

Consumo Energético no Posicionamento de Estações Rádio-Base Móveis Baseadas em VANTs Aplicando Aprendizagem por Reforço

Ana Flávia dos Reis, Glauber Brante, Rafaela Parisotto e Richard Demo Souza

Resumo— Veículos aéreos não tripulados (VANTs) apresentam como vantagens a capacidade de adaptação, flexibilidade e possibilidade de auto-organização, podendo atuar como estações rádio-base para aumentar cobertura em redes sem fio. Sob essa perspectiva, o presente artigo emprega algoritmos de posicionamento baseados em *Q-learning*, uma técnica de aprendizagem por reforço capaz de otimizar o posicionamento dos VANTs, com objetivo de maximizar a cobertura da rede. Com isso, analisamos o consumo energético do *Q-learning* combinado com a alocação de potência de transmissão nos VANTs. Os resultados evidenciam o aumento de cobertura através da alocação de potência, bem como o menor consumo de energia.

Palavras-Chave— Veículos aéreos não tripulados, Posicionamento ótimo, Consumo energético, *Q-learning*.

Abstract— Unmanned aerial vehicles (UAVs) present as advantages the adaptability, flexibility and possibility of self-organization, being able to act as mobile base stations in order to increase coverage in wireless networks. In this perspective, this paper employs *Q-learning*-based positioning algorithms, a reinforcement learning technique capable of optimizing the positioning of the UAVs, in order to maximize network coverage. Thus, we analyzed the energy consumption of the *Q-learning* combined with the allocation of the transmission power in the UAVs. The results show the increased coverage obtained through power allocation, as well as lower energy consumption.

Keywords— Unmanned aerial vehicles, Optimum positioning, Power consumption, *Q-learning*.

I. INTRODUÇÃO

As novas gerações de sistemas de comunicações sem fio podem, como uma de suas tecnologias para aumentar a cobertura das redes, utilizar veículos aéreos não tripulados (VANTs) como estações rádio-base (ERBs) móveis. Por serem, em geral, de rápida implementação e apresentarem possibilidade de reconfiguração, estes equipamentos estão se tornando uma solução para prover comunicação sem fio de maneira eficiente, *e.g.*, em caso de falhas na comunicação terrestre. Por exemplo, em [1] propõe-se a aplicação de VANTs como ERBs móveis que melhorem a comunicação através de um aumento da cobertura em eventos temporários, como eventos esportivos e outros eventos ao ar livre. De forma similar, em [2] são citadas aplicações de VANTs em cenários de desastre: o monitoramento, a previsão, a avaliação de danos, o

compartilhamento de informações e, dentre outros, o suporte para a evacuação de sobreviventes. Dadas estas aplicações, outros benefícios do uso de VANTs são evidenciados pela eliminação de torres e cabos, resultando em uma economia na infraestrutura de comunicação [3].

Nesse contexto, o posicionamento dos VANTs é uma questão fundamental, pois pode-se encontrar a posição que maximiza a cobertura da comunicação. O estudo desenvolvido por [4], por exemplo, determinou o posicionamento ótimo de VANTs para situações de comunicação sem fio com e sem interferência a partir de uma solução analítica e, assim, maximizou o desempenho dessa solução. Métodos analíticos de posicionamento exigem conhecimento prévio de parâmetros do cenário, como a quantidade de usuários da rede e suas posições, inviabilizando suas aplicações em cenários reais, como os de emergência ou de eventos temporários.

O cenário de aplicação de VANTs para comunicação sem fio apresenta-se ainda mais favorável quando a capacidade de auto-organização da rede é considerada. Uma rede auto-organizável é definida como uma rede capaz de se adaptar, ter funções autônomas e ser ágil o suficiente para manter os objetivos propostos para sua atuação, mesmo quando mudanças ocorram no ambiente [5]. Nesse contexto, técnicas de aprendizado de máquina podem ser aplicadas para possibilitar a auto-organização da rede de comunicação sem fio a partir da utilização de VANTs, conforme apresentado em [6], e a escolha das trajetórias dos VANTs podem ser otimizadas de forma a maximizar a eficiência do consumo energético, assim como apresentado por [7]. Em especial, o *Q-learning*, uma técnica de aprendizagem por reforço, apresenta-se como promissora, tendo em vista que a aplicação dessa técnica permite que os VANTs se auto-organizem e aprendam a partir do resultado obtido em situações anteriores [8].

Nesse sentido, o trabalho apresentado em [9] empregou *Q-learning* para otimizar a posição de VANTs em um cenário de emergência, onde o principal objetivo da solução é maximizar a quantidade de usuários cobertos pela rede de comunicação. Quando comparada a diferentes estratégias de posicionamento fixo, como em posições aleatórias, posições em torno de um círculo no centro do cenário, ou nas localidades dos *hot spots* da rede previamente destruída, a solução de aprendizagem por reforço apresentou desempenho superior, minimizando o número de usuários em indisponibilidade e convergindo de maneira mais rápida. No entanto, o consumo energético não é investigado em [9], apesar de ser um fator fundamental tendo em vista a quantidade de energia limitada e a necessidade de

Ana Flávia dos Reis e Glauber Brante, Universidade Tecnológica Federal do Paraná (UTFPR), Curitiba, PR, Brasil (e-mails: anareis@alunos.utfpr.edu.br e gbrante@utfpr.edu.br). Rafaela Parisotto e Richard Demo Souza, Universidade Federal de Santa Catarina (UFSC), Florianópolis, SC, Brasil (e-mails: rafaela.parisotto@linse.ufsc.br e richard.demo@ufsc.br). O presente trabalho foi realizado com apoio da CAPES, Código de Financiamento 001 e do CNPq.

recarga constante dos VANTs.

O presente artigo apresenta uma extensão do trabalho desenvolvido em [9]. Aqui, é incluída a possibilidade de adaptar a potência de transmissão dos VANTs e uma análise de consumo energético levando em consideração um modelo característico de VANTs movidos por hélices [7]. O objetivo é maximizar a cobertura da rede, minimizando o número de usuários em indisponibilidade. Além disso, é considerado um cenário de eventos temporários, com toda a conectividade sendo provida pelos VANTs. Os resultados evidenciam que a estratégia de alocação de potência de transmissão gera uma maior cobertura da rede e implica em uma menor movimentação dos VANTs gerando, portanto, um menor consumo de energia.

O resto do artigo está organizado da seguinte maneira. O modelo do sistema, incluindo o cenário urbano com a presença de edifícios e usuários da rede, o modelo de comunicações e o modelo de consumo de energia dos VANTs devido as movimentações é apresentado na Seção II. O algoritmo *Q-learning* proposto é apresentado na Seção III, enquanto resultados e respectivas discussões são apresentados na Seção IV. Por fim, a Seção V apresenta as considerações finais.

II. MODELO DO SISTEMA

A. Modelo do Cenário Urbano

O cenário é baseado no modelo urbano de distribuição de prédios e usuários. A União Internacional de Telecomunicações define em [10] que os parâmetros para caracterizar o ambiente urbano são:

- α , a razão entre a área construída e a área total;
- β , o número médio de edifícios por km^2 ;
- γ , o parâmetro de escala da distribuição das alturas dos edifícios, segundo função de densidade de probabilidade de Rayleigh.

Para um cenário urbano, conforme [11], assume-se que a largura de um edifício é dada por $W = 1000\sqrt{\alpha/\beta}$, enquanto o espaço entre edifícios é dado por $S = 1000/\beta - W$. Todos os edifícios que compõem o modelo urbano desse artigo são considerados quadrados e as variáveis W e S são dimensionadas em metros, conforme a Fig. 1. Ainda, são considerados K usuários distribuídos aleatoriamente pelo cenário urbano, em que $\mathcal{K} = \{1, 2, \dots, K\}$ denota o conjunto de usuários ativos.

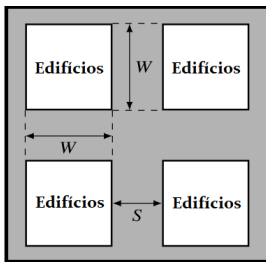


Fig. 1. Modelo urbano de distribuição de edifícios utilizado na aplicação de VANTs proposta nesse artigo, conforme [11].

B. Modelo de Comunicações

É considerado um conjunto de $\mathcal{D} = \{1, 2, \dots, D\}$ VANTs, em que D denota o número de VANTs disponíveis, os quais operam como ERBs móveis e possuem uma antena com ângulo de abertura θ . Dessa forma, o raio de cobertura de cada VANT $j \in \mathcal{D}$ é dado por

$$\varphi_j = h_j \cdot \tan \frac{\theta}{2}, \quad (1)$$

em que h_j é a altura de cada VANT.

A perda de percurso da comunicação é causada pela dissipação da potência irradiada pelo transmissor e pelos efeitos do canal de propagação [12]. A perda de percurso, em dB, entre um VANT $j \in \mathcal{D}$ e um usuário da rede de comunicação sem fio $i \in \mathcal{K}$ separados por uma distância d_{ij} é dada por

$$\kappa_{ij} = 20 \cdot \log_{10} \left(\frac{4\pi f_c d_{ij}}{c} \right) + \xi, \quad (2)$$

onde f_c é a frequência portadora, c é a velocidade da luz e ξ é a perda adicional para a propagação do sinal, que assume diferentes valores dependendo da existência ou não de linha de visada entre o VANT e o usuário [4]. É importante ressaltar que o cálculo da linha de visada depende da realização instantânea do modelo, levando em consideração a posição tridimensional do VANT, o posicionamento e altura dos prédios, bem como o posicionamento dos usuários.

A relação sinal-ruído mais interferência (SINR) de um usuário i alocado em um VANT j é

$$\varrho_{i,j} = \frac{P_{r_{ij}}}{B N_0 + \sum_{k=1, k \neq j}^D P_{r_{i,k}}}, \quad (3)$$

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 $P_{t,j}$. Ainda, o somatório no denominador de (3) apresenta a potência recebida a partir da transmissão de todos os VANTs, ou seja, essa parcela da equação introduz a interferência no sinal recebido causada pelos demais transmissores [12].

A alocação de usuários em cada VANT é feita de acordo com as suas SINRs. Se o usuário atinge determinada SINR limite para um VANT que possui recurso suficiente para aloçá-lo, então esse usuário é alocado para esse VANT. Entretanto, se o VANT em questão não possui recursos disponíveis ou a SINR do usuário está abaixo do limite, o próximo VANT é testado. Caso todas os VANTs sejam testados e o usuário for incapaz de se alocar, esse usuário é considerado em indisponibilidade, ou seja, está fora da cobertura da rede.

C. Modelo de Consumo de Energia

O consumo energético de VANTs de hélice depende, em geral, da velocidade de voo e de sua aceleração. Conforme [7], 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 é dada por [7]

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

onde Ω é a velocidade angular das lâminas, R o raio do rotor, v_0 a velocidade média induzida pelo rotor, d_0 a taxa de arrasto da fuselagem, ρ é a densidade do ar, s a solidez do rotor e A é a área de disco do rotor. Além disso, P_0 é a potência do perfil da lâmina, dada por

$$P_0 = \frac{\delta}{8} \rho s A \Omega^3 R^3, \quad (5)$$

em que δ é o coeficiente de arrasto de perfil, ao passo que a potência induzida é

$$P_i = (1 + k) \frac{\chi^{\frac{3}{2}}}{\sqrt{2\rho A}}, \quad (6)$$

onde k e χ são o fator de correção da potência induzida e peso do VANT, respectivamente.

III. ALGORITMO PROPOSTO

O *Q-learning*, algoritmo utilizado para a auto-organização da rede de comunicação sem fio a partir de VANTs proposta nesse artigo, é uma solução inteligente baseada em aprendizagem por reforço, a qual permite que os agentes do sistema aprendam a atuar de forma autônoma a partir do conhecimento das consequências das suas ações [8].

O presente artigo apresenta uma extensão do trabalho desenvolvido em [9], onde cada VANT é considerado um *agente* da solução *Q-learning* e o *cenário* é composto por edifícios e usuários que podem se movimentar. O conjunto de *estados* possíveis para os agentes consiste nas posições tridimensionais dos VANTs, além dos níveis de potência de transmissão a serem empregados. Dessa forma, nove *ações* são possíveis para cada VANT: movimentar-se para frente, para trás, direita, esquerda, baixo, cima, aumentar ou diminuir a potência, ou não fazer nada. O objetivo das ações dos VANTs é a de aumentar sua *recompensa*, definida como

$$r = \sum_{j=1}^D U_j, \quad (7)$$

em que U_j é o número de usuários conectados ao VANT j . Ainda, assume-se que cada VANT executa seu algoritmo de forma independente, mas tem acesso à quantidade de usuários alocados aos outros VANTs através via conexão de *backhaul*¹.

Assim, cada agente é independente e mantém a sua própria tabela- Q , composta por elementos $Q(s_t, a_t)$ que representam o valor de estar em um *estado* s_t específico, no instante de

tempo t , enquanto realizando uma *ação* a_t específica, sendo atualizada conforme

$$Q(s_t, a_t) = Q(s_t, a_t) + \lambda \left[r_{t+1} + \phi \max_a \{Q(s_{t+1}, a)\} - Q(s_t, a_t) \right], \quad (8)$$

onde λ é a taxa de aprendizagem, r_{t+1} é a recompensa esperada para o próximo instante de tempo, ϕ é o fator de desconto e $\max_a \{Q(s_{t+1}, a)\}$ é uma função de estimação do valor ótimo para a ação do instante futuro.

A chance do algoritmo escolher uma ação aleatória é definida pelo parâmetro ϵ , que estabelece se o agente da solução deve adotar uma estratégia denominada *explore*, onde diferentes ações em diferentes estados são testadas, ou adotar a estratégia *exploit*, onde o agente deve buscar a melhor ação possível dentro de um estado específico. Além disso, cada agente deve observar o número total de usuários alocados pelo sistema de forma a ser incentivado a alocar mais usuários.

Na inicialização do algoritmo, os VANTs são posicionados aleatoriamente e com suas tabelas- Q zeradas. A otimização é dividida em *episódios*, os quais são uma realização instantânea do ambiente em que os usuários são considerados parados. Assim, os VANTs são capazes de realizar algumas ações e de observar a recompensa obtida. Estas ações são realizadas até que os critérios de parada sejam atendidos, os quais são: (i.) um número máximo de iterações (ι_{\max}) seja atingido; (ii.) o valor da recompensa não ter melhorado depois de um certo número de iterações ($\iota_{r,\max}$); (iii.) que todos os usuários tenham sido alocados aos VANTs. Ao final de um episódio, os VANTs se movimentam para uma posição final e alocam a potência de acordo com o estado que possui a maior recompensa. Para o episódio seguinte, assume-se que os usuários se movimentam de forma aleatória, em velocidade pedestre e a não mais do que 3 metros da posição anterior.

A solução proposta é sumarizada no Algoritmo 1.

Algoritmo 1: Algoritmo *Q-learning* proposto.

- 1 Inicializa VANTs em posições aleatórias;
 - 2 Inicializa tabelas- Q com zeros;
 - 3 **para** todo episódio **faça**
 - 4 **enquanto** critérios de parada não forem atingidos **faça**
 - 5 VANTs selecionam ação com maior Q ;
 - 6 Usuários são alocados e observa-se r_{t+1} ;
 - 7 Tabelas- Q são atualizadas conforme (8);
 - 8 **fim**
 - 9 VANTs se movimentam e alocam potência de acordo com a melhor recompensa observada;
 - 10 Usuários são alocados;
 - 11 **fim**
-

IV. RESULTADOS E DISCUSSÕES

O cenário simulado consiste em uma área urbana de 250.000 m² com $K = 250$ usuários. A cada rodada de simulação, o cenário de prédios e posições de usuários são definidos aleatoriamente, de forma que são realizadas 10

¹Assume-se que existe um *backhaul* que conecta as VANTs até a central de controle da rede. Porém, o tratamento deste aspecto foge ao escopo deste artigo. Além disso, a considera-se que a conectividade é provida exclusivamente por ERBs compostas por VANTs.

TABELA I
 PARÂMETROS DE SIMULAÇÃO.

Parâmetro	Valor
Área	500 m × 500 m
Número de usuários (K)	250 usuários
Capacidade de cada VANT	50 usuários
Passo do VANT (eixos X e Y)	50 m
Passo do VANT (eixo Z)	100 m
Altura dos VANTs	[100, 1000] m
Razão entre área construída e área total (α)	0,3 [10]
Número médio de edifícios (β)	500 edifícios/km ² [10]
Parâmetro da dist. de alturas do edifícios (γ)	15 m [10]
Perda adicional de propagação LOS (ξ)	1 dB [11]
Perda adicional de propagação NLOS (ξ)	20 dB [11]
Abertura das antenas (θ)	60°
Largura de banda (B)	180 kHz
Frequência de portadora (f_c)	1 GHz
Número máximo de iterações (t_{max})	200
Iter. sem aumento de recompensa ($t_{r,max}$)	20
Taxa de aprendizagem (λ)	0,9
Fator de desconto (ϕ)	0,9
Fator entre <i>exploit</i> e <i>explore</i> (ϵ)	0,5
Velocidade do VANT (V)	5 m/s
Peso do VANT (χ)	20 N
Densidade do ar (ρ)	1,225 kg/m ³ [7]
Raio do rotor (R)	0,4 m [7]
Área do disco do rotor (A)	0,503 m ² [7]
Velocidade angular das lâminas (Ω)	300 rad/s [7]
Velocidade média induzida pelo rotor (v_0)	4,03 m/s [7]
Taxa de arrasto da fuselagem (d_0)	0,6 [7]
Solidez do rotor (s)	0,05 [7]
Coefficiente de arrasto do perfil (δ)	0,012 [7]
Fator de correção da potência induzida (k)	0,1 [7]

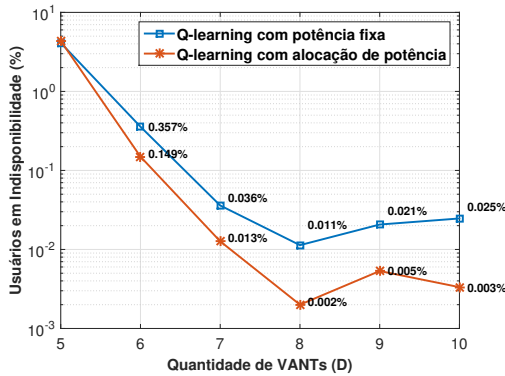


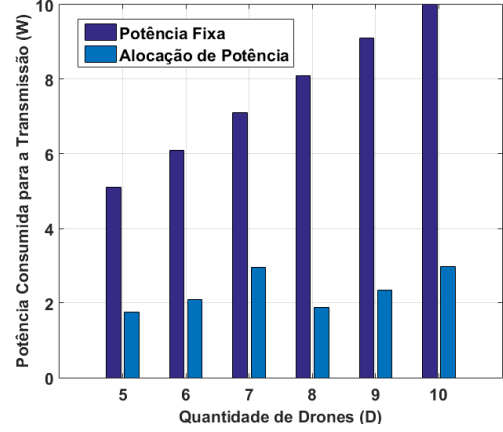
Fig. 2. Porcentagem média de usuários em situação de indisponibilidade.

rodadas de 200 episódios independentes. Seguindo [9], este artigo implementa o algoritmo de *Q-learning* com potência de transmissão fixa em 0 dB, a ser comparado ao método proposto considerando alocação de potência de transmissão dentre três possibilidades: -20 dB, -10 dB e 0 dB. Os parâmetros de simulação são apresentados na Tabela I.

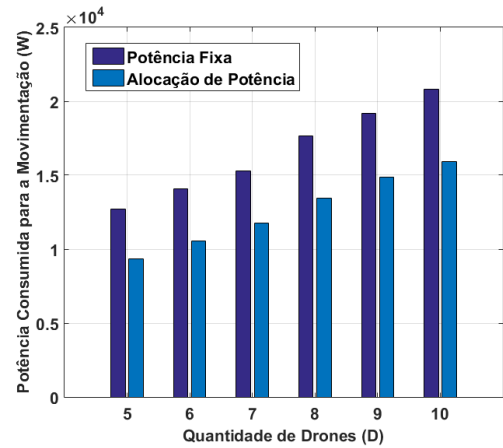
Considerando os dois métodos de posicionamento de VANTs propostos, com potência fixa e com alocação de potência, a porcentagem média de usuários em situação de indisponibilidade, ou seja, não cobertos pela rede, em função da quantidade de VANTs (D) é apresentada na Fig. 2. Como podemos observar, o posicionamento a partir do método *Q-learning* considerando a alocação da potência de transmissão

 TABELA II
 POTÊNCIA DE TRANSMISSÃO ALOCADA PARA CADA VANT.

VANT j	1	2	3	4	5	6	7	8
$P_{t,j}$ (dB)	-10	-10	-20	-20	-20	-10	-10	-10



(a) Potência total de transmissão.



(b) Potência total de movimentação.

Fig. 3. Consumo total de potência de transmissão e de movimentação, considerando todos os VANTs.

dos VANTs apresentou os melhores resultados em relação à indisponibilidade de usuários, com porcentagem de usuários nessa situação atingindo um mínimo de 0,002% para uma rede formada por 8 VANTs. Tal redução de usuários em indisponibilidade se deve principalmente à redução da interferência entre VANTs próximos devido à menor potência utilizada. A Tabela II mostra a potência de transmissão alocada para cada VANT do cenário com 8 VANTs, a qual mostra uma potência média de -11,79 dB.

A Fig. 3 apresenta o consumo total de potência, calculado como a média a partir das 10 rodadas, somando-se as potências utilizadas dentro dos 200 episódios, devido à transmissão de dados (Fig. 3a) e da movimentação dos VANTs (Fig. 3b). Destaca-se que o método de posicionamento com alocação de potência apresentou consumos menores que o método com potência fixa, tanto para a transmissão quanto para a movimentação. Em particular, o algoritmo com potência fixa proposto

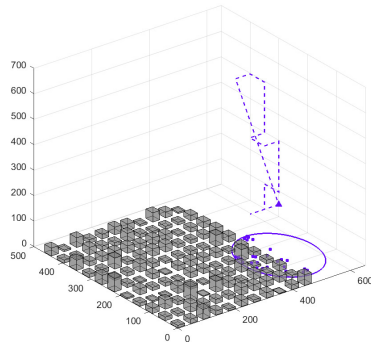
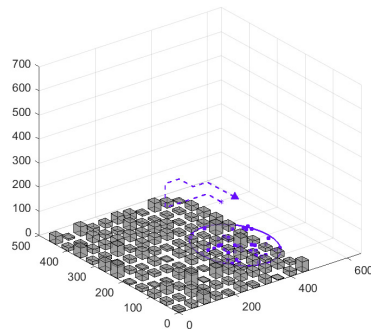
(a) *Q-learning* com potência fixa.(b) *Q-learning* com potência variável.

Fig. 4. Exemplo da trajetória do 8° VANT durante um episódio.

por [9] utiliza a movimentação em altura para reduzir a interferência entre os VANTs, ao passo que a alocação de potências menores pelo algoritmo proposto nesse artigo permite um resultado semelhante, com a vantagem de movimentar menos os VANTs, economizando bateria. Por exemplo, a Fig. 4 ilustra a trajetória do 8° VANT durante um dos episódios executados com os cenários com potência fixa (Fig. 4a) e com alocação de potência (Fig. 4b). Conforme observado, a movimentação do VANT com alocação de potência é consideravelmente menor, justificando o menor consumo de energia.

Por fim, a Fig. 5 apresenta o consumo médio de movimentação por VANT em função de D . Observa-se que o consumo por VANT diminui com o aumento da quantidade de VANTs, indicando que estes precisam se movimentar um pouco menos quando D aumenta dentro dos critérios de convergência adotados. Ainda, cabe ressaltar que a Fig. 2 apresenta um número ótimo de VANTs que minimiza o percentual de usuários em indisponibilidade, ao passo que o consumo de movimentação diminui quando D aumenta, o que pode indicar otimizações em termos de eficiência energética neste cenário.

V. CONSIDERAÇÕES FINAIS

O estudo de técnicas para o posicionamento de VANTs para compor redes sem fio é justificado pela sua flexibilidade e possibilidade de aumentar cobertura em vários cenários, tais quais os de eventos esporádicos considerados nesse trabalho. Nesse contexto, esse artigo consiste na análise comparativa de estratégias para o posicionamento a partir de *Q-learning*, que inclui mobilidade dos VANTs entre as iterações. Os

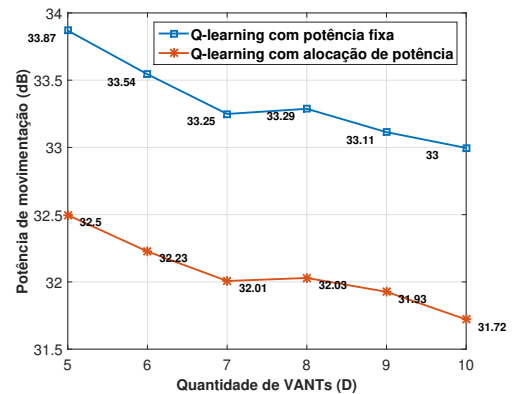


Fig. 5. Consumo médio de movimentação por VANT.

resultados obtidos mostram que a alocação de diferentes níveis de potência transmissão minimizou a quantidade de usuários em indisponibilidade e o consumo de energia para todas as quantidades de VANTs consideradas, evidenciando a relevância do tratamento da variação de potência de transmissão na implementação de algoritmos de auto-organização de redes sem fio. Entre os trabalhos futuros são destacadas as possibilidades de alteração da recompensa do algoritmo de aprendizagem por reforço, *e.g.*, para considerar a eficiência energética, e adição de outras possibilidades de ações para os agentes da solução *Q-learning*, como variação abertura das antenas dos VANTs.

REFERÊNCIAS

- [1] Y. Zeng, R. Zhang, and T. J. Lim, “Wireless communications with unmanned aerial vehicles: Opportunities and challenges,” *IEEE Communications Magazine*, vol. 54, no. 5, pp. 36–42, 2016.
- [2] M. Erdelj, E. Natalizio, K. R. Chowdhury, and I. F. Akyildiz, “Help from the sky: Leveraging UAVs for disaster management,” *IEEE Pervasive Computing*, vol. 16, no. 1, pp. 24–32, 2017.
- [3] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, “Unmanned aerial vehicle with underlaid device-to-device communications: Performance and tradeoffs,” *IEEE Transactions on Wireless Communications*, vol. 15, no. 6, pp. 3949–3963, 2016.
- [4] M. Mozaffari, W. Saad, M. Bennis, and M. Debbah, “Drone small cells in the clouds: Design, deployment and performance analysis,” in *IEEE Global Communications Conference (GLOBECOM)*, 2015, pp. 1–6.
- [5] O. G. Aliu, A. Imran, M. A. Imran, and B. Evans, “A survey of self organisation in future cellular networks,” *IEEE Communications Surveys & Tutorials*, vol. 15, no. 1, pp. 336–361, 2013.
- [6] P. V. Klaine, M. A. Imran, O. Onireti, and R. D. Souza, “A survey of machine learning techniques applied to self-organizing cellular networks,” *IEEE Communications Surveys & Tutorials*, vol. 19, no. 4, pp. 2392–2431, 2017.
- [7] Y. Zeng, J. Xu, and R. Zhang, “Energy minimization for wireless communication with rotary-wing UAV,” *IEEE Transactions on Wireless Communications*, vol. 18, no. 4, pp. 2329–2345, April 2019.
- [8] C. J. C. H. Watkins, “Q-Learning,” vol. 292, pp. 279–292, 1992.
- [9] P. V. Klaine, J. P. Nadas, R. D. Souza, and M. A. Imran, “Distributed drone base station positioning for emergency cellular networks using reinforcement learning,” *Cognitive computation*, vol. 10, no. 5, pp. 790–804, 2018.
- [10] ITU-R Rec., “Propagation Data and Prediction Methods for The Design of Terrestrial Broadband Millimetric Radio Access Systems,” pp. 1410–2, 2003.
- [11] A. Al-Hourani, S. Kandeepan, and A. Jamalipour, “Modeling air-to-ground path loss for low altitude platforms in urban environments,” in *IEEE Global Communications Conference*, 2014, pp. 2898–2904.
- [12] A. Goldsmith, *Wireless communications*. Cambridge University Press, 2005.