5.3 Results

The results for the ramp scenario are reported in Figure 2, and those for the churn scenario in Figure 3. We see from Figures 2(a) and 3(a) that ALOHA-dQT and ALOHA-dQT-NE yield high network utilization, generally over 75%. If nodes can detect energy in network slots, as in the ALOHA-dQT setup, and thus differentiate empty slots from collisions slots, the performance is generally higher than in the ALOHA-dQT-NE setup, where energy cannot be detected. The performance of ALOHA-dQT approaches that of ALOHA-QTF, indicating that our delayed acknowledgements mechanism yields an efficiency which is almost as good as the ideal case of immediate acknowledgements. The performance for ALOHA-dQT-NE is slightly inferior to that of ALOHA-dQT, indicating that the ability to differentiate empty slots from collisions confers a clear, if relatively small, performance advantage.

In detail, for the ramp scenario, we see that after a brief transient, the network utilization for ALOHA-dQT is above 80% except in a brief transient when nodes become inactive, after about 200 time blocks. The utilization of ALOHA-dQT-NE is similar, but 10% to 15% lower. ALOHA-QTF has overall a slightly greater utilization than ALOHA-dQT. As for the other procols, ALOHA-EB steadily tracks its optimal performance of 37%. ALOHA-Q does not offer optimal performance when the number of active nodes is 50, as one might expect. The reason is that when the number of active nodes is close to the frame length, even though the potential utilization is close to 1, the adaptation time is very long, on the order of hundred of thousands of time slots [5]. Instead, ALOHA-Q is able to reach better performance when the number of active nodes is 30. All the protocols exhibit acceptable fairness, except for a temporary dip when the number of active nodes is increasing. ALOHA-EB, due to its symmetry, offers superior fairness, if not superior utilization.

The utilization in the churn scenario follows a similar pattern, with ALOHA-QTF having highest utilization, closely followed by ALOHA-dQT, which at steady state offers utilization above 75%, and then by ALOHA-dQT-NE with utilization around 65%. ALOHA-EB is once again around 37%, and ALOHA-Q just below 50%. While in the ramp scenario the fairness of ALOHA-dQT-NE was slightly better than the one of ALOHA-dQT, the opposite is true in churn scenario.

In general, the fairness of ALOHA-dQT protocol can be improved at the cost of lower utilization, and vice versa. We can adjust both by using fairness parameter ϵ_r described in section 3.2.

6 CONCLUSIONS

We introduce ALOHA-dQT, a novel channel access protocol based on the use of reinforcement-learning in the context of slotted ALOHA operating in a single-channel fully-connected wireless network. All previous variants of slotted ALOHA based on reinforcement learning, including ALOHA-Q [4, 5], ALOHA-QTF [7], and the deep-RL based approach of [19], assume that a transmitter knows the fate of its transmission at the conclusion of the time slot. In practice, this requires the presence of a repeater that rebroadcasts on a separate channel all packets or explicit acknowledgments. In contrast, ALOHA-dQT is based on an explicit acknowledgement protocol. The acknowledgement protocol is based on the nodes broadcasting, and iteratively merging, their information about the channel history, and updates to the information history drive the reinforcement learning and node adaptation. ALOHA-dQT offers high network utilization, generally above 75%, with fair allocation of bandwidth among active network nodes.

Reinforcement-learning based channel access protocols hold the potential of offering high channel utilization, as the nodes can coordinate their behavior, and we view ALOHA-dQT as a first step in making these protocol suitable for practical use in ad-hoc networks.

ACKNOWLEDGMENTS

This material is based upon work sponsored by the Defense Advanced Research Projects Agency (DARPA) and the Air Force Research Laboratory (AFRL). Any opinions, findings, conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of DARPA or AFRL.

REFERENCES

[1] P. Alvaro, N. Conway, J. M. Hellerstein, and W. R. Marczak. Consistency Analysis in Bloom: a CALM and Collected Approach. In *CIDR*, pages 249–260, 2011.
[2] O. Bousquet and M. K. Warmuth. Tracking a small set of experts by mixing past posteriors. *Journal of Machine Learning Research*, 3(Nov):363–396, 2002.
[3] J. Capetanakis. Generalized tdma: The multi-accessing tree protocol. *IEEE Transactions on Communications*, 27(10):1476–1484, 1979.
[4] Y. Chu, S. Kosunalp, P. D. Mitchell, D. Grace, and T. Clarke. Application of reinforcement learning to medium access control for wireless sensor networks. *Engineering Applications of Artificial Intelligence*, 46:23–32, 2015.
[5] Y. Chu, P. D. Mitchell, and D. Grace. ALOHA and q-learning based medium access control for wireless sensor networks. In *2012 International Symposium on Wireless Communication Systems (ISWCS)*, pages 511–515. IEEE, 2012.
[6] N. Conway, W. R. Marczak, P. Alvaro, J. M. Hellerstein, and D. Maier. Logic and lattices for distributed programming. In *Proceedings of the Third ACM Symposium on Cloud Computing*, pages 1–14, 2012.
[7] L. de Alfaro, M. Zhang, and J. Garcia-Luna-Aceves. Approaching fair collision-free channel access with slotted aloha using collaborative policy-based reinforcement learning. In *IEEE IFIP Networking Conference*, 2020.
[8] D. P. Helmbold, D. D. Long, and B. Sherrod. A dynamic disk spin-down technique for mobile computing. In *Proceedings of the 2nd annual international conference on Mobile computing and networking*, pages 130–142. ACM, 1996.
[9] M. Herbster and M. K. Warmuth. Tracking the best expert. *Machine learning*, 32(2):151–178, 1998.
[10] Huaizhou Shi, R. V. Prasad, E. Onur, and I. G. M. M. Niemegeers. Fairness in Wireless Networks:Issues, Measures and Challenges. *IEEE Communications Surveys & Tutorials*, 16(1):5–24, 2014.
[11] R. K. Jain, D.-M. W. Chiu, and W. R. Hawe. A quantitative measure of fairness and discrimination. *Eastern Research Laboratory, Digital Equipment Corporation, Hudson, MA*, 1984.
[12] G. Jakllari, M. Neufeld, and R. Ramanathan. A framework for frameless TDMA using slot chains. In *2012 IEEE 9th International Conference on Mobile Ad-Hoc and Sensor Systems (MASS 2012)*, pages 56–64, Las Vegas, NV, USA, Oct. 2012. IEEE.
[13] E. E. Khaleghi, C. Adjih, A. Alloum, and P. Mühlethaler. Near-far effect on coded slotted aloha. In *2017 IEEE 28th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, pages 1–7. IEEE, 2017.
[14] L. Kleinrock. *Queueing systems. Volume I: theory.* wiley New York, 1975.
[15] G. Liva. Graph-based analysis and optimization of contention resolution diversity slotted aloha. *IEEE Transactions on Communications*, 59(2):477–487, 2010.
[16] E. Paolini, G. Liva, and M. Chiani. Coded slotted aloha: A graph-based method for uncoordinated multiple access. *IEEE Transactions on Information Theory*, 61(12):6815–6832, 2015.
[17] L. G. Roberts. ALOHA packet system with and without slots and capture. *ACM SIGCOMM Computer Communication Review*, 5(2):28–42, 1975.
[18] F. Schoute. Dynamic frame length aloha. *IEEE Transactions on communications*, 31(4):565–568, 1983.
[19] Y. Yu, T. Wang, and S. C. Liew. Deep-reinforcement learning multiple access for heterogeneous wireless networks. *IEEE Journal on Selected Areas in Communications*, 2019.
[20] M. Zhang, L. de Alfaro, and J. Garcia-Luna-Aceves. Collision-free channel access with delayed acknowledgements using collaborative policy-based reinforcement learning. In *ACM SIGCOMM Conference, NetAI Workshop*, 2020.