

Um controlador para um protocolo MAC híbrido em redes sem-fio baseado em aprendizado por reforço

Camilla B. Sousa¹, Marcelo M. Sobral¹

¹ Área de Telecomunicações – Instituto Federal de Santa Catarina Câmpus São José (IFSC)
Rua José Lino Kretzer 608 – 88.103-902 – São José – SC – Brazil

camilla.b@aluno.ifsc.edu.br

msobral@ifsc.edu.br

Abstract. *This paper proposes a reinforcement learning mechanism to select the operational mode of a hybrid MAC protocol, depending on traffic workload. This kind of protocol combines CSMA/CA and TDMA approaches, aiming to provide better throughput and lower latencies, and has been used in IEEE 802.11 wireless access networks. This paper presents a model of a controller for such MAC, which uses experience to learn to select the best operational model depending on the perceived state of the network.*

Resumo. *Neste trabalho, propõe-se um mecanismo baseado em aprendizado por reforço para escolher o modo de operação de um protocolo MAC híbrido, dependendo da carga de tráfego existente na rede sem-fio. Protocolos como esse combinam MAC do tipo CSMA/CA e TDMA, buscando proporcionar maiores vazões (throughput) e menores latências, e têm sido usados em redes de acesso sem-fio com tecnologia IEEE 802.11. Este trabalho apresenta um modelo para um controlador para esse tipo de MAC, o qual usa experiência para aprender a escolher o melhor modo de operação dependendo da percepção sobre o estado da rede.*

1. Introdução

Em diversos lugares no Brasil e no mundo, onde não há infraestrutura de rede cabeada, usam-se redes de acesso sem-fio para fornecer serviços de dados para clientes [Simó et al. 2007]. As suas estruturas são formadas por estações base, que se situam nas dependências do provedor de acesso, e estações clientes, que ficam nas localidades dos clientes. As distâncias entre estação base e clientes tipicamente são de alguns quilômetros. As estações clientes se associam à estação base, usando-a para enviar e receber quadros de dados. A topologia resultante é denominada PTMP (*Point-To-MultiPoint*), e centralizada na estação base.

Essas redes se baseiam na tecnologia IEEE 802.11 [IEEE 2016], porém com algumas modificações para adaptar o uso dessa tecnologia em longas distâncias e obter melhor eficiência no uso do canal sem-fio [Patra et al. 2007]. Com as distâncias envolvidas, interferências entre estações clientes e estação base, e por transmissões de outras redes, são frequentes. Equipamentos para esse tipo de rede incluem um protocolo MAC modificado, com a finalidade de coordenar as transmissões dos clientes na rede sem-fio ¹. A abor-

¹Alguns exemplos são iPoll da empresa LigoWave, AirMax da empresa Ubiquiti, e NV2 da empresa Mikrotik.

dagem comumente adotada envolve uma forma de TDMA ou *polling*, com a estação base concedendo ciclicamente tempos de acesso aos clientes [Panigrahi and Raman 2009]. Essa forma de operação melhora a utilização do canal sem-fio, evitando a ocorrência de colisões entre transmissões de clientes e da base.

O acesso ao meio sem-disputa, na forma de TDMA ou *polling*, pode ser ineficiente em momentos em que poucos clientes transmitem, ou se as transmissões de clientes são esporádicas [Israr et al. 2012] [Wang et al. 2016]. A consequência é uma maior latência percebida por clientes para transmitirem os seus quadros de dados, pois precisam esperar receberem autorização da estação base. Assim, mesmo estando a rede ociosa, um cliente não pode transmitir enquanto a estação base não o autorizar. Nesse caso, o MAC mais adequado seria o CSMA/CA usual, ao invés de TDMA. Devido a essa questão, alguns fabricantes implementam protocolos MAC híbridos capazes de comutar de um modo CSMA/CA para TDMA e vice-versa ², dependendo da carga de tráfego na rede sem-fio.

Sendo assim, as duas estratégias de acesso ao meio possuem vantagens relativas, dependendo da carga de tráfego na rede. O acesso ao meio mais vantajoso é aquele que, a cada momento, pode proporcionar maior taxa de transmissão e menor latência. Neste trabalho, propõe-se um método para que o protocolo MAC híbrido seja capaz de aprender a escolher dinamicamente o melhor modo de operação. Isso deve ser obtido com uma técnica de Aprendizado por Reforço (*Reinforced Learning*), fazendo com que o MAC melhore o seu desempenho com a experiência obtida com os resultados de escolhas anteriores.

2. Aprendizado por Reforço em protocolos MAC

O Aprendizado por Reforço é um dos paradigmas de Aprendizado de Máquina (*Machine Learning*) e este “usa a estrutura formal dos processos de decisão de Markov para definir a interação entre um agente de aprendizado e seu ambiente em termos de estados, ações e recompensas” [Sutton and Barto 2018]. No diagrama da Figura 1 é possível visualizar a dinâmica entre agente e ambiente. Esses dois blocos interagem periodicamente seguindo a sequência: i) o agente recebe o estado do ambiente (S) e toma uma ação (A), ii) ambiente reage e retorna ao agente o seu novo estado junto com a recompensa (R).

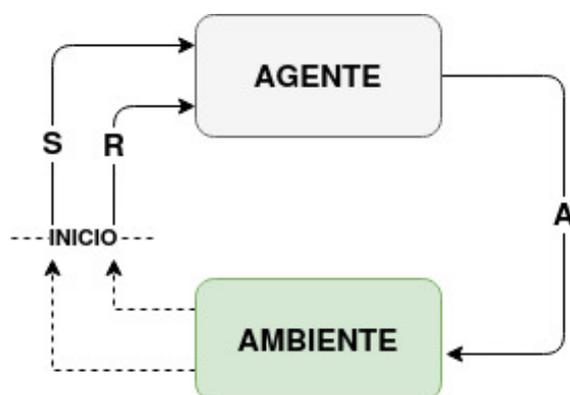


Figura 1. Dinâmica entre agente e ambiente

²Esse é o caso do iPollv3 da empresa LigoWave.

Um Processo de Decisão de Markov (MDP – *Markov Decision Process*) dispõe de um conjunto de ações e um conjunto de estados. As ações são as formas como o agente poderá agir nas tomadas de decisão. Os estados representam todas as configurações possíveis do ambiente durante o processo. A recompensa é um número inteiro que orienta o agente ao objetivo, indicando se a decisão tomada foi boa ou não. A tarefa do agente é maximizar as recompensas, ou seja, em cada estado deverá encontrar a ação que gera resultados com maior valor. O aprendizado ocorre à medida que o agente associa valores a cada ação ou estado, os quais informam o ganho potencial proporcionado por sua escolha. No caso de sistemas com número de estados tratáveis, algoritmos de aprendizado tabulares, tais como *SARSA*, *Q-Learn* e *Expected SARSA*, podem apresentar bons resultados com poucos recursos computacionais [Sutton and Barto 2018].

O uso de aprendizado de máquina em protocolos MAC tem aumentado recentemente. Um mecanismo para ajuste dinâmico de tempos de backoff no protocolo MAC CSMA/CA em redes IEEE802.11 com aprendizado por reforço é apresentado em [Amuru et al. 2015]. Em [Choe et al. 2020] propõe-se um algoritmo adaptativo para um protocolo MAC, em que se usa Deep-Q Network (DQN) para melhorar o desempenho de broadcast em redes veiculares. Em [Edalat and Obraczka 2019] usa-se aprendizado de máquina para ajustar a janela de disputa (CW – *Contention Window*) do protocolo MAC CSMA/CA em redes IEEE 802.11, obtendo maiores vazões e menores latências. O mesmo problema é tratado com uma técnica baseada em redes neurais em [Ali et al. 2017]. O uso de aprendizado por reforço foi proposto por [Pressas et al. 2019], com um algoritmo do tipo Q-Learning para que nodos em uma rede veicular maximizem as entregas de quadros. Esses trabalhos buscam ajustar dinamicamente parâmetros de um protocolo MAC em rede sem-fio, para que se torne capaz de se adaptar às condições do ambiente, melhorando aspectos de seu desempenho. Essas abordagens diferem do proposto neste trabalho, em que se busca selecionar o tipo de protocolo MAC com melhor desempenho para as condições de tráfego existentes em uma rede de acesso sem-fio.

3. Modelo para o MAC adaptativo

A rede de acesso sem-fio considerada neste trabalho tem topologia PTMP com algumas características encontradas em redes de acesso sem-fio implantadas por provedores de acesso. A estação base e as estações clientes (CPEs) possuem antenas setoriais. As distâncias entre CPEs e base é de até 10 km. Com isso, a estação base pode receber as transmissões dos CPEs, mas os CPEs conseguem receber as transmissões somente da base.

O protocolo MAC existente na estação base e nos CPEs é do tipo híbrido, com modos de operação CSMA/CA e TDMA. Quando em modo CSMA/CA, base e CPEs disputam o acesso ao meio igualmente. Em modo TDMA, a base comanda o acesso ao meio, ao enviar ciclicamente a cada CPE uma mensagem de controle para autorizá-lo a transmitir. A ativação de um ou outro modo de operação é comunicada pela base aos CPEs. Dessa forma, o controle do MAC híbrido se faz somente na base.

Para o propósito deste trabalho, supõe-se que a estação base mantenha estatísticas sobre as comunicações com os CPEs. Essas informações incluem as quantidades de erros de transmissão, a latência média e a taxa de quadros por segundo (PPS – *Packets Per Second*). Os erros de transmissão correspondem a quadros não confirmados e quadros

corrompidos. A latência é dada pelo tempo desde que o ingresso de um quadro na fila de transmissão, até o recebimento de sua confirmação quando enfim transmitido. A taxa PPS é calculada pela quantidade de quadros transmitidos em um intervalo de tempo. Essas estatísticas podem ser usadas pelo mecanismo de controle do MAC híbrido para subsidiar sua decisão sobre qual o melhor modo de operação a cada intervalo de tempo.

O mecanismo de controle do MAC híbrido foi modelado como um MDP, cujas probabilidades de transição são determinadas iterativamente por um algoritmo de aprendizado por reforço. Os elementos desse modelo estão descritos a seguir:

- **Estado:** o estado foi definido por um par dado por PPS médio (ajustado com EWMA) e número de CPEs ativos. O conjunto de estados E é dado pelo produto cartesiano pelo conjunto de valores de PPS P , e pelo conjunto de quantidades de CPEs L (i.e. $E = P \times L$). Os valores de PPS são definidos como múltiplos de um valor básico configurável. Por exemplo, se o valor básico for 20 PPS, os valores de PPS para estados seriam 0, 20, 40, até um valor máximo configurado para PPS. A quantidade máxima de CPEs pode ser configurada.
- **Ação:** uma ação pode ser de dois tipos, CSMA/CA ou TDMA. Assim, uma ação determina qual o modo de operação deve ser usado pelo MAC híbrido no próximo intervalo de tempo.
- **Recompensa:** a recompensa foi definida pela soma dos erros de transmissão e da latência média, porém com sinal negativo:

$$r = -(\text{erros} + \text{latencia})$$

Com isso, quantidades menores de erros de transmissão e menores latências, proporcionam maiores recompensas.

- **Política:** a política é do tipo ϵ -greedy, para possibilitar que o mecanismo de controle se adapte a mudanças no padrão de carga de tráfego na rede sem-fio
- **Modelo de ambiente:** desconhecido, pois não se pode caracterizar de antemão o padrão de carga de tráfego.
- **Algoritmo:** o algoritmo de aprendizado escolhido foi *Expected Sarsa* não-episódico e com decaimento ($\gamma < 1$). A escolha se baseou na simplicidade desse algoritmo, que é apropriado devido à relativamente pequena quantidade de estados, e na melhor convergência desse algoritmo em relação a Sarsa e Q-Learn.

4. Resultados preliminares

O mecanismo de controle foi investigado preliminarmente por meio de um simulador do MAC híbrido, o qual foi desenvolvido pelos autores. O simulador implementa os modos de operação CSMA/CA e TDMA, e possibilita criar cenários com uma estação base e um ou mais CPEs dispersos a distâncias aleatórias da base. Não há interferências por transmissões de outras redes ou fontes de radiação eletromagnética, nem erros de transmissão devido a corrompimento de quadros. Portanto, os únicos erros possíveis nesse simulador se devem a colisões. Com isso, se busca avaliar o mecanismo proposto em condições ideais, em que a eficiência do protocolo depende somente da coordenação de transmissões da base e dos clientes. A avaliação do mecanismo em condições reais, com influência de impedimentos por outras fontes de sinal, deve ser feita em um estudo posterior.

O modelo de geração de tráfego utilizado gera quadros com tamanhos distribuídos uniformemente entre 1500 e 16384 octetos, com intervalos também uniformemente distribuídos. Os intervalos entre quadros foram usados para simular diferentes cargas de tráfego na rede. As simulações foram realizadas com redes compostas por 4 e 6 CPEs, distribuídos a distâncias de até 4km da estação base. Para cada rede, executaram-se ensaios com cargas de tráfego crescentes e duração de 10 minutos.

As Figuras 2 e 3 mostram respectivamente os resultados das simulações em redes com 4 e 6 CPEs. Em ambos os casos, o modo TDMA apresenta um comportamento regular, em que a latência média independe da carga de tráfego. O modo CSMA, por outro lado, apresenta latências crescentes à medida que a carga de tráfego aumenta, porém, para cargas baixas, em torno de 200 pps, as latências são inferiores às obtidas com TDMA. O MAC híbrido, controlado pelo mecanismo proposto, apresenta resultados semelhantes ao TDMA, com exceção de cargas de tráfego muito baixas (em torno de 120 pps), quando as latências se aproximam das obtidas com CSMA.

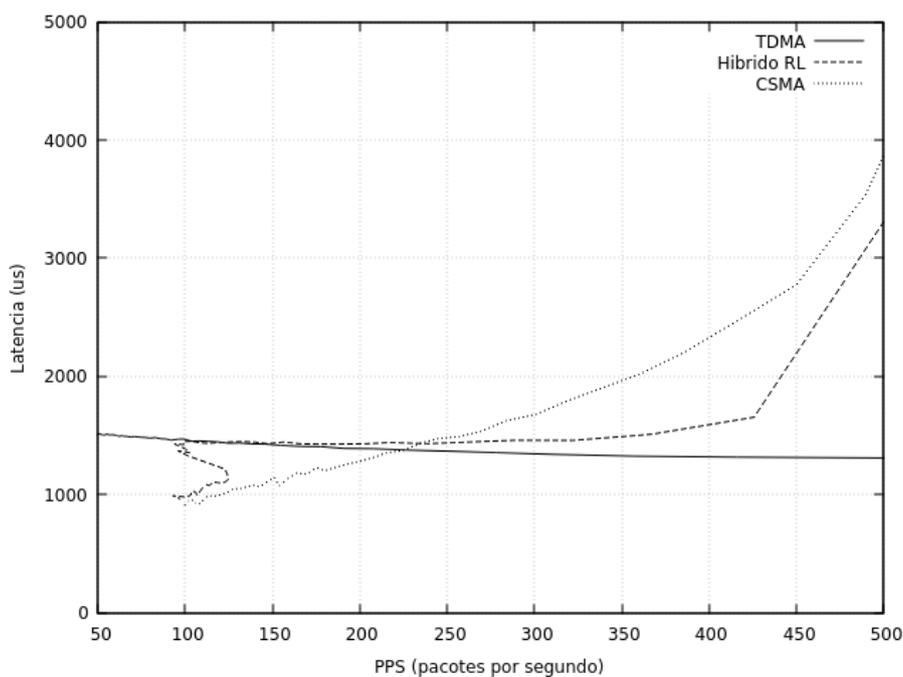


Figura 2. Rede com 4 CPEs

Esses resultados são consistentes com os desempenhos dos modos CSMA/CA e TDMA com baixas e altas cargas de tráfego. Em baixas cargas, tanto CSMA/CA quanto TDMA apresentam baixas taxas de perdas de quadros. Nesse caso, porém, a latência obtida com CSMA/CA é menor, por isso ele se torna a melhor escolha para o MAC. Em cargas de tráfego mais altas tanto a taxa de perdas quanto a latência aumentam com CSMA/CA, mas se mantém estáveis com TDMA, que passa então a ser a melhor escolha. Por outro lado, o controle do MAC híbrido com aprendizado por reforço tendeu a escolher o modo TDMA. Isso ficou evidente em cargas relativamente baixas de tráfego, entre 120 e 250 pps, quando CSMA/CA se desempenha melhor com respeito à latência (e equivalente quanto a taxa de perdas). O MAC híbrido escolheu o modo CSMA/CA somente em baixas taxas de PPS, entre 100 e 120 pps. Porém, a partir dessa taxa, o controlador alternou

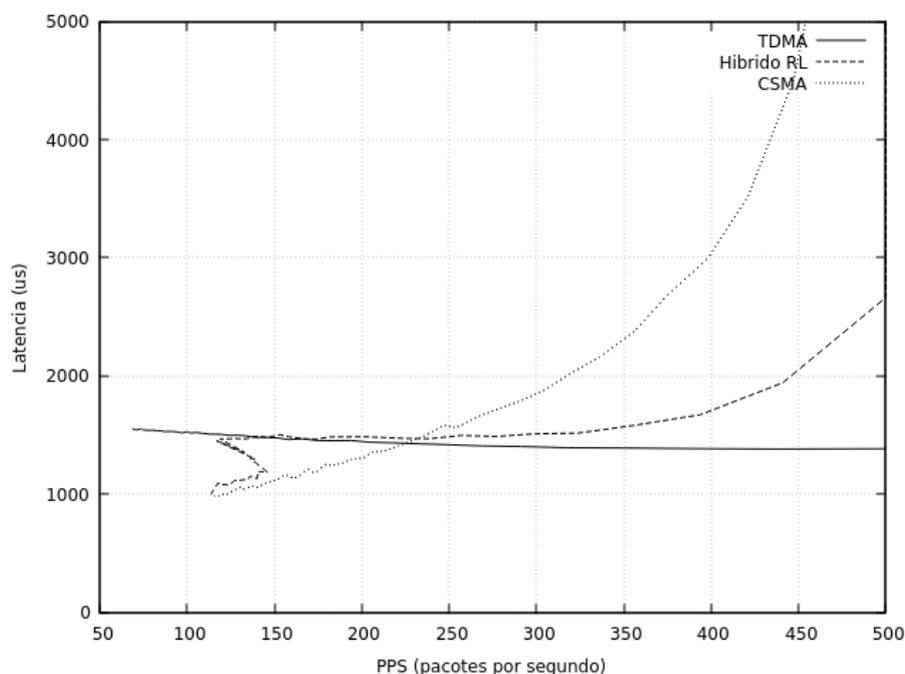


Figura 3. Rede com 6 CPEs

entre os modos CSMA/CA e TDMA, até estabilizar na escolha por TDMA. Durante esse intervalo, colisões e erros de transmissão levaram a taxa de PPS efetiva ser inferior à taxa gerada pela base e clientes. Essa questão deve ser melhor investigada antes de continuar o trabalho com um protótipo em equipamentos reais.

5. Conclusão

Neste trabalho apresentou-se um controlador para um protocolo MAC híbrido para redes de acesso sem-fio baseadas no padrão IEEE 802.11. Esse controlador seleciona o melhor modo de operação para o MAC, dentre as opções CSMA/CA e TDMA. A seleção do modo de operação se faz com uma técnica de aprendizado de máquina baseada em aprendizado por reforço do tipo *Expected SARSA*. Um protótipo do controlador foi implementado e testado com um simulador criado pelos autores. Os resultados preliminares mostraram que o controlador seleciona o melhor modo de operação em baixas e altas cargas de tráfego, porém não em cargas intermediárias. O próximo passo neste estudo será revisar o modelo para que o controlador melhore suas escolhas com cargas intermediárias de tráfego. Em seguida, um protótipo do controlador deverá ser implementado para ser executado em equipamentos reais.

References

- Ali, Z., Giupponi, L., Mangues, J., and Bojovic, B. (2017). Machine Learning Based Scheme for Contention Window Size Adaptation in LTE-LAA. In *in Proceedings of 28th Annual IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (IEEE PIMRC 2017)*, 8-13 October 2017, Montreal (Canada).
- Amuru, S., Xiao, Y., van der Schaar, M., and Buehrer, R. M. (2015). To send or not to send - learning mac contention. In *2015 IEEE Global Communications Conference (GLOBECOM)*, pages 1–6.

- Choe, C., Choi, J., Ahn, J., Park, D., and Ahn, S. (2020). Multiple channel access using deep reinforcement learning for congested vehicular networks. In *2020 IEEE 91st Vehicular Technology Conference (VTC2020-Spring)*, pages 1–6.
- Edalat, Y. and Obraczka, K. (2019). Dynamically tuning ieee 802.11's contention window using machine learning. In *MSWIM '19: Proceedings of the 22nd International ACM Conference on Modeling, Analysis and Simulation of Wireless and Mobile Systems*, pages 19–26.
- IEEE (2016). Ieee standard for information technology—telecommunications and information exchange between systems local and metropolitan area networks—specific requirements - part 11: Wireless lan medium access control (mac) and physical layer (phy) specifications. *IEEE Std 802.11-2016 (Revision of IEEE Std 802.11-2012)*, pages 1–3534.
- Israr, I., Yaqoob, M. M., Javaid, N., Qasim, U., and Khan, Z. A. (2012). Simulation analysis of medium access techniques. In *2012 Seventh International Conference on Broadband, Wireless Computing, Communication and Applications*, pages 602–607.
- Panigrahi, D. and Raman, B. (2009). Tdma scheduling in long-distance wifi networks. In *IEEE INFOCOM 2009*, pages 2931–2935.
- Patra, R., Nedeveschi, S., Surana, S., Sheth, A., Subramanian, L., and Brewer, E. (2007). Wildnet: Design and implementation of high performance wifi based long distance networks. In *4th USENIX Symposium on Networked Systems Design & Implementation (NSDI 07)*, Cambridge, MA. USENIX Association.
- Pressas, A., Sheng, Z., Ali, F., and Tian, D. (2019). A q-learning approach with collective contention estimation for bandwidth-efficient and fair access control in IEEE 802.11p vehicular networks. *IEEE Trans. Veh. Technol.*, 68(9):9136–9150.
- Simó, J., Figuera, C., Seoane, J., and Mart'inez, A. (2007). Distance limits in ieee 802.11 for rural networks in developing countries. *proc. IEEE WRECOM*.
- Sutton, R. S. and Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT press.
- Wang, Q., Jaffrès-Runser, K., Xu, Y., Scharbarg, J., An, Z., and Fraboul, C. (2016). Tdma versus csma/ca for wireless multi-hop communications: A comparison for soft real-time networking. In *2016 IEEE World Conference on Factory Communication Systems (WFCS)*, pages 1–4.