

Problème de remplacement d'équipements avec des machines seconde main

Chen Xiong¹, Guillaume Massonnet¹, Simon Thevenin¹

IMT Atlantique, LS2N-CNRS, La Chantrerie, 4 rue Alfred Kastler, Nantes, France
{chen.xiong, guillaume.massonnet, simon.thevenin}@imt-atlantique.fr

Mots-clés : *Apprentissage par renforcement, constraint markov decision process, multi-agent reinforcement learning.*

1 Introduction et problématique

De nos jours, les lignes de fabrication subissent fréquemment des reconfigurations. Il n'est pas rare de voir une entreprises modifier sa gamme de produits tous les 6 mois, entraînant le retrait de ressources industrielles de leur ligne d'assemblage avant d'atteindre leur fin de vie. Ces ressources se retrouvent ensuite sur des marchés de seconde main, où d'autres fabricants sont susceptibles de les acquérir pour un tarif plus abordable qu'un équipement neuf. L'achat de machines d'occasion présente bien sûr des inconvénients, car l'équipement peut ne pas être aussi fiable qu'un neuf. Cependant, elle prolonge également leur durée de vie utile, contribuant ainsi à l'amélioration de l'empreinte écologique globale des équipements concernés. Par conséquent, le marché mondial des robots d'occasion et remis à neuf devrait passer de 1,9 milliard de dollars en 2022 à 5 milliards de dollars américains en 2031 [1].

Notre travail se concentre sur la conception d'une politique de remplacement d'équipement qui prend en compte la possibilité de se fournir sur un marché d'occasion. Nous étendons le problème de remplacement de machine pour modéliser son utilisation dans une ligne d'assemblage et intégrer l'action de remplacer un robot par un robot d'occasion. Nous modélisons le problème sous la forme d'un processus de décision Markovien (MDP). Comme le nombre de machines dans une chaîne de montage peut être très large, la taille du MDP pour beaucoup d'instances réalistes est trop grande pour être résolue de manière optimale. Nous considérons donc une approche d'apprentissage par renforcement profond pour optimiser la politique.

2 Modèle

Nous étendons le modèle de maintenance préventive [2] pour l'adapter au contexte de ligne d'assemblage avec des machines de seconde main. Nous utilisons le processus de Markov pour simuler l'évolution de l'état de chaque machine. Nous introduisons les définitions suivantes :

- Espace d'état : $\mathcal{S} = \{0(New), 1(Good), 2(Average), 3(Defect), 4(Fail)\}$
- Espace d'action : En considérant le remplacement par machine de second-mains, on peut définir $\mathcal{A} = \{K(keep), R(repair), RN(replace by new), RO(replace by second-hand)\}$.
- On modélise l'évolution de la dégradation au cours du temps par un processus gamma.

Le problème étudié intègre les récompenses et les coûts suivants :

- Un revenu de production par unité de temps (V_p) identique pour tous les états sauf lorsque la machine ne fonctionne pas (état *Fail*)
- Un coût d'inspection par unité de temps pour l'état 1, 2 et 3 égal à C_{m1} , C_{m2} and C_{m3}
- Un coût de réparation pour une machine et pour la remettre dans l'état précédent depuis l'état 1, 2, 3 et 4, égaux respectivement à $C_{r1} \leq C_{r2} \leq C_{r3} \leq C_{r4}$.
- Un coût de défaillance C_f encouru si la machine se tombe en panne.

- Les prix d’achat d’une nouvelle machine et d’une machine de seconde main dans un état intermédiaire (1,2 ou 3) sont $C_{pn} \geq C_{po1} \geq C_{po2} \geq C_{po3}$.
- Un coût de pénalité C_p appliqué en cas d’arrêt du système lors du remplacement ou de la réparation d’une machine entre deux inspection.

Pour calculer la probabilité de transition, on intègre les densités de probabilité pour chaque distribution sur les intervalles définies par les seuils de dégradation $H_1 < H_2 < H_3$.

3 Méthodologies

Nous avons développé une approche de Q-learning classique pour résoudre le problème de maintenance des machines sur une ligne d’assemblage. Cette approche fonctionne bien pour le cas où il y a peu de machines sur la ligne, mais elle n’est pas capable de passer à l’échelle lorsque le nombre de machines augmente. En effet, la Q-table devient trop grande, ce qui pose des problèmes de mémoire et de temps de calcul. Ainsi, nous adoptons le Deep Q-learning avec remémoration d’expérience [3], qui utilise des réseaux de neurones profonds pour une modélisation efficace de la fonction de valeur Q dans des contextes à haute dimension. En complément de l’utilisation d’un réseau de neurones conventionnel, nous explorons l’apprentissage par renforcement multi-agents [4]. Cette approche maintient des Q-tables indépendantes pour chaque agent (machine), les utilisant pour construire la Q-table globale du système. Cette approche permet d’atteindre une stratégie approximativement optimale dans le contexte complexe de la maintenance des machines sur des lignes d’assemblage à plusieurs agents.

Dans le domaine de la production industrielle, notre objectif est d’optimiser la productivité tout en minimisant les coûts. Cependant, lors des opérations de maintenance des lignes d’assemblage, les arrêts de machines peuvent entraîner des pertes de productivité. Pour remédier à cela, nous envisageons d’introduire une contrainte visant à minimiser la probabilité d’interruption de la ligne. Nous envisageons de modéliser le problème avec un Processus de décision Markovien contraint [5] et d’utiliser l’apprentissage par renforcement sous contraintes. À l’avenir, nous prévoyons d’intégrer cette approche avec l’apprentissage par renforcement multi-agents.

Remerciements

Ce travail a été réalisé dans le cadre du projet ALICIA financé par la Commission européenne, sous le numéro de subvention 101091577. Les auteurs tiennent également à remercier la région Pays de la Loire pour son soutien financier et le centre de calcul CCIPL pour la généreuse allocation de ressources informatiques.

Références

- [1] Astute Analytica *Used and refurbished robot market – Industry Dynamics, Market Size, and Opportunity Forecast up to 2031*. "<https://www.astuteanalytica.com/industry-report/toc/used-and-refurbished-robots-market>", 2023.
- [2] Nooshin Yousefi et al. *Reinforcement learning for dynamic condition-based maintenance of a system with individually repairable components*. *Quality Engineering*, 2020.
- [3] Volodymyr Mnih et al. *Playing Atari with Deep Reinforcement Learning*. *CoRR*, 2013.
- [4] Peter Sunehag et al. *Value-Decomposition Networks For Cooperative Multi-Agent Learning*. *CoRR*, 2017.
- [5] Yonathan Efroni et al. *Exploration-Exploitation in Constrained MDPs*. *CoRR*, 2020.