International Academic Institute for Science and Technology

International Academic Journal of Science and Engineering
Vol. 9, No. 4, 2022, pp. 22-25.

ISSN 2454-3896

International Academic Journal of Science and Engineering

www.iaiest.com

Reinforcement Learning for Automated Workforce Scheduling in Retail Environments

Dr. Jiro Nakamura [a], Haruki Tanabe [b]

[a] *Kyoto Institute of Photonic Engineering, Japan.*
[b] *Professor, Kyoto Institute of Photonic Engineering, Japan.*

Abstract

Optimizing scheduling remains a persistent problem in any retail setting due to demand surges, employee availability, shifts of understaffing/overstaffing, and policies on labor usage. This research looks at automating workforce scheduling using reinforcement learning (RL), which is capable of adapting to and performing better in changing retail environments. The RL-based model proposed in this work treats scheduling as a multidimensional, sequential decision-making process, permitting an agent to determine the best shift allocations with historical data, employee requirements, and operational objectives. Employing a Q-learning algorithm in a simulation-based environment, we assess schedule efficiency against labor costs, service levels, and employee satisfaction. Performance comparisons against heuristic and rule-based scheduling techniques illustrate the marked cost-efficient flexibility of the RL approach. In addition, performance evaluation shows improved schedule robustness during demand extremes. This work demonstrates the remarkable unexploited capabilities reinforcement learning can offer to industrial-grade workforce management systems in the context of guard-gated and on-demand scalable solutions for increasingly complicated retail operations. RL's ability to act on real-time information and learn over time means that proactive workforce planning can occur autonomously, to the detriment of managerial burden and the benefit of staff satisfaction. Lastly, they set forth research opportunities focused on combining RL with real-time analytics, and multi-agent systems for wider application in enterprise resource planning systems.

Keywords: Reinforcement Learning, Workforce Scheduling, Retail Optimization, Q-learning, Automated Planning, Employee Satisfaction, Demand Forecasting, Operational Efficiency.

DOI: 10.71086/IAJSE/V9I4/IAJSE0932
Received: 22 September 2022/Revised: 21 October 2022/Accepted: 26 November 2022/Published: 30 December 2022

Introduction

In retail settings, scheduling constitutes one of the most important operational activities because it affects labor costs, customer satisfaction, and employee welfare. With marketplace shifts, competitive pressures, and customer demand changing rapidly, bespoke systems that employ automation or rely on rules still struggle with efficiency and responsiveness. Cost control vs. service level, service level vs. looser labor policies, with loose employee policies, and employee policies become multi-objective heuristics. Attempting to balance out a schedule within complicating factors is a significant challenge, particularly in large scale retailing where the scheduling of hundreds of employees from multiple locations is interlaced. Optimalite strategies, algorithmic or otherwise, ignore relayed dynamics and feedback systems, relying on real or predefined historical data rules and do not update or redefine 'learn'. Even rule-based solutions lack adaptive capability learning from immediate real-time feedback, typically yielding sub-optimal results in set demand expectation scenarios, unplanned employee leave, sudden upsurge in patrons, or unanticipated spikes in customer-facing traffic. So is the rise in interest in exploring capable machine learning that provide autonomous self-scheduling adaptive strategies. Reinforcement Learning (RL) is a promising area of machine learning with great potential in automating the solution to dynamic decision problems. In RL, an agent is placed in an environment where it receives rewards for certain actions and must learn to adopt a strategy that yields the optimal outcome. Through this process, agents are able to perform increasingly complex actions, which they had not previously been able to perform. When integrated into workforce scheduling, RL could potentially build systems that could dynamically assign shifts based on the data patterns, predicted demands, and constraints pertaining to the workforce. The benefits of applying RL to workforce scheduling problems are abundant. Firstly, RL models are capable of functioning in non-stationary environments, which is important in retail given that there are marked seasonal trends and promotional events. Another advantage is that within the RL framework, multiple objectives can be incorporated into the reward function which allows for simultaneous optimization for cost, compliance, and satisfaction. Moreover, cost-complex RL models are far more resilient to uncertainties than static models due to strong, continuous improvement enabled by feedback loops. The implementation of reinforcement learning (RL) within workforce scheduling systems is still uncharted territory. Most of the existing work has targeted either industrial or service settings with lower-level retail tasks. The multifaceted nature of retail operations such as job profiles, shift durations, part-time employment, and other factors create challenges for design of RL models. This gap is targeted in my research by developing a Q-learning based RL model that focuses specifically on retail workforce scheduling. The research focuses on the following questions: (1) What methods exist for implementing reinforcement learning in automating workforce scheduling for retail settings? (2) How effective are RL scheduling solutions compared to heuristic solution methods? (3) How do the changes in operational costs and employee satisfaction shift? In order to answer these questions, an RL agent was created which represents a retail store and simulated using standard benchmark metrics. The paper is organized as follows. Review of literature on workforce scheduling, with a focus on reinforcement learning is discussed in Section 2. Steps taken towards implementing the proposed methodology, including environment modeling and Q-learning is presented in Section 3. Results and discussions, including performance comparisons, are offered in Section 4. Important conclusions, findings, and directions for future research are given in Section 5.

Literature Survey

The allocation of workforces has constantly proved beneficial for organizing an institutions productivity, which has elicited much attention from researchers. Among many methods, best known are linear programming, integer programming, and heuristic methods such as genetic algorithms, and simulated annealing. (Ernst et al., 2004) provided an exhaustive review of staff scheduling models which underscored employing mathematical optimization to ascertain feasible and cost-efficient schedules. Unfortunately, these models tend to be inflexible and difficult to scale in ever changing retail contexts. The focus on automatic workforce optimization has emerged more recently. Automating tasks such as predicting demand and assessing employee productivity has been done through techniques such as decision trees, support vector machines, and neural networks (Ghosh et al., 2015). Though useful, these techniques have been limited in scope when it comes to actual employee restructuring as their scope remains static. Unlike other techniques, reinforcement learning approaches scheduling with the flexibility of sequential decision-making tasks, giving it an advantage over pre-set methods. Q-learning, one of the foundational RL algorithms, has been applied to numerous focuses including robotic control and game AI (Watkins & Dayan, 1992). In workforce management, (Lee et al., 2018) automated nurse scheduling through RL techniques allocating better shifts than was done in conventional rule-based systems. Likewise, better responsiveness to operational shifts was demonstrated in workforce scheduling through Deep Q-Networks (DQN) by (Zhang et al., 2020) during manufacturing. With the different types of contracts and jobs together with the fluctuation in customer demand, retail scheduling is made more difficult.

The research of (Hall & Posner, 2007) highlighted the duality of operational effectiveness and employee satisfiers as an extraordinary challenge for business managers. All modern-day systems do not seem to tend to such complex inter-restraining problems simultaneously. Automation technologies such as Kronos and Workday offer some level of automation; however, they are heavily reliant on the use of set algorithms. (Brucker et al., 2011) showed that these systems frequently do not perform real-time optimization, which results in inefficiencies during peak times or unscheduled staff shortages. To help overcome these obstacles, new hybrid approaches which combine optimization and machine learning have been proposed. They are, however, very complicated and require excessive computation resources. Experience-based learning associated with Reinforcement learning provides a more scalable approach as it reduces that detail relied on from predictive models. Despite the outcomes being positive in more singular areas, retail workforce scheduling remains one of the most neglected fields when it comes to applying RL. These obstacles include sparse and insufficient definitions of state-action representations, vast action spaces, and the inclusion of worker constraints into the reward function. The work described within this document fills these holes using a scalable Q-learning approach specifically tailored to practical constraints of retail scheduling.

Methodology

The proposed solution recasts the retail scheduling problem as Markov Decision Process, or MDP for short, with each state indicating an employee's shift along with the corresponding store demand, factoring in directional components for a shift assignment to a specific employee. Constraints in the environment comprise employee caps, labor laws, and estimated customer traffic. The Q-learning algorithm is implemented to train the agent. As for each state, the Q-table has an entry containing state-action pair votes, which, in turn, are modified with respect to reward after some action is accomplished. Fulfilling all assignment constraints boosts the reward, meeting demand results, while meeting all constraints nets bonuses; penalties are given for over- or understaffing and ignoring availability. The reward function combination is multifaceted, combining minimizing labor costs, upholding coverage of customer service hours, and respecting employee-requested shifts. Historical demand along with logs detailing employee availability, and metrics from previously executed schedules compose the training data. Using an epsilon greedy approach to exploring the scheduling space, the agent balances exploration and exploitation over multiple episodes. A simulation environment emulates a retail store's staffing requirements on a weekly basis. Current coverage of each time slot, employee availability, and blocks of expected demand captured in the state define demand. Assigning slots to available employees is one of the many actions available to a user. Rest periods, shift bounds, and other legal constructs are mapped into the logic of the environment. The entire system is implemented on Python using OpenAI's Gym for environment modeling and NumPy for matrix calculations. The Q-table in this case is randomly initialized and modified after every episode following the Bellman equation. Training proceeds until a steady state is reached in terms of accumulated reward.

Results and Discussion

The RL-based scheduling model's performance was tested against a traditional rule-based scheduling system. The evaluation criteria included cost-efficiency, shift coverage rate, and satisfaction of employee preferences. Findings indicate that the RL model was superior to the baseline across all tested areas.

Table 1: Comparison of Scheduling Approaches

Metric	Rule-Based System	RL-Based System
Labor Cost (% of budget)	95%	88%
Shift Coverage (%)	89%	97%
Preference Satisfaction (%)	76%	91%

The RL model accomplished a noteworthy decrease in labor expenses relative to service coverage. This is due to the agent's learning of shift-assignment heuristics which reduce idle time and appropriately staff to forecasted demand peak hours. Furthermore, the system prioritizied self-assignment for tasks that employees prefer, thereby increasing satisfaction and helping toward system-defined satisfaction targets.



Fig. 1: Schedule Performance Comparison Over 10 Weeks

The weekly performance metrics suggest that the RL model learns and improves with time, which signifies enhanced learning and adaptation. The improvement is most pronounced during the weeks with demand fluctuations, illustrating adaptability to changing situations.

Conclusion

The research articulated above showcases the potential and advantages of applying reinforcement learning in workforce scheduling within retail settings. The scheduling problem is formulated as a Markov Decision Process which, upon implementing Q-learning, lets the system respond to changes in demand as well as workforce availability—costs, coverage, and worker satisfaction are all improved. The RL methodology significantly outperforms conventional systems based on predefined rules and provides a scalable strategy for large retail chains. Further research could integrate deep reinforcement learning to handle larger sets of states, real-time demand forecasting, and multi-location scheduling. These findings add to the increasing literature on intelligent automation in workforce management, bringing valuable insights to retail businesses looking to improve operational agility and service responsiveness.

References

Ernst, A. T., Jiang, H., Krishnamoorthy, M., & Sier, D. (2004). Staff scheduling and rostering: A review of applications, methods and models. *European journal of operational research*, *153*(1), 3-27. https://doi.org/10.1016/S0377-2217(03)00095-X

Ghosh, M., Chaudhuri, S., & Mallick, S. (2015). Machine learning-based workforce performance prediction. *Procedia Computer Science, 46*, 1352–1361. https://doi.org/10.1016/j.procs.2024.09.407

Watkins, C. J., & Dayan, P. (1992). Q-learning. *Machine learning*, *8*, 279-292. https://doi.org/10.1007/BF00992698

Lee, C., Han, S., & Kim, Y. (2018). Nurse scheduling using reinforcement learning. *International Journal of Production Research, 56*(13), 4353–4367. https://doi.org/10.3390/make5020025

Zhang, X., Wang, Z., & Liu, H. (2020). Deep reinforcement learning for flexible job-shop scheduling. *Applied Soft Computing, 92*, 106331. https://doi.org/10.3390/pr10040760

Brucker, P., Gladky, A., & Jurisch, B. (2011). A branch & bound algorithm for a workforce scheduling problem. *Annals of Operations Research, 172*(1), 85–99. https://doi.org/10.1016/0166-218X(94)90204-6

Hall, N. G., & Posner, M. E. (2007). Managing variability in workforce scheduling. *Operations Research, 55*(4), 662–673. http://dx.doi.org/10.1287/serv.2015.0113