

MACHINE LEARNING APLICADO AO CONTROLE AUTÔNOMO DE PROCESSOS

Douglas Alves Goulart¹;
Renato Dutra Pereira Filho²

¹Universidade Federal do Rio Grande – douglasagoulart@gmail.com

²Universidade Federal do Rio Grande – renatodutrapereira@gmail.com

1. INTRODUÇÃO

Em ambientes industriais, com problemas reais e complexos, o caráter não-linear é preponderante (AGUIRRE et al., 2007). Construir modelos matemáticos em tais cenários pode se tornar complicado e inviável, o que desfavorece abordagens/métodos convencionais de controle automático. Além do que, sistemas de controle incapazes de aprender e se adaptar dependem fortemente da existência de modelos matemáticos que descrevam com precisão o problema tratado.

Em problemas reais, é fundamental que o sistema consiga se ajustar de maneira a se aproximar do comportamento esperado, fazendo mínimo uso possível de interferência humana. O processo de obtenção de tal aproximação é, basicamente, o processo de aprendizado (BECKY, 2005).

As técnicas de controle autônomos e inteligentes são assim classificadas pois apresentam capacidades de aprendizado e adaptação especialmente adequadas aos problemas complexos e dinâmicos. Essas capacidades são oriundas da interação massiva do controlador com o ambiente, como no caso do *Reinforcement Learning*.

Este trabalho teve como objetivo principal avaliar a aplicabilidade de técnica de Aprendizado de Máquina (*Machine Learning*) para controle autônomo de nível de tanque, baseado em Aprendizado por Reforço (*Reinforcement Learning - RL*) completamente automatizado.

Os propósitos secundários foram: validar e comparar as performances dos controladores, autônomo baseado em RL e o clássico Proporcional-Integral. Além de adaptar o algoritmo de Otimização por Enxame de Partículas (*Particle Swarm Optimization – PSO*) para otimização dos parâmetros dos controladores RL e PI.

2. METODOLOGIA

Algoritmos de inteligência artificial com técnicas de aprendizado, tem sido uma alternativa às técnicas de controle tradicionais há bastante tempo, como mostrado já por HOSKINS; HIMMELBLAU (1992). Uma das mais promissoras técnicas é o *Reinforcement Learning*, uma técnica de aprendizado livre de modelos baseado em interações do agente controlador com o ambiente.

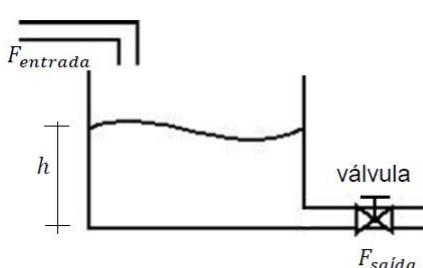
De maneira resumida, a dinâmica do RL segundo BUSONIU et. al. (2017) é descrita como: a cada instante de tempo, o agente recebe do ambiente uma representação do seu estado, e com essa informação ele seleciona uma ação. No instante de tempo seguinte, o agente recebe uma recompensa e um novo estado consequências da ação.

De maneira geral, o agente implementa um mapeamento dos estados em probabilidades de realizar ações. Tal mapeamento é realizado com o aprendizado do agente de acordo com sua experiência de ações tomadas e as respectivas recompensas recebidas.

No algoritmo de RL, neste trabalho desenvolvido, as ações representavam a abertura de uma válvula que regula a vazão de saída de um tanque com uma vazão de entrada fixa, os estados são os níveis do tanque, e as recompensas são baseadas no qual perto o nível está do nível desejado.

A simulação do processo deu-se considerando um tanque de $3m^2$ de área com 1m de altura, onde o nível objetivo a ser alcançado é de 0,5m, o nível inicial é de 0,2m e a vazão de entrada de $100m^3/h$. A equação diferencial que modela o sistema baseada na conservação de massa foi solucionada pelo método de Runge-Kutta de 4^a ordem.

Figura 1: Descrição do Tanque de Líquido e seu modelo matemático.



Tanque aberto de descarga aberta, isotérmico, sem reações químicas. Fluido newtoniano.

Variável controlada: nível de líquido.

Variável manipulada: vazão de saída de líquido

$$\frac{dh}{dt} = \frac{F_{\text{entrada}} - F_{\text{saída}}}{A}$$

$$F_{\text{saída}} = f(x) \cdot Cv \sqrt{\Delta P}$$

$$\Delta P = (P_0 + \rho \cdot g \cdot h) - P_0$$

$$\Delta P = \rho \cdot g \cdot h$$

onde:

h - nível de líquido (m);

t - tempo (h);

F - vazão (m^3/h);

A - Área transversal ao escoamento no tanque (m^2);

$f(x)$ - abertura da válvula - advinda do controlador;

Cv - característica da válvula;

P - Pressão;

ρ - massa específica do fluido.

O RL utilizou a metodologia de aprendizado Ator-Crítico, a qual melhor se encaixa com o trabalho de estados e ações no espaço contínuo, devido a separação das estruturas de avaliação das ações e do mapeamento de recompensas.

O controlador RL é considerado autônomo já que todos seus hiperparâmetros que dependem do ambiente e do seu método de aprendizado foram determinados por um algoritmo de otimização. Assim garantindo seu melhor desempenho, e fazendo com que o único conhecimento do processo exigido para sua execução seja a função de recompensa, para informar o objetivo a ser atingido pelo controlador.

Este algoritmo apresenta sete parâmetros a serem otimizados, são eles:

- Número de funções utilizadas para descrever os espaços contínuos;
- 2 Ganhos de aprendizado do Ator-Crítico;
- Fator de exploração de novas ações e estados;
- Peso de ajuste da recompensa;
- Fator de desconto de recompensas futuras;
- Parâmetro de decaimento do traço de elegibilidade.

Estes são otimizados pelo PSO, adaptado pra trabalhar com 7 dimensões, como o objetivo de minimizar o somatório do erro quadrático, ISE, do nível do tanque com a altura objetivo.

O algoritmo original de otimização por enxame de partículas (PSO), foi desenvolvido por KENNEDY; EBERHART (1995) e faz uso de um paradigma computacional fundamentado no fenômeno da inteligência coletiva exibida por enxames de insetos e bando de pássaros. No PSO, cada partícula representa uma possível solução do problema de otimização, é inicializada a partir da atribuição de posições e velocidades randômicas, soluções essas que irão “transitar” através do espaço multidimensional de soluções possíveis.

A partícula é avaliada através do somatório do erro quadrático (ISE) de uma simulação final, após 50 simulações de treinamento do algoritmo do aprendizado. As partículas com os melhores índices de desempenho influenciam as demais, convergindo o enxame à solução ótima. Com a mesma métrica e algoritmo, os parâmetros do controlador PI também foram otimizados.

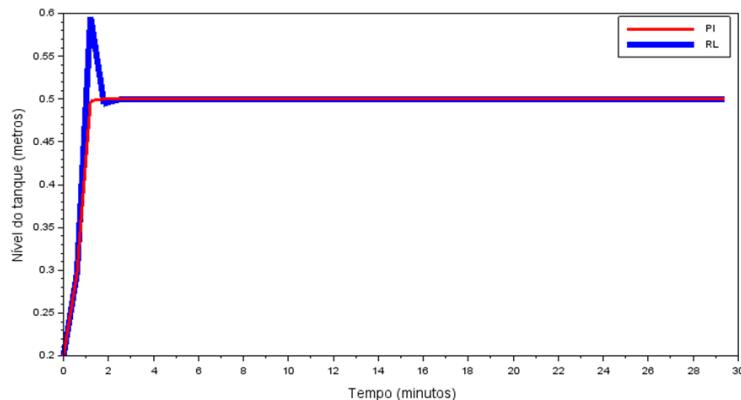
Os algoritmos “Controlador baseado em *Reinforcement Learning*”, “Simulador de tanque com correntes de entrada e saída” e “Otimizador por enxame de partículas” foram desenvolvidos, acoplados na linguagem C++ e executados na plataforma de computação em nuvem AWS – *Amazon Web Services*, mais especificamente na *IDE Cloud9*.

A instância virtual de processamento na AWS foi a instancia gratuita t2.micro, que consistem em 1 CPU virtual com 6 créditos de CPU/hora, memória de 1GiB e armazenamento no *Amazon Elastic Block Store*.

3. RESULTADOS E DISCUSSÃO

Com a configuração para o otimizador de 100 particulas com 100 iterações máximas, o algoritmo levou cerca de 23h para ser concluído. Os comportamentos dinâmicos da variável controlada, o nível do tanque, é mostrado na Figura 2.

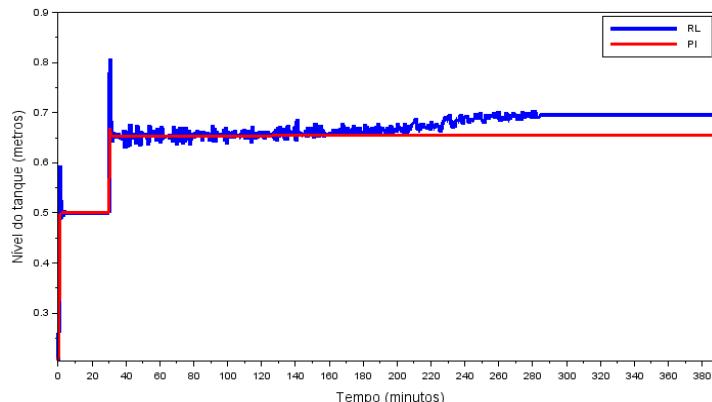
Figura 2: Nível do tanque controlado pelas metodologias PI e RL.



Apesar do RL levar um pouco mais de tempo para estabilizar a resposta, observa-se a boa atuação dos controladores nas condições otimizadas.

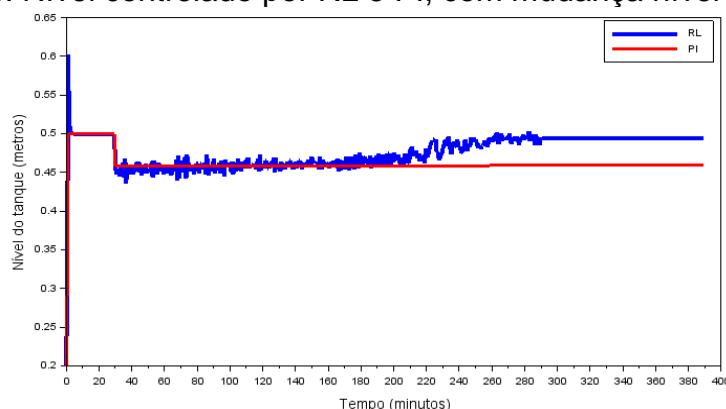
Entretanto, fez-se o teste para simulações onde há mudanças nas condições de operação. Executou-se a simulação, mas a partir dos 30 min a vazão inicial passa de 100m³/h para 85m³/h. O aprendizado com a nova situação e adaptabilidade do controlador autônomo (RL) é ilustrada na Figura 3.

Figura 3: Nível controlado por RL e PI, com mudança na vazão de entrada.



Por último os controladores foram submetidos a uma operação servo, onde aos 30 minutos o nível objetivo passava de 0,5m para 0,7m (Figura 4).

Figura 4: Nível controlado por RL e PI, com mudança nível objetivo.



O controlador baseado em RL seguiu aprendendo com o ambiente e atingiu o novo *setpoint* - objetivo sem intervenções no mesmo. Enquanto, claramente o controlador clássico Proporcional-Integral precisaria de novo processo de otimização para atingir, com excelência, o novo objetivo.

4. CONCLUSÕES

A viabilidade e adaptatividade do controlador autônomo baseado em *Machine Learning*, com o algoritmo de RL foi mostrada neste trabalho.

O controlador automático baseado em *Reinforcement Learning* mostrou-se robusto à variações no processo, no estudo de caso abordado, mudança na vazão de entrada do tanque – perturbação na carga do sistema dinâmico - e no nível de referência do líquido – setpoint - , sendo capaz de aprender a partir da interação com o ambiente durante o processo.

Deverão ser realizados estudos posteriores envolvendo processos com comportamento transiente, em especial processos operando em batelada, bem como processos com comportamentos fortemente não-lineares, para avaliar a efetividade do controle autônomo de processos baseado em *Machine Learning*.

5. REFERÊNCIAS BIBLIOGRÁFICAS

- AGUIRRE, LA; BRUCIAPAGLIA, AH; MIYAGI, PE; HIROSHI, R; TAKAHASHI,C. **Enciclopédia de automática: controle & automação**. Blucher, 2007.
- BECKY, GA. **Autonomous robots: from biological inspiration to implementation and control**. MIT press, 2005.
- HOSKINS, J; HIMMELBLAU, D. Process control via artificial neural networks and reinforcement learning. **Computers & Chemical Eng.**, v.16, n.4, p.241-251, 1992.
- BUSONIU, L; BABUSKA, R; SCHUTTER, B; ERNST, D; **Reinforcement learning and dynamic programming using function approximators**. CRC press, 2017.
- KENNEDY, J; EBERHART, R; Particle swarm optimization (PSO). **IEEE International Conference on Neural Networks**, Australia, p.1942-1948, 1995.