

Allocation dynamique de ressources : comparaison d’approches de recherche opérationnelle et d’apprentissage par renforcement

L. Trassoudaine^{1,2}, M-H. Ibrahim¹, O. Ben-Ammar¹, S. Harispe¹

¹ SyCoIA, IMT Mines Ales, Ales, France
lois.trassoudaine@mines-ales.fr

² ESII, Laverune, France

Mots-clés : *Allocation Dynamique, Gestion de files d’attentes, Optimisation Combinatoire, Intelligence Artificielle, Apprentissage par Renforcement*

1 Introduction

De nombreux secteurs s’intéressent à la gestion de files d’attente physiques ou digitales. L’optimisation de ces files d’attente est par exemple particulièrement cruciale dans les hôpitaux, où la gestion des services d’urgence est un enjeu majeur. Nous exposons dans cette note le positionnement de nos travaux sur la gestion de files d’attente et l’allocation de ressources. Ceux-ci, résolument appliqués, sont réalisés en collaboration avec la société ESII, qui apporte ses données ainsi que son expertise en matière d’optimisation de files d’attente.

Nous nous intéressons spécifiquement au problème d’allocation d’agents physiques à des usagers arrivant dynamiquement et stochastiquement sur un site. Nous nous appuyons pour cela sur des approches d’optimisation issues de la Recherche Opérationnelle (RO) et de l’Intelligence Artificielle (IA). L’objectif est de proposer la meilleure allocation - au sens de la minimisation du temps passé par les usagers dans le système - en un temps de calcul quasi-instantané.

L’optimisation de l’allocation de ressources dans des contextes dynamiques a été largement étudiée dans la littérature scientifique depuis plus de 60 ans. Une part importante de la littérature en RO sur le sujet provient notamment du domaine de la gestion de files d’attente.¹ Plusieurs développements méthodologiques se sont en particulier concentrés sur des politiques simples associées à de bonnes garanties de performances théoriques dans des structures de réseaux spécifiques et au regard de fonctions objectif particulières, e.g. approches de type (*generalised*) $c\mu$ -rule ou *MaxWeight* [1] dans des contextes de fonctions de coût convexes par rapport aux tailles des files d’attentes.

Bien que très intéressantes sur plusieurs aspects (e.g., garanties théoriques, compréhension mathématique fine), ces politiques sont parfois difficilement utilisables en pratique ou ne permettent pas un traitement satisfaisant pour différentes classes ou variantes de problèmes d’allocation dynamique retrouvés dans de nombreux secteurs. Cela s’explique par le fait qu’elles soient difficilement adaptables à certains cas spécifiques, en particulier dans le cas de non-stationnarités, e.g. des processus stochastiques qui gouvernent l’arrivée de type de clients ou demandes et les temps de traitement, ou lorsque des contraintes propres au métier s’ajoutent, e.g., gestion de files d’attente couplée à des rendez-vous définis à l’avance.

En réponse à ces limitations, une diversité d’approches à base d’IA et plus spécifiquement à base d’Apprentissage Automatique sont explorées depuis quelques années. L’objectif est dans ce cas d’apprendre les politiques dans des contextes simulés ou à partir de données, plutôt que de les définir de manière ad hoc. Les approches étudiées adoptent très souvent le paradigme de l’apprentissage par renforcement (RL) adapté au traitement de problématiques de

1. Notons par ailleurs que les travaux de la communauté RO qui portent sur l’optimisation de l’ordonnement dynamique d’un atelier flexible (DFJSSP) [3] sont aussi dans une certaine mesure d’intérêt.

décision séquentielle. Elles reposent dans ce cas la plupart du temps sur des techniques avancées de modélisation à base de processus de décision markoviens et d'apprentissage (profond) [1]. L'état de l'art souligne qu'elles offrent par ailleurs des performances très encourageantes voire intéressantes dans certains contextes difficiles à aborder avec les approches standard [2] - au prix, rappelons-le, de la perte d'un nombre important de propriétés séduisantes et parfois importantes, e.g. garanties de convergence.

Nos travaux portent sur l'analyse, la comparaison et l'hybridation de méthodes issues de la RO et de l'IA pour la résolution de problèmes d'allocation de ressources en temps réel au sein de systèmes dynamiques à files d'attente d'intérêt pour le monde économique.

2 Exemple de problème

Nous étudions, dans un premier temps, une formulation simplifiée d'un problème d'intérêt du domaine, reposant sur une modélisation en temps discret et sur un ensemble de tâches unitaires, c'est-à-dire non décomposées en séquences d'opérations intermédiaires. Le système comporte plusieurs ressources hétérogènes, chacune étant capable de traiter uniquement certains types de tâches, avec des temps de service incertains dépendant du type de tâche considéré.

Le problème satisfait un ensemble de contraintes structurelles : (i) chaque ressource ne peut traiter qu'une seule tâche à la fois ; (ii) chaque tâche ne peut être prise en charge que par une unique ressource ; (iii) le traitement d'une tâche est non préemptif ; et (iv) les ressources sont supposées disponibles sur l'ensemble de l'horizon d'allocation. Les arrivées d'utilisateurs et leurs tâches associées, ainsi que les temps de traitement, sont décrites par des lois de probabilité connues, permettant de caractériser la dynamique stochastique du système.

L'objectif de nos travaux est de développer des approches d'optimisation intégrant des composantes RO et IA, capables d'améliorer les performances du système selon différents indicateurs clés, et en particulier le temps total passé par un utilisateur dans le système (temps d'entrée moins temps de sortie). Ces travaux visent à concevoir des stratégies décisionnelles réactives et robustes, adaptées aux environnements à forte variabilité et à ressources hétérogènes.

3 Contribution

Nous comparons les approches de l'état de l'art au regard de différentes configurations de test, et de modélisations de politiques appartenant au domaine de la RO et de l'IA. Les évaluations réalisées, et entraînements des politiques à base d'apprentissage, sont effectués dans des contextes de simulations à temps discret configurées au regard de problématiques d'intérêt de notre partenaire ESII. Les résultats de comparaison seront exposés lors de la conférence.

Références

- [1] Haozhe Chen, Ang Li, Ethan Che, Jing Dong, Tianyi Peng, and Hongseok Namkoong. Qgym : Scalable simulation and benchmarking of queuing network controllers. *Advances in Neural Information Processing Systems*, 37 :92401–92419, 2024.
- [2] Jim Dai and Mark Gluzman. Queueing Network Controls via Deep Reinforcement Learning. *Stochastic Systems*, 12(1) :30–67, March 2022.
- [3] Candice Destouet, Houda Tlahig, Belgacem Bettayeb, and Bélahcène Mazari. Systematic review and future directions in dynamic flexible job shop scheduling : a decade of research. *Journal of Intelligent Manufacturing*, October 2025.