation optimization in LTE-A/5G networks using big data analytics. In: **International Conference on Information Networking (ICOIN)**. 2016. p. 254–259.
- KLAINÉ, P. V.; IMRAN, M. A.; ONIRETI, O.; SOUZA, R. D. A survey of machine learning techniques applied to self organizing cellular networks. **IEEE Communications Surveys and Tutorials**, IEEE, 2017.
- KLAINÉ, P. V.; IMRAN, M. A.; ONIRETI, O.; SOUZA, R. D. A survey of machine learning techniques applied to self-organizing cellular networks. **IEEE Communications Surveys & Tutorials**, v. 19, n. 4, p. 2392–2431, 2017.
- KLAINÉ, P. V.; NADAS, J. P.; SOUZA, R. D.; IMRAN, M. A. Distributed drone base station positioning for emergency cellular networks using reinforcement learning. **Cognitive computation**, Springer, v. 10, n. 5, p. 790–804, 2018.
- MERWADAY, A.; GUVENC, I. UAV assisted heterogeneous networks for public safety communications. In: **IEEE wireless communications and networking conference workshops (WCNCW)**. New Orleans, USA, 2015. p. 329–334.

- MIOZZO, M.; GIUPPONI, L.; ROSSI, M.; DINI, P. Switch-on/off policies for energy harvesting small cells through distributed Q-learning. In: **IEEE Wireless Communications and Networking Conference Workshops (WCNCW)**. 2017. p. 1–6.
- MOZAFFARI, M.; SAAD, W.; BENNIS, M.; DEBBAH, M. Drone small cells in the clouds: Design, deployment and performance analysis. In: **IEEE Global Communications Conference (GLOBECOM)**. San Diego, USA, 2015. p. 1–6.
- ORO, S. D.; ZAPPONE, A.; PALAZZO, S.; LOPS, M. A learning approach for low-complexity optimization of energy efficiency in multicarrier wireless networks. **IEEE Transactions on Wireless Communications**, IEEE, v. 17, n. 5, p. 3226–3241, 2018.
- PAPOULIS, A.; PILLAI, S. U. **Probability, random variables, and stochastic processes**. Tata McGraw-Hill Education, 2002.
- PARISOTTO, R. de P.; KLAINÉ, P. V.; NADAS, J. P. B.; SOUZA, R. D.; BRANTE, G.; IMRAN, M. Drone base station positioning and power allocation using reinforcement learning. In: **16th International Symposium on Wireless Communication Systems**. Oulu, Finland, 2019.
- PATRA, R. K.; MUTHUCHIDAMBARANATHAN, P. Optimisation of spectrum and energy efficiency in UAV-enabled mobile relaying using bisection and PSO method. In: **3rd International Conference for Convergence in Technology (I2CT)**. Pune, India, 2018. p. 1–7.
- PELEKANOU, A.; ANASTASOPOULOS, M.; TZANAKAKI, A.; SIMEONIDOU, D. Provisioning of 5G services employing machine learning techniques. In: **International Conference on Optical Network Design and Modeling (ONDM)**. 2018. p. 200–205.
- PEREIRA, Z. C.; TON, T. H.; REBELATTO, J. L.; SOUZA, R. D.; UCHÔA-FILHO, B. F. Generalized network-coded cooperation in OFDMA communications. **IEEE Access**, IEEE, v. 6, p. 6550–6559, 2018.
- REIS, A. F.; BRANTE, G.; PARISOTTO, R. P.; SOUZA, R. D. Consumo energético no posicionamento de estações rádio-base móveis baseadas em VANTs aplicando aprendizagem por reforço. In: **XXXVII Simpósio Brasileiro de Telecomunicações e Processamento de Sinais (SBrT 2019)**. Petrópolis, 2019.
- RICHTER, F.; FEHSKE, A. J.; FETTWEIS, G. P. Energy efficiency aspects of base station deployment strategies for cellular networks. In: **IEEE Vehicular Technology Conference Fall (VTC 2009-Fall)**. 2009. p. 1–5.
- SINCLAIR, N.; HARLE, D.; GLOVER, I. A.; IRVINE, J.; ATKINSON, R. C. An advanced SOM algorithm applied to handover management within LTE. **IEEE Transactions on Vehicular Technology**, IEEE, v. 62, n. 5, p. 1883–1894, 2013.
- SUTTON, R. S.; BARTO, A. G. et al. **Introduction to reinforcement learning**. MIT press Cambridge, 1998.

- TULLBERG, H.; POPOVSKI, P.; LI, Z.; UUSITALO, M. A.; HOGLUND, A.; BULAKCI, O.; FALLGREN, M.; MONSERRAT, J. F. The METIS 5G system concept: Meeting the 5G requirements. **IEEE Communications Magazine**, v. 54, n. 12, p. 132–139, Dec. 2016. ISSN 0163-6804.
- WANG, Q.; ZHANG, W.; LIU, Y. Multi-UAV dynamic wireless networking with deep reinforcement learning. **IEEE Communications Letters**, p. 1–4, September 2019.
- WATKINS, C. J. C. H. Q-Learning. v. 292, p. 279–292, 1992.
- XU, J.; QIU, L. Energy efficiency optimization for MIMO broadcast channels. **IEEE Transactions on Wireless Communications**, IEEE, v. 12, n. 2, p. 690–701, 2013.
- YILMAZ, O. N.; HAMALAINEN, S.; HAMALAINEN, J. Comparison of remote electrical and mechanical antenna downtilt performance for 3GPP LTE. In: **IEEE Vehicular Technology Conference Fall (VTC 2009-Fall)**. 2009. p. 1–5.
- ZAPPONE, A.; DEBBAH, M.; ALTMAN, Z. Online energy-efficient power control in wireless networks by deep neural networks. In: **IEEE 19th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)**. 2018. p. 1–5.
- ZAPPONE, A.; JORSWIECK, E. et al. Energy efficiency in wireless networks via fractional programming theory. **Foundations and Trends in Communications and Information Theory**, Now Publishers, Inc., v. 11, n. 3-4, p. 185–396, 2015.
- ZENG, Y.; XU, J.; ZHANG, R. Energy minimization for wireless communication with rotary-wing UAV. **IEEE Transactions on Wireless Communications**, v. 18, n. 4, p. 2329–2345, 2019.
- ZENG, Y.; ZHANG, R. Energy-efficient UAV communication with trajectory optimization. **IEEE Transactions on Wireless Communications**, v. 16, n. 6, p. 3747–3760, 2017.
- ZENG, Y.; ZHANG, R.; LIM, T. J. Wireless communications with unmanned aerial vehicles: Opportunities and challenges. **IEEE Communications Magazine**, IEEE, v. 54, n. 5, p. 36–42, 2016.
- ZHAN, C.; LAI, H. Energy minimization in internet-of-things system based on rotary-wing UAV. **IEEE Wireless Communications Letters**, v. 8, n. 5, p. 1341–1344, October 2019.
- ZHENG, K.; YANG, Z.; ZHANG, K.; CHATZIMISIOS, P.; YANG, K.; XIANG, W. Big data-driven optimization for mobile networks toward 5G. **IEEE Network**, v. 30, n. 1, p. 44–51, 2016.