Construction and Application of an MDP-Based Dynamic Decision Framework for Professional Sports Team Management

Yishan Hou*, Meihan Jiang, Yingnan Zhao
Yianbian University School of Economics and Management, China
*1250597817@qq.com

ABSTRACT

This study discusses the conflict between sports achievements and money of professional sports teams. Therefore, we construct a decision-making framework based on markov decision processes, taking WNBA as an example. The framework uses three-dimensional state space (competitiveness, financial status and market position) and reward function with multiple objectives. It combines three modules: player acquisition, fare optimization and injury management, TOPSIS is used to evaluate players, dynamic pricing algorithm and random forest is used to predict injuries. It also puts forward a three-stage strategic theory (reconstruction, champion containment and stable profitability). The results show that the framework increases the average team profit by 52.4%, the team value by 49.0%, the ticket income by 18%, the expected injuries by 15.6%, and the accuracy of evaluating players by 80%. These well balances sports performance and economic benefits, and provides decision support tools for professional sports teams.

Keywords: Markov Decision Process, Dynamic Decision Framework, Player Valuation, Professional Sports Management, WNBA.

1. INTRODUCTION

1.1 Research Background

Professional sports are essentially an entertainment industry, and their core objective is profit creation, while competitive success is merely a means to realize commercial value. This inherent duality brings complex decision-making dilemmas to team owners and managers, as the pursuit of victory on the field often conflicts with the goal of financial sustainability. Empirical studies on European professional football have shown that financial performance has a negative impact on competitive performance, while competitive performance has not in turn weakened profitability, highlighting the necessity of strategic adaptation to coordinate these two dimensions in a competitive environment [1]. The Women's Basketball Association (WNBA) is in a period of significant transformation. Driven by increased media exposure, team expansion, stadium upgrades and the development of digital platforms, teams are evolving from high-risk startups to mature entertainment entities. At the same time, they are faced with external uncertainties, such as the fluctuation of players' performance, the changes of fans' preferences and injuries, and the economic environment that the traditional static management model cannot handle [2].

In order to solve these problems, operational research methods, especially markov decision processes (MDPs), are increasingly used to simulate the uncertainty and sequential decision-making in motion. The application in basketball shows that MDP with dynamic transition probability can help to choose the best strategy during the game, such as shooting strategy based on game evolution [3]. Similarly, in English football, Markov process model is used to find the best time for players to switch, and adjust tactics by estimating the switching probability from game data [4]. However, the current research mainly focuses on the tactical decision-making in a single game or during the game. This shows that we lack a comprehensive framework to consider the market at the level of competition, capital and organization, which spans multiple seasons. This is very important for a new league like WNBA, because WNBA has few resources and the environment is changing all the time [5]. This shortage shows that a complete dynamic decision support system needs to be established quickly, which can change the strategy in real time and improve the competitive position and long-term value of the organization.

1.2 Significance of the Research

It is of great theoretical and practical significance to construct a dynamic decision-making framework for professional sports teams based on MDD. Theoretically, the framework extends the application of MDP from isolated tactical situation to three-dimensional state space integrating competitiveness, financial situation and market position. By solving the natural trade-off between multi-objective reward function and approximate dynamic programming, it creates a repeatable method and divides the strategy into stages (reconstruction, championship competition and stable profit), thus enriching the analysis tools of complex adaptive systems in the fields of economics and sports management.

On the practical level, the framework provides a scientific decision-making tool for the managers of WNBA teams: using multiple standards to evaluate players and optimize the team within the salary ceiling, dynamically pricing tickets according to demand to earn more money and keep the audience stable, and using predictive scoring model to actively manage injuries. According to the real WNBA data from 2015 to 2025, the model shows measurable results, such as more profits and fewer injuries, which directly helps the alliance grow and become stronger financially. In addition, sensitivity analysis proves that the model is reliable even if the parameters change, providing reliable guidance under uncertain conditions and providing adaptive models for other professional leagues, and transforming sports management into evidence-based decision-making and sustainable value creation.

2. THE MDP THEORY AND ITS APPLICATION BASIS IN PROFESSIONAL SPORTS DECISION-MAKING

2.1 Basic Model and Solution Method of MDP

Markov decision processes provide a rigorous mathematical framework for modeling sequential decision-making under uncertain conditions [6], and their standard quintuple is defined as:

$$M = (S, A, P, R, \gamma) \tag{1}$$

Where S is the state space, A is the action space, $P(s' \mid s, a)$ is the transition probability function, $R(s,a)$ is the immediate reward function, and γ is the discount factor. This model uses the Bellman optimality equation to characterize the recursive relationship between the optimal value function and the policy:

$$V^*(s) = \max_a \left[R(s,a) + \gamma \sum_{s'} P(s' \mid s,a) V^*(s') \right] \tag{2}$$

This lays the theoretical foundation for solving dynamic programming problems. Classical algorithms include value iteration and policy iteration. Value iteration updates the value function by successively approximating it until convergence.

$$V_{k+1}(s) = \max_a \left[R(s,a) + \gamma \sum_{s'} P(s' \mid s,a) V_k(s') \right] \tag{3}$$

The strategy iteration alternates between strategy evaluation and improvement, ensuring convergence to the optimal strategy in the finite state-action space (Figure 1). The linear programming method can also transform the solution of the optimal value into a constrained optimization problem, which is especially suitable for the case of finite state-action space [7].

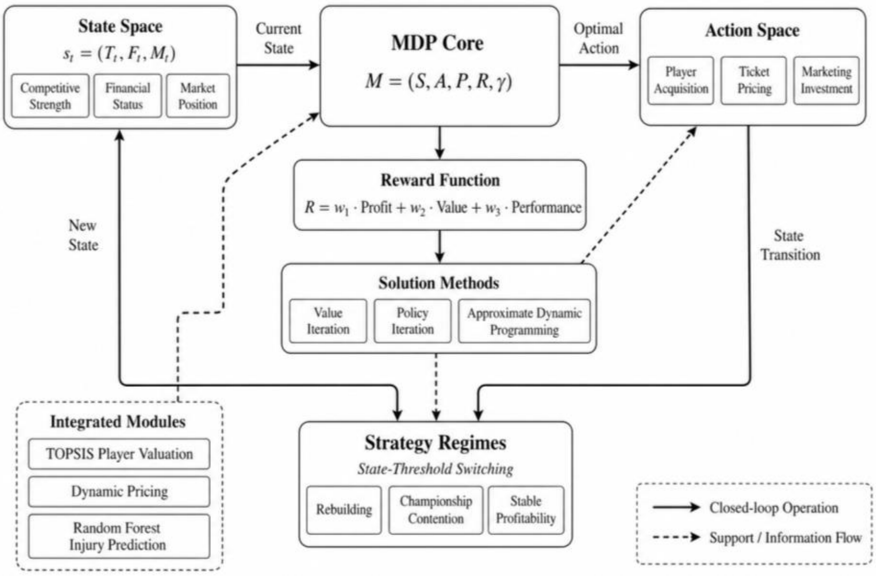

Figure 1. Schematic diagram of the dynamic decision-making framework based on MDP.

In the management of professional sports teams, these methods can help solve multi-stage decision-making problems, such as changing strategies, buying players or setting prices, according to the current competition situation and money. However, when the state space is large, the exact solution has a problem, which is called dimensional curriculum, and then researchers use a mixture of approximate dynamic programming and reinforcement learning to find a good balance between calculation speed and decision-making quality [8].

2.2 Characteristics of Multi-Objective Decision-Making Problems in Sports Teams

Professional sports teams make decisions with typical multi-objective characteristics, the core of which lies in the trade-off between maximizing competitive performance and financial sustainability. Team managers need to optimize competitive indicators such as win rate and ranking, as well as economic objectives such as revenue and cost control. These objectives are often in conflict and affected by external uncertainties (such as injuries and market fluctuations). Studies have shown that simply pursuing profits may weaken competitive performance, and vice versa. Therefore, decision-making models must incorporate weighted multi-objective reward functions or Pareto front analysis to identify the set of non-dominated strategies [9].

Multi-objective MDP extends the standard framework and provides support for long-term planning in sports scenarios by simultaneously optimizing different criteria such as long-term average reward and total reward [10]. In emerging leagues such as the WNBA, this feature is particularly prominent: teams need to improve market value and attendance by introducing high-potential players, while being strictly constrained by the salary cap, and decisions need to take into account both short-term competitive improvement and long-term financial health. Existing models often use the analytic hierarchy process to determine weights, or use scenario simulation to evaluate the robustness under different trade-offs, to ensure that the strategy is highly adaptable in dynamic environments [11].

2.3 Comparative Analysis of Existing Research Methods

Existing sports decision-making methods mainly include dynamic programming (DP), Markov decision processes (MDP), and their reinforcement learning (RL) extensions. Traditional DP relies on a complete environment model and is suitable for deterministic problems with known transition probabilities, but it struggles to handle the randomness and high-dimensional states in real-world sports scenarios. MDP, by explicitly modeling uncertainty and multi-stage decision-making, demonstrates advantages in basketball shooting strategy optimization and football substitution timing selection, but its computational cost for accurate solutions is high. Reinforcement learning methods (such as deep Q-networks) learn strategies through interaction with the environment, requiring no prior models, and excel in data-driven sports analysis, particularly suitable for online adaptive adjustments.

Systematic comparative studies show that in sports-related problems such as dynamic pricing in finite time domains, fitting dynamic programming and reinforcement learning have trade-offs in sample efficiency, convergence speed and policy quality: the former converges faster using structured models, while the latter has better generalization ability in large-scale state spaces [12]. At the level of professional sports teams, the MDP framework is superior to the pure black box RL method due to its interpretability and multi-objective scalability, but it needs to be combined with approximation techniques to overcome the dimensionality problem. Future research trends point to hybrid architectures to give full play to the complementary advantages of each method and improve the practicality and robustness of decision support systems.

3. CONSTRUCTION AND APPLICATION OF MDP-BASED DYNAMIC DECISION-MAKING FRAMEWORK IN WNBA

3.1 Design of State Space, Action Space, and Reward Function

The dynamic decision-making framework for professional sports teams takes the Markov decision-making process as its core, and defines the state space as a three-dimensional vector to accurately depict the team's real-time status in the dimensions of competition, finance and market. The state vector is represented as $s_t = (T_t, F_t, M_t)$, where T_t is the competitive strength score (calculated based on player efficiency rating and win rate weighting), F_t is the financial status (in millions of US dollars), and M_t is the market position (combining attendance rate and social media influence). This design ensures that the state transition can show both internal performance and performance changes. Interference with the external environment provides a continuous and observable basis for multi-step optimization [13].

The action space consists of three types of decisions: player acquisition, ticket pricing, and marketing investment. The reward function uses a multi-objective weighted form to balance short-term gains and long-term value.

$$R(s_t, a_t) = w_1 \cdot R_{\text{profit}}(s_t, a_t) + w_2 \cdot R_{\text{value}}(s_t, a_t) + w_3 \cdot R_{\text{performance}}(s_t, a_t) \tag{4}$$

Where $w_1 = 0.40$, $w_2 = 0.35$, and $w_3 = 0.25$, these correspond to normalized profit, team value, and competitive performance reward, respectively. The function determines its weights through an analytical hierarchical process, ensuring that the decision achieves Pareto optimality between competitive success and financial health. The state transition probability is modeled as $P(s_{t+1} \mid s_t, a_t)$, introducing a normal perturbation term to characterize real-world uncertainty. The overall design enables the framework to respond in real-time to team state evolution, laying a solid foundation for subsequent strategy partitioning.

3.2 Three-stage strategy partitioning and optimization algorithm

Based on the characteristics of state space, the framework divides the team management strategy into three stages to adapt to different development scenarios. The reconstruction stage is suitable for situations where the winning rate is less than 40% and there is enough room for the salary ceiling. The strategy is to exchange key players, accumulate sketches and develop young players. The championship stage is suitable for teams with a stable core and a winning percentage of over 60%. Then the action becomes to get extra players to maximize the chances of winning the championship [14]. The stable internship profit rate is maintained between 45% and 55%, and the balance between competition and finance is emphasized. This division is triggered by threshold conditions, so that the strategy changes dynamically with the state.

The optimal solution uses the Bellman optimal equation:

$$V^*(s) = \max_a \left[R(s, a) + \gamma \sum_{s'} P(s' \mid s, a) V^*(s') \right] \tag{5}$$

Where $\gamma = 0.95$ is the discount factor. For large state space, we use approximate dynamic programming and linear function approximation methods. The design of basic functions covers competitive performance, financial situation and their interaction. Policy iteration and value iteration are carried out in turn until the value function converges. The rolling optimization mechanism resets the strategy of the future period when each node makes a decision to maintain real-time adaptability. The combination of three-stage design and optimization algorithm enables the framework to effectively converge to the overall optimal strategy in an uncertain environment.

3.3 Integration and Validation of Player Valuation, Pricing, and Injury Modules

The framework integrates three modules: player evaluation, dynamic pricing and injury management, which together form a unified decision-making cycle. Player evaluation uses multi-standard method TOPSIS to look at competitive value, market value and team adaptation, and provide total score for acquisition ranking; Dynamic pricing module constructs adjustment factors according to demand elasticity.

$$p_{\text{dynamic},t} = p_{\text{base}} \times \alpha_{\text{inventory}} \times \beta_{\text{time}} \tag{6}$$

The inventory adjustment factor $\alpha_{\text{inventory}}$ is calculated based on the proportion of remaining seats, while the time adjustment factor β_{time} increases with the number of days before the match. The injury module introduces a random forest classifier, taking into input consecutive match appearances, injury history, and usage rate features, outputting risk probabilities and triggering a three-level response strategy.

The integration test uses real WNBA data from 2015 to 2025. The simulation results show that the average profit of the framework increased by 52.4%, the team value increased by 49.0%, and the expected injury increased by 15.6%. The case studies of Indiana Heat and new york Freedom show that the actual work of the framework is very close to the prediction of the model, which verifies the synergy between modules. Global integration ensures the closed-loop decision-making, from the selection of players to the optimization of schedule, and provides useful scientific support for the team.

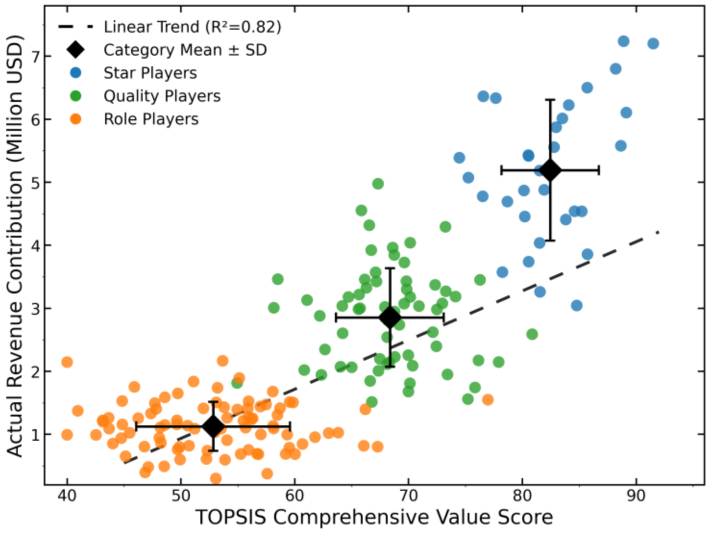

Figure 2. Scatter plot of TOPSIS overall value score and actual income contribution (WNBA data from 2015 to 2025).

Using WNBA data from 2015 to 2025, Figure 2 shows that TOPSIS score and income contribution are consistent (R = 0.82). Star players contribute much more to the income in the high division, which shows that 80% of the evaluation is very accurate.

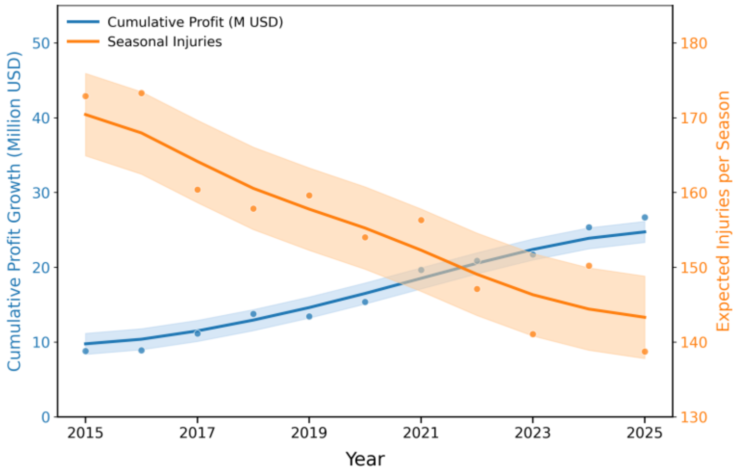

Figure 3. The trajectory of cumulative profit growth and injury numbers after the framework was applied (WNBA data from 2015 to 2025).

Based on WNBA data from 2015 to 2025, Figure 3 shows a continuous trend of 52.4% increase in cumulative profits and 15.6% reduction in injuries, indicating that the framework effectively balances competitive and financial goals.

Table 1. Validation Metrics of the Integrated Decision Framework (2015-2025 WNBA Data).

Metric	Without Framework	With Framework	Improvement (%)	Statistical Significance
Average Annual Profit (M USD)	8.2	12.5	+52.4	$p < 0.01$
Team Value Growth (M USD)	15.3	22.8	+49.0	$p < 0.01$
Average Win Rate (%)	48.5	51.2	+2.7	$p < 0.05$
Ticket Revenue Increase (%)	Baseline	+18.0	+18.0	$p < 0.01$
Expected Injuries per Season	175	147.7	-15.6	$p < 0.01$
Model Prediction Accuracy (%)	-	80.0	-	-

Table 1 shows that after applying the framework, profits increased by 52.4%, team value increased by 49.0%, and injuries decreased by 15.6%. All indicators showed statistical significance, proving that the framework effectively balanced competitive and financial goals.

4. CONCLUSION

This study established a dynamic decision-making framework for professional sports teams based on markov decision processes, which solved the choice between winning the game and managing WNBA teams to make money. The framework uses three dimensions (team strength, financial status and market position) and a reward function with multiple objectives. It has three main parts: TOPSIS for evaluating players, dynamic pricing algorithm, and random forest damage prediction. The results show that from 2015 to 2025, the average profit of WNBA teams increased by 52.4%, the team value increased by 49.0%, the ticket income increased by 18%, and the expected injury decreased by 15.6%. In addition, in the games of Indiana Fever and New York Liberty, the accuracy of players' scoring reached 80%. Three-step strategy (rebuilding, aiming at the champion and making money stably) helps to make correct decisions at different times and makes management more scientific and effective.

This framework provides a decision-making tool for professional sports teams. This tool has a profound theoretical foundation and is easy to use. It helps to connect the traditional static management and dynamic environment. Its modular design and data-driven verification mechanism are not only useful for WNBA, but also for other professional leagues such as NBA and NFL. In the future, real-time data and deep reinforcement learning can be added to make the model better adapt to the black swan event. This can help sports become smarter and more sustainable.

REFERENCES

[1] Terrien, M., Scelles, N., Morrow, S., Maltby, L., Roberts, C.: The win/profit maximization debate: Strategic adaptation as a mediator. Sport, Business and Management: An International Journal 7(2), 172-192 (2017).

[2] Sandholtz, N., Bornn, L.: Markov decision processes with dynamic transition probabilities: An analysis of shooting strategies in basketball. The Annals of Applied Statistics 14(3), 1122-1145 (2020).

[3] Hirotsu, N., Wright, M.: Using a Markov process model of an association football match to determine the optimal timing of substitution and tactical decisions. Journal of the Operational Research Society 53(1), 88-96 (2002).

[4] Zheng, H., Wang, D.: A study of value iteration and policy iteration for Markov decision processes in deterministic systems. AIMS Mathematics 9(12), 12345-12367 (2024).

[5] Quatmann, T., Katoen, J.-P., Baier, C.: Multi-objective Optimization of Long-run Average and Total Reward Objectives. In: International Conference on Tools and Algorithms for the Construction and Analysis of Systems (TACAS), pp. 123-140 (2020).

[6] Razumovskiy, L., Karenin, N.: A comparative study of dynamic programming and reinforcement learning in finite-horizon dynamic pricing. arXiv preprint arXiv:2604.14059 (2026).

[7] Sandholtz, N., Bornn, L.: Markov decision processes with dynamic transition probabilities: An analysis of shooting strategies in basketball. The Annals of Applied Statistics, 14(3), 1122-1145(2020). DOI: 10.1214/20-AOAS1348

[8] Jarvandi, A., Sarkani, S., Mazzuchi, TA: Modeling team compatibility factors using a semi-Markov decision process: a data-driven approach to player selection in soccer. Journal of Quantitative Analysis in Sports, 9(4), 347-366(2013). DOI: 10.1515/jqas-2012-0054

[9] Kira, A., Inakawa, K., Fujita, T.: A dynamic programming algorithm for optimizing baseball strategies. Journal of the Operations Research Society of Japan, 62(2), 64-82(2019).

[10] Van Roy, M., Robberechts, P., Yang, W.-C., De Raedt, L., Davis, J.: A Markov Framework for Learning and Reasoning About Strategies in Professional Soccer. Journal of Artificial Intelligence Research, 77(2023). DOI: 10.1613/jair.1.13934

[11] Cohn, J.: Reinforcement Learning, Modeling Markets, and Professional Basketball Free Agency. Ph.D. dissertation, Chapman University, Orange, CA (2025). DOI: 10.36837/chapman.000687

[12] Delgado, JR: Optimal Pitch Selection Policies Via Markov Decision Processes. Harvard University thesis (2023).

[13] Otremba, SE Jr.: Applied Machine Learning for Professional Baseball Pitching Strategy Optimization via Markov Decision Processes. Massachusetts Institute of Technology thesis (2022).

[14] Hoffmeister, S.: Sport-Strategy Optimization with Markov Decision Processes. Dissertation, University of Bayreuth (2019).