

APLICAÇÕES DE ALGORITMOS DE APRENDIZAGEM POR REFORÇO EM PROCESSOS INDUSTRIAIS CONTÍNUOS: UMA REVISÃO

APPLICATIONS OF REINFORCEMENT LEARNING ALGORITHMS IN CONTINUOUS INDUSTRIAL PROCESSES: A REVIEW

André Nakashimaⁱ
Daniel Barbuto Rossatoⁱ
Pedro André Braga de Oliveiraⁱ

RESUMO

A Aprendizagem por Reforço (RL) apresenta um potencial crescente para otimizar o controle de processos industriais contínuos. Este trabalho revisa a literatura recente para sintetizar o estado da arte, analisando os algoritmos, sucessos e desafios de sua aplicação. Os resultados mostram que, embora o RL supere métodos tradicionais em simulações, melhorando a adaptabilidade e o desempenho, sua transição para a prática ainda é um desafio com grande potencial para desenvolvimento. As principais barreiras identificadas são a alta complexidade computacional, a garantia de estabilidade, a necessidade de dados de alta qualidade e a integração com sistemas legados, indicando uma lacuna a ser preenchida entre a pesquisa teórica e a aplicação industrial.

Palavras-chave: aprendizagem por reforço; processos industriais contínuos; revisão sistemática; controle de processos; algoritmos de RL.

ABSTRACT

Reinforcement Learning (RL) shows growing potential for optimizing the control of continuous industrial processes. This paper reviews recent literature to synthesize the state-of-the-art, analyzing the algorithms, successes, and challenges of its application. Results show that while RL outperforms traditional methods in simulations by improving adaptability and performance, its transition to practice remains a challenge with great potential for development. The main barriers identified are high computational complexity, ensuring stability, the need for high-quality data, and integration with legacy systems, highlighting a gap to be filled between theoretical research and industrial application.

Keywords: reinforcement learning; continuous industrial processes; systematic review; process control; RL algorithms.

1 INTRODUÇÃO

A Aprendizagem por Reforço (RL – *Reinforcement Learning*), uma área da Inteligência Artificial (IA) e vertente da Aprendizagem de Máquina (*Machine Learning*), baseia-se em um processo de tomada de decisão inspirado no comportamento humano. O aprendizado ocorre quando um agente executa uma ação em um ambiente e recebe uma recompensa (positiva ou negativa) que influencia suas decisões futuras. Os componentes centrais deste ciclo—agente, ambiente, ação e recompensa—são fundamentados no Processo de Decisão de Markov (MDP), conforme ilustrado na

Figura 1. O objetivo do agente é aprender uma política de ações que maximize a recompensa acumulada ao longo do tempo.

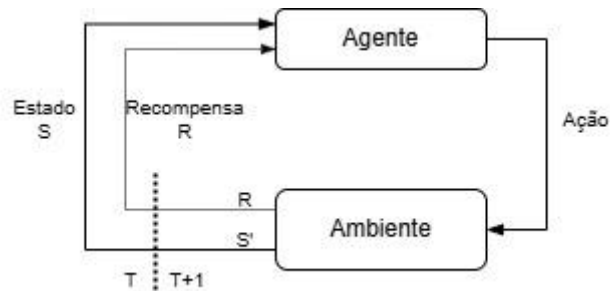


Figura 1. Representação do processo de decisão de Markov

Recentemente, com o avanço e a sofisticação dos algoritmos, a aplicação de RL tem se expandido de áreas como jogos e robótica para se tornar uma solução promissora para os desafios dinâmicos do controle de processos industriais contínuos, onde a otimização da produção, a segurança e a redução de custos são fundamentais.

1.1 Problema de pesquisa

Apesar do sucesso em ambientes simulados, a aplicação bem-sucedida da Aprendizagem por Reforço no mundo real, especialmente em processos contínuos com alta variabilidade, enfrenta desafios significativos. A transição da simulação para a planta industrial é dificultada pela necessidade de simuladores de alta precisão e pela dificuldade em obter dados de alta qualidade para o treinamento, visto que as condições operacionais podem variar de maneira imprevisível. Existe, portanto, uma necessidade de compreender como os desafios de implementação estão sendo enfrentados, quais algoritmos estão sendo utilizados e quais as lacunas de conhecimento que ainda impedem a adoção mais ampla desta tecnologia no setor industrial.

1.2 Objetivos

O objetivo deste trabalho é analisar as características e sintetizar o estado atual do conhecimento sobre os algoritmos de RL e suas aplicações para o controle de processos industriais contínuos. Busca-se, especificamente, destacar quais problemas estão sendo resolvidos, os resultados obtidos e as dificuldades encontradas, com a intenção de identificar o progresso em direção a implementações reais e contribuir para o avanço do conhecimento no campo.

1.3 Justificativa

A revisão da literatura sobre a aplicação de RL em processos industriais contínuos é fundamental para consolidar o conhecimento atual e orientar pesquisas futuras. Ao sintetizar os algoritmos, identificar as dificuldades e apresentar uma visão do estado da arte, este estudo oferece um panorama claro sobre a maturidade da tecnologia. Esta análise é essencial para pesquisadores e profissionais da indústria, pois permite compreender os desafios práticos e as soluções que estão sendo desenvolvidas, acelerando a transição de implementações bem-sucedidas em simulações para aplicações robustas e confiáveis no ambiente industrial.

2 REVISÃO DE LITERATURA

A base da Aprendizagem por Reforço está na interação entre um agente e um ambiente para maximizar uma recompensa cumulativa, um conceito formalizado por (Winder, 2020). Essa abordagem, que aprende por tentativa e erro, distingue-se por sua capacidade de desenvolver políticas de controle ótimas sem um modelo matemático pré-definido do processo. A literatura recente demonstra uma variedade de algoritmos sendo adaptados para o controle industrial, com destaque para famílias como *Policy Gradient* (PG), que otimiza a política de ações diretamente, e *Actor-Critic* (AC), que utiliza uma arquitetura de ator-crítico para refinar as decisões. Extensões com *deep learning*, como DDPG e PPO, surgem para lidar com a alta dimensionalidade e os espaços de ação contínuos dos processos industriais. Uma prática comum nos estudos é comparar o desempenho desses algoritmos com controladores tradicionais, como PID e MPC, para validar sua eficácia (Nievas et al., 2024). No entanto, o desenvolvimento dessas soluções depende fortemente de simulações. (Nian; Liu; Huang, 2020) ressaltam que um simulador preciso é um pré-requisito para o treinamento seguro de um agente de RL. Complementarmente, Dogru et al. (2024) apontam que a falta de dados de alta qualidade para alimentar essas simulações e a dificuldade em modelar a variabilidade dos processos reais são desafios centrais, reforçando a complexidade da transição do ambiente simulado para o prático, onde comportamentos não previstos podem levar a ações incorretas e impactantes.

3 METODOLOGIA

Para garantir uma abordagem sistemática e transparente, este estudo adotou as recomendações da declaração PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses). O processo metodológico iniciou-se com a fase de identificação, na qual foi realizada uma busca na base de dados Google Scholar por artigos publicados nos últimos 5 anos (2019-2024), utilizando as palavras-chave “Reinforcement Learning ” e “industrial plant process control”, o que retornou aproximadamente 17.500 resultados. Em seguida, na etapa de triagem, os 50 primeiros artigos ordenados por relevância foram avaliados a partir de seus títulos e resumos, resultando na seleção de 21 trabalhos alinhados ao escopo. Por fim, na fase de elegibilidade e inclusão, os 21 artigos foram lidos na íntegra, e após a remoção de mais 5 estudos que não atendiam completamente aos critérios, chegou-se ao total de 16 artigos para a análise qualitativa final.

4 RESULTADOS E DISCUSSÕES

A análise dos 16 artigos selecionados revelou uma diversidade de algoritmos de RL aplicados, incluindo variações de Actor-Critic (AC), Deep Deterministic Policy Gradient (DDPG) e Proximal Policy Optimization (PPO), com a maioria dos estudos demonstrando desempenho superior aos controladores tradicionais como PID e MPC em ambientes simulados. Um achado central é que apenas 2 dos 16 estudos realizaram aplicações em ambiente real, evidenciando a lacuna entre a pesquisa teórica e a prática industrial. A partir da análise, consolidou-se que as maiores dificuldades são a complexidade computacional (citada em 12 artigos) e a garantia de estabilidade e convergência (11 artigos), além da necessidade de dados de alta qualidade e integração. Em contrapartida, os principais problemas solucionados foram

a melhoria na qualidade do controle e a adaptação a variações do processo (ambos em 12 artigos), otimização em tempo real (6 artigos) e a redução da intervenção humana (4 artigos). Esses dados indicam que, embora a tecnologia de RL seja eficaz para resolver problemas complexos de controle, sua implementação prática depende da superação de barreiras técnicas significativas.

5 CONCLUSÃO

Conclui-se que a RL demonstra um enorme potencial para otimizar processos, superando controladores tradicionais em cenários simulados ao oferecer maior adaptabilidade e desempenho. No entanto, a transição para aplicações industriais reais ainda é incipiente. Os principais obstáculos são a complexidade computacional, a necessidade de garantir estabilidade e a integração com sistemas legados. Os resultados reforçam que, para que a RL se torne uma realidade consolidada na indústria, pesquisas futuras devem focar na validação prática dos algoritmos e no desenvolvimento de metodologias que enderecem os desafios de implementação no mundo real.

REFERÊNCIAS

- ALHAZMI, K.; ALBALAWI, F.; SARATHY, S. M. A Reinforcement Learning-based Economic Model Predictive Control Framework for Autonomous Operation of Chemical Reactors. arXiv, , 6 maio 2021. Disponível em: <<http://arxiv.org/abs/2105.02656>>. Acesso em: 26 nov. 2024, <https://doi.org/10.1016/j.cej.2021.130993>
- ALVES GOULART, D.; DUTRA PEREIRA, R. Autonomous pH control by reinforcement learning for electroplating industry wastewater. *Computers & Chemical Engineering*, v. 140, p. 106909, set. 2020. <https://doi.org/10.1016/j.compchemeng.2020.106909>
- BAO, Y.; ZHU, Y.; QIAN, F. A Deep Reinforcement Learning Approach to Improve the Learning Performance in Process Control. *Industrial & Engineering Chemistry Research*, v. 60, n. 15, p. 5504–5515, 21 abr. 2021. <https://doi.org/10.1021/acs.iecr.0c05678>
- Blum, N., Krespach, V., Zapp, G., Oehse, C., Rehfeldt, S. and Klein, H. (2021), Investigation of a Model-Based Deep Reinforcement Learning Controller Applied to an Air Separation Unit in a Production Environment. *Chemie Ingenieur Technik*, 93: 1937-1948. <https://doi.org/10.1002/cite.202100094>
- CHEN, K. et al. Optimal control towards sustainable wastewater treatment plants based on multi-agent reinforcement learning. *Chemosphere*, v. 279, p. 130498, set. 2021. <https://doi.org/10.1016/j.chemosphere.2021.130498>
- KUBOSAWA, S.; ONISHI, T.; TSURUOKA, Y. Synthesizing Chemical Plant Operation Procedures using Knowledge, Dynamic Simulation and Deep Reinforcement Learning. arXiv, , 6 mar. 2019. Disponível em: <<http://arxiv.org/abs/1903.02183>>. Acesso em: 26 nov. 2024 <https://doi.org/10.48550/arXiv.1903.02183>
- LEE, D. et al. Comparison of Deep Reinforcement Learning and PID Controllers for Automatic Cold Shutdown Operation. *Energies*, v. 15, n. 8, p. 2834, 13 abr. 2022. <http://dx.doi.org/10.3390/en15082834>

- LI, J. et al. Nonzero-Sum Game Reinforcement Learning for Performance Optimization in Large-Scale Industrial Processes. *IEEE Transactions on Cybernetics*, v. 50, n. 9, p. 4132–4145, set. 2020. <https://doi.org/10.1109/TCYB.2019.2950262>
- LIU, C.; DING, J.; SUN, J. Reinforcement Learning Based Decision Making of Operational Indices in Process Industry Under Changing Environment. *IEEE Transactions on Industrial Informatics*, v. 17, n. 4, p. 2727–2736, abr. 2021. <http://dx.doi.org/10.1109/TII.2020.3005207>
- MOWBRAY, M. et al. Using process data to generate an optimal control policy via apprenticeship and reinforcement learning. *AIChE Journal*, v. 67, n. 9, p. e17306, set. 2021. <http://dx.doi.org/10.1002/aic.17306>
- NIEVAS, N. et al. Reinforcement Learning for Autonomous Process Control in Industry 4.0: Advantages and Challenges. *Applied Artificial Intelligence*, v. 38, n. 1, p. 2383101, 31 dez. 2024. <https://doi.org/10.1080/08839514.2024.2383101>
- PAGE, M. J. et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. *BMJ*, p. n160, 29 mar. 2021a.
- POWELL, B. K. M.; MACHALEK, D.; QUAH, T. Real-time optimization using reinforcement learning. *Computers & Chemical Engineering*, v. 143, p. 107077, dez. 2020. <https://doi.org/10.1016/j.compchemeng.2020.107077>
- SACHIO, S. et al. Integrating process design and control using reinforcement learning. *arXiv*, , 11 ago. 2021. Disponível em: <<http://arxiv.org/abs/2108.05242>>. Acesso em: 26 nov. 2024 <https://doi.org/10.1016/j.cherd.2021.10.032>
- SHUPRAJHAA, T.; SUJIT, S. K.; SRINIVASAN, K. Reinforcement learning based adaptive PID controller design for control of linear/nonlinear unstable processes. *Applied Soft Computing*, v. 128, p. 109450, out. 2022. <https://doi.org/10.1016/j.asoc.2022.109450>
- SPIELBERG, S. et al. Deep Reinforcement Learning for Process Control: A Primer for Beginners. 11 abr. 2020. <https://doi.org/10.48550/arXiv.2004.05490>
- WINDER, P. Reinforcement Learning: Industrial Applications of Intelligent Agents. 1st edition ed. Sebastopol, CA: O'Reilly Media, Inc., 2020.
- YANG, Q. et al. Reinforcement-Learning-Based Tracking Control of Waste Water Treatment Process Under Realistic System Conditions and Control Performance Requirements. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, v. 52, n. 8, p. 5284–5294, ago. 2022. <https://doi.org/10.1109/TSMC.2021.3122802>
- ZHU, L. et al. Alleviating parameter-tuning burden in reinforcement learning for large-scale process control. *Computers & Chemical Engineering*, v. 158, p. 107658, fev. 2022. <https://doi.org/10.1016/j.compchemeng.2022.107658>

SOBRE O(S)AUTOR(ES)

André Nakashima



Possui graduação em engenharia de controle e automação pela Univesidade São Judas Tadeu. Pós-graduação em automação e controle do Centro Universitário SENAI-SP, Campus “Mariano Ferraz”. Possui experiência em projetos de engenharia de automação de sistemas com foco em instalação e configuração de Sistemas de controle distribuído de Processos.

Daniel Barbuto Rossato



Possui doutorado em Engenharia Eletrônica e Computação pelo ITA (2022), mestrado em Engenharia Elétrica – Sistemas pela EPUSP (2009) e bacharelado em Engenharia Elétrica – Automação e Controle pela EPUSP (2002). Atualmente é docente na Faculdade SENAI São Paulo – Campus Mariano Ferraz

Pedro André Braga de Oliveira



Possui graduação em Matemática Aplicada e Computacional com habilitação em sistemas e controle pela Universidade de São Paulo - USP (2009), especialização em Ciência de Dados (*Big Data Analytics*) pela Universidade Presbiteriana Mackenzie (2019). Atualmente é professor da Faculdade SENAI São Paulo “Mariano Ferraz”.