A Dynamic Framework for Optimizing Reward Policies in the Sharing Economy

Angela Li
Applied Mathematics & Physics
Stony Brook University
Stony Brook, NY
angela.li.4@stonybrook.edu

Chengkun Yao, Cheng Li
Katz School of Science and Health
Yeshiva University
New York, NY
cyao1@mail.yu.edu, cli3@mail.yu.edu

David Li
Katz School of Science and Health
Yeshiva University
New York, NY
david.li@yu.edu

***Abstract*—This study introduces a dynamic framework for optimizing reward policies in the sharing economy, leveraging game theory to enhance user participation, improve resource utilization, and sustain platform growth. By analyzing strategic interactions among participants, the research employs game-theoretic models, including Nash equilibrium, to optimize resource allocation and ensure equitable outcomes. The proposed framework addresses the limitations of static reward models by developing dynamic reward policies that adapt to real-time data and evolving user behaviors. Using a dynamic programming algorithm, the study formulates the reward allocation problem as an optimization challenge, recursively selecting users to incentivize in a way that maximizes platform efficiency and profitability within budget constraints. Scenario-based analyses validate the framework's effectiveness, demonstrating its ability to adapt to varying market conditions and user behaviors. The integration of game theory ensures that reward policies are not only effective in driving engagement but also fair and scalable, offering actionable strategies for designers and policymakers to unlock the full potential of sharing economy ecosystems.**

***Index Terms*—Game Theory, Nash Equilibrium, Probability, Dynamic Programming, Optimization.**

I. Introduction

The sharing economy has emerged as a transformative force in the modern socio-economic landscape, reshaping the way individuals and communities access and utilize resources [1]. Rooted in the principle of shared consumption, the sharing economy offers a framework where goods, services, and skills are exchanged on a temporary basis, often facilitated by technology-driven platforms. By emphasizing access over ownership, this model enables a dynamic shift in resource utilization, promoting both economic and environmental sustainability.

At its core, the sharing economy fosters resource efficiency and reduces the environmental footprint by minimizing overconsumption. Simultaneously, it creates new opportunities for social engagement and additional income generation, enhancing the livelihoods of individuals across diverse demographics. For instance, platforms facilitating ride-sharing, home rentals, and gig-based tasks empower users to monetize underutilized assets and skills while meeting the demands of their communities [2].

Several key players exemplify the success of the sharing economy across various sectors. Uber has disrupted traditional transportation models through its app-based ride-sharing services, while Airbnb has revolutionized the accommodation industry by enabling property owners to connect directly with short-term renters. TaskRabbit facilitates efficient service exchanges for day-to-day tasks, and Freelancer provides a global marketplace for professional freelance work. These platforms demonstrate the breadth and adaptability of the sharing economy, as well as its potential to redefine conventional industries.

In the sharing economy ecosystem, sustainability is heavily reliant on the active engagement of resource providers [3]. These providers often exhibit interdependencies within the same geographic area, creating correlations in their participation levels. This dynamics is further influenced by external factors, such as fluctuations in the job market and broader economic conditions [4], which can introduce significant instability to the sharing economy.

This paper seeks to investigate the mechanisms driving the success of the sharing economy, focusing on the role of reward policies in enhancing resource optimization, sustainability, and socio-economic growth. By examining the design and impact of reward policies, this research aims to develop strategies for optimizing rewards to further stimulate participation and growth within the sharing economy.

The remainder of this paper is organized as follows: Section II reviews related work and highlights the limitations of existing approaches. Section III details the proposed reward optimization design, including its mathematical formulation and implementation. Section IV presents scenario-based analysis. Section V provides the algorithm to determine the value of α. Finally, Section VI concludes the paper and discusses potential directions for future research.

II. Related Work

The sharing economy has attracted significant attention from researchers across multiple disciplines, including economics, sociology, and computer science. Studies have highlighted the transformative potential of sharing platforms in enabling resource optimization and promoting sustainable consumption. However, the success of these platforms often hinges on user participation, engagement, and effective resource allocation, which are influenced by reward policies [5].

A. Reward Mechanisms in the Sharing Economy

Prior research has explored various reward mechanisms to enhance user participation. For example, static reward models have been used in platforms like Uber and Airbnb to incentivize providers and balance supply-demand dynamics. These approaches, while effective in certain scenarios, often fail to adapt to changing user behaviors and market conditions. Dynamic reward policies, which adjust incentives based on real-time data, have demonstrated greater efficacy in sustaining user engagement and optimizing resource utilization. However, challenges persist in designing these policies to account for fairness, transparency, and the diverse motivations of platform participants [6].

B. Behavioral Economics and User Engagement

Behavioral economics provides a valuable framework for understanding user behavior in sharing economy platforms. Studies [7] have shown that users are influenced not only by monetary rewards but also by social recognition, platform trust, and ease of use [8]. Incorporating these behavioral factors into reward design has been shown to significantly improve user retention and participation rates. However, integrating these insights into dynamic reward policies remains an area of ongoing research.

C. Game Theory and Resource Optimization

Game-theoretic models [9] have been widely applied to analyze strategic interactions among participants in sharing economy platforms. These models provide insights into pricing strategies, competition, and cooperation among users. For instance, Nash equilibrium concepts have been used to optimize resource allocation and ensure equitable outcomes. Despite their utility, many game-theoretic models rely on simplifying assumptions, limiting their applicability to real-world, dynamic sharing ecosystems.

D. Machine Learning in Reward Policy Design

The advent of machine learning has opened new avenues for designing adaptive reward policies. Techniques such as reinforcement learning [10] and multi-armed bandit algorithms [11] have been employed to optimize incentives based on user interactions and platform performance. These methods enable platforms to dynamically adjust rewards in response to evolving user behaviors and market trends. However, balancing computational complexity with scalability and interpretability remains a challenge.

E. Limitations of Existing Approaches

While substantial progress has been made, existing methods often fall short in addressing the complexities of dynamic, multi-user environments. Many reward policies are designed for specific platforms, limiting their generalizability across different contexts. Additionally, the lack of integration between behavioral insights, game theory, and machine learning models restricts the development of holistic reward frameworks.

III. The Proposed Reward Optimization Design

A. Variables under Dynamic Nash Equilibrium Assumption

In this subsection, we define the key inputs and concepts used throughout the methodology. These notations form the basis for modeling the incentive allocation problem, where the objective is to determine how incentives should be distributed among resource providers to maximize engagement while balancing cost efficiency for the platform.

Consider a grocery delivery and pick-up service platform where customers place orders, and shoppers in the region fulfill these orders by purchasing the items from stores and delivering them to customers. When the platform faces a high volume of orders but a limited number of shoppers, it may increase cash incentives to motivate shoppers to take on more orders. Conversely, when there are too few orders and an excess of shoppers, the platform may reduce incentives to save costs. In both scenarios, the platform must ensure that a minimum percentage of orders is fulfilled to maintain customer satisfaction, while simultaneously minimizing incentive spending. This balance is crucial for optimizing operational efficiency and sustaining platform growth.

Assume that at a given time step t, shoppers in this region adopt their own optimal strategies based on the assumption of an instantaneous Nash equilibrium. Importantly, this Nash equilibrium is not static but dynamic, evolving over time [12]. By the next time step $t+1$, we can reasonably assume that the Nash equilibrium condition remains largely valid, albeit with a slight shift due to changes in shopper behavior, market conditions, or platform dynamics. This dynamic equilibrium reflects the continuous adaptation of shoppers to the evolving environment while allowing the platform to optimize the reward policies adaptively on daily basis. At the time step t, let's define,

- N is the total number of available shoppers
- i is the ith shopper
- c_i is the incentive (cash) that the ith shopper will take
- p_i is the probability that the ith shopper will take the incentive c_i
- x_i where $x \in [0, 1]$, 1 means to advertise the incentive c_i to the ith shopper, and 0 means not.

Notes:

- In reality, the incentive can be a continuous range with probability density. Then the summation below will become integration on distribution.
- For now, for simplicity, this study assumes each shopper having a single incentive and the associated probability.
- The incentive and the associated probability at time step t can be predicted by machine learning models or by statistics on historical data.
- c_i could be zero, for the shoppers who don't need incentive
- p_i could be 1, for the shoppers who will definitely take the order.

B. One-Time Purchase (OTP) vs Shoppers

This subsection explores the relationship between one-time purchases (OTP) and the prediction of shopper behavior. By modeling OTP as a function of the number of participating shoppers, the analysis examines how shopper engagement influences OTP outcomes. The discussion also considers edge cases where highly active shoppers could skew this relationship, emphasizing the importance of balancing shopper participation to optimize OTP rates.

Assume OTP is a monotonically increasing function of shoppers, i.e., more shoppers, higher OTP; less shoppers, lower OTP.

Generally speaking, this assumption is true, except that some shoppers may take many orders, then higher OTP doesn't necessarily imply more shoppers.

Then we can think there is a function

$$OTP = O(n) \tag{1}$$

where n is the required number of shoppers for this OTP.

C. Constraints

This subsection outlines the critical constraints that govern the reward optimization process. It emphasizes maintaining minimum OTP fulfillment percentages, adhering to shopper engagement thresholds, and operating within budget limits. These constraints ensure the reward policy's effectiveness while balancing financial feasibility and platform objectives.

Usually OTP has a minimum percentage O_L to fulfill, for example, $O_L = 81\%$.

$$O(n) \geq O_L \tag{2}$$

which implies that n has a lower bound N_L too, according to the monotonical assumption above.

$$n \geq N_L \tag{3}$$

which further implies the below

$$\sum_{i=1}^{N} p_i \times x_i \geq N_L \tag{4}$$

Total spent reward has a budget upper bound B_H

$$\sum_{i=1}^{N} c_i \times p_i \times x_i \leq B_H \tag{5}$$

However, the total advertised incentive could be higher than B_H. This is not a constraint.

$$\sum_{i=1}^{N} c_i \times x_i > B_H \tag{6}$$

The platform only cares about the actual spent incentive, i.e., the amount that the shoppers take, so we multiply the probability to have the actual spending.

D. The Objectives

The ultimate goal is to have the highest profit within the incentive budget and maintain the minimum OTP fulfillment percentage.

Alternatively, we want the highest OTP within the budget

$$\max_{n \in [0,N]} O(n) \tag{7}$$

which is equivalent to solve all integer x_i, in order to

$$\max_{x \in [0,1]} \sum_{i=1}^{N} p_i \times x_i \tag{8}$$

with the constraint

$$\sum_{i=1}^{N} c_i \times p_i \times x_i \leq B_H \tag{9}$$

E. Dynamic Programming Optimization

Assume each shopper is independent individual shopper. We can recursively advertise to the best next shopper, until we hit the budget. Now it becomes a subproblem: How to define and find the next optimal shopper?

Let's assume we now already advertised to M shoppers. The task is to pick the best $(M+1)$-th shopper from the remained list.

For simplicity without loss of generality , let's assume the M shoppers have the index from 1 to M.

The M shoppers associated with their probability count to the actual number of shoppers by:

$$\sum_{i=1}^{M} p_i \times x_i \tag{10}$$

If we pick the j-th shopper from index $M+1$ to N, then, $x_j = 1$, the contribution to the actual total number of shoppers is

$$p_j \times x_j, \text{ which is } p_j \tag{11}$$

Then the total actual shoppers become

$$\sum_{i=1}^{M} p_i \times x_i + p_j \tag{12}$$

The marginal contribution to the current OTP is

$$\left.\frac{\partial O(n)}{\partial n}\right|_{n=\sum_{i=1}^{M} p_i \times x_i + p_j} \tag{13}$$

which can be estimated by the OTP vs shoppers prediction using linear interpolation or cubic spline interpolation for smooth curve.

Numerically, the marginal contribution of the jth shopper toward the OTP is

$$\delta O_j = O(\sum_{i=1}^{M} p_i \times x_i + p_j) - O(\sum_{i=1}^{M} p_i \times x_i) \tag{14}$$

Futher assume there is a multipler α that represents the dollar value of profit for OTP. It can be as simple as a constant factor, or can be a function of OTP. I will describe how to estimate the reasonable value of it later.

So the dollar amount contribution by advertising to the jth shopper is

$$\alpha \times \delta O_j \tag{15}$$

And the marginal actual cost toward the budget is

$$c_j \times p_j \times x_j, \text{ which is } \delta B_j = c_j \times p_j \tag{16}$$

In order to pick the best j-th shopper, from all the remained shoppers from $M+1$ to N, we want the j-th shopper to bring the best profit by choosing the j so that

$$\max_{j \in [M+1,N]} (\alpha \times \delta O_j - \delta B_j) \tag{17}$$

Rewrite to,

$$\max_{j \in [M+1,N]} (\alpha \times [O(\sum_{i=1}^{M} p_i \times x_i + p_j) - O(\sum_{i=1}^{M} p_i \times x_i)] - c_j \times p_j) \tag{18}$$

And the marginal actual cost must be under budget, which is

$$c_j \times p_j \le B_H - \sum_{i=1}^{M} c_i \times p_i \times x_i \tag{19}$$

where $x_1 = x_2 = x_3 = ... = x_M = 1$ which comes to the optimal solution of the subproblem by simply picking the shopper with the highest profit within budget. However, the best profit can still be negative.

To get global optimal solution, this process recursively continues if any one of below conditions is true,

- OTP is below O_L and the marginal actual cost is under budget. In this condition, the shopper may bring negative profit.
- OTP is above O_L and from the remained list we can find a shopper that brings max positive profit and the marginal actual cost is under budget

This algorithm also works for soly maximizing OTP by simply assigning α a very big number like $\$1,000,000$ so that the marginal profit from each shopper is always positive, i.e, any increment, big or tiny, on OTP is worth to pursue.

IV. Scenario-Based Analysis

To evaluate the effectiveness of the proposed reward optimization framework, a scenario-based analysis was conducted using simulated data. The analysis focused on understanding the interplay between user engagement, incentive allocation, and overall platform performance under varying conditions. Key metrics included one-time purchase (OTP) fulfillment rates, total incentive costs, and marginal profit contributions.

The relationship between OTP rates and shopper participation was examined to identify the optimal number of incentivized users. As expected, OTP rates increased monotonically with the number of participating shoppers, albeit with diminishing returns at higher participation levels, as shown in Fig. 1. The analysis highlighted a critical threshold beyond which additional incentives yield negligible gains in OTP fulfillment.

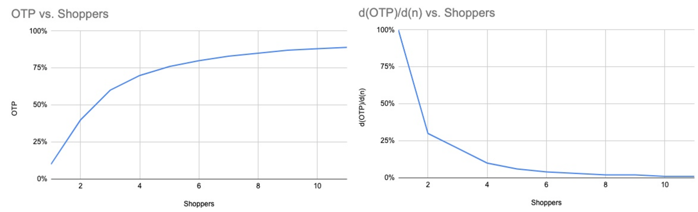

Fig. 1: OTP fulfillment percentage v.s. the number of shoppers

This finding underscores the importance of targeting the right users to maximize efficiency. To ensure financial feasibility, the reward policy was tested under varying budget constraints. The results demonstrated that the proposed dynamic optimization algorithm effectively allocated resources to maximize OTP rates within the given budget. Specifically, the algorithm prioritized high-impact users, who contributed significantly to platform performance while requiring minimal incentive costs. Scenarios with tighter budgets revealed the algorithm's ability to maintain OTP fulfillment rates by strategically reallocating rewards. A detailed analysis of marginal cost versus marginal

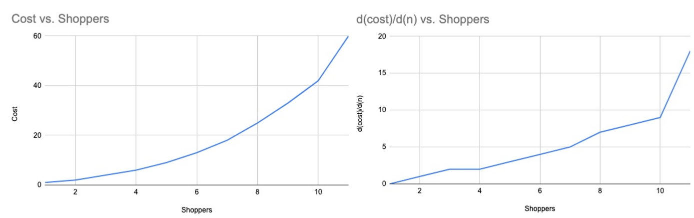

Fig. 2: Total incentive cost v.s. the number of shoppers

profit in Fig. 3 showed the algorithm's capability to identify the optimal stopping point for incentivization. By continuously evaluating the dollar value of OTP contributions relative to the cost of incentives, the model ensured that rewards were allocated only when they generated positive returns as shown in Fig. 4. The goal is to keep adding shoppers until $\frac{d(\text{OTP})}{d(n)}$

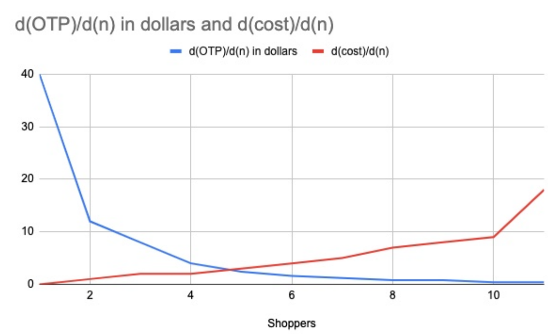

Fig. 3: Marginal OTP and Marginal Cost

in dollars meets $\frac{d(\text{cost})}{d(n)}$, which makes the marginal profit zero for the last added shopper. In the best senario, the incentive spending is still below budget and we don't intend to use all budget. If the last added shopper still brings positive marginal profit, we will keep adding until running out of budget as shown in Fig. 5. In scenarios where minimum OTP thresholds were unmet, the model extended incentivization to

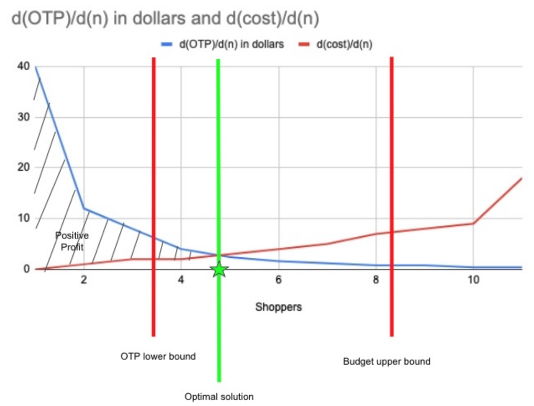

Fig. 4: Choose the Next Optimal Shopper

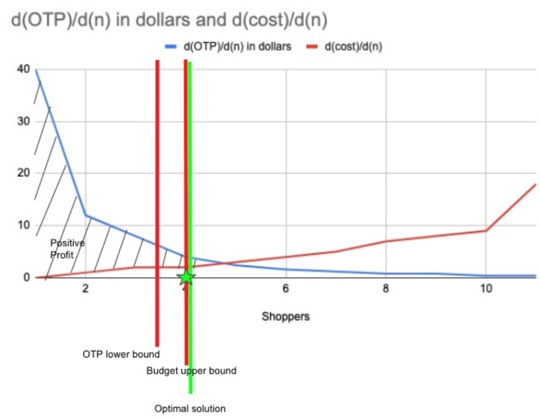

Fig. 5: Add More Shoppers

less profitable users to ensure compliance, even at a slight loss. In Fig. 6, if the last added shopper brings negative marginal profit, but we are still below the required OTP, then we still add shoppers until we meet the minimum OTP, even we start losing money.

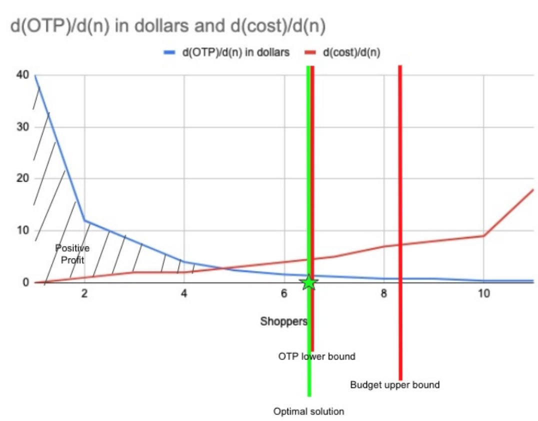

Fig. 6: Keep Adding Shoppers at a Loss

V. Determine the Value of α

If we don't know the value of α or we just want to maximize OTP by using all incentive budget, we can simply assign big value to α and apply the same algorithm, which basically shifts up the curve of $\frac{d(\mathrm{OTP})}{d(n)}$ so that the cross point moves to the far right side of budget upper bound, as shown in Fig. 7.

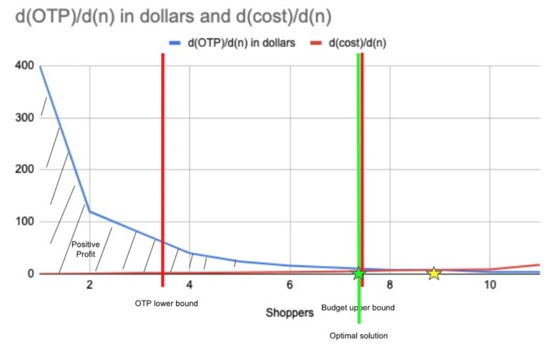

Fig. 7: Maximize OTP by Using All Incentive Budget

A. Precise Value of α is Unnecessary

In the last scenario, we need a reasonable value of α.

- Different α values may lead to different "optimal solutions". Only reasonable value of α will lead to reasonable optimal solution.
- Small drift of α value won't significantly change the optimal solutions since in every step we simply pick the shopper with the highest profit within budget, i.e, the ranking of shopper's profitability matters, not the actual marginal profit. So we can say, the algorithm and the solution have some level of invariance to small change of α.
- So we only need to estimate a reasonable α in order to execute the above algorithm. It doesn't have to be very accurate but should be accurate enough in the right range.

B. Risk Neutral Pricing Methodology

In order to estimate a reasonable value of α, we introduce the risk neutral pricing assumptions:

- The business suggests the incentive budget from their best empirical data and knowledge.
- On average, spending all the incentive budget will be expected to maximize the profit.

With the above assumptions, one can calibrate the reasonable value of α. In order to find the α that satisfies the above assumptions, the optimal solution must be at the cross point of the two curves, and it has to overlap with the budget upper bound as shown in Fig. 8. In other words, in the ideal situation, when we spend the last dollar of our budget, the marginal profit just comes down to zero. However, normally it won't be the case by the initial guess of α, but we can adjust α to make the above condition happen by the following steps. Fig. 9 shows the process.

- Execute the algorithm until we spend all budget.
- Then based on the d_1 and d_2 in Fig. 9, we can adjust α to be

$$\alpha \Leftarrow \frac{d_2}{d_1} \times \alpha \tag{20}$$

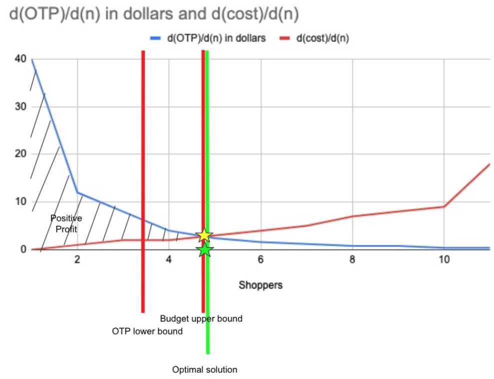

Fig. 8: Estimate a Reasonable Value of α

- Which parallelly shifts up or down the OTP curve so that the cross point will be right on the budget upper bound
- This update on α may change the optimal solution because the incentive curve will be slightly different depending on the optimal shopper selection, so we need to run the algorithm from scratch again. Repeat the process until
$$\left|\frac{d_2}{d_1} - 1\right| \leq \epsilon \tag{21}$$
where ϵ is a user-defined small and positive tolerance value used to determine when the algorithm should stop iterating. Apparently, when $d_2 = d_1$, α will not be updated any more, then the algorithm stops.
- Then we get both the reasonable α and its optimal solution simultaneously. This estimated α can be specific for each incentive optimization problem, or can be an average estimated value from historical data. Since the solution has some level of invariance to small change of α, the optimal solution should be stable without knowing the accurate α.

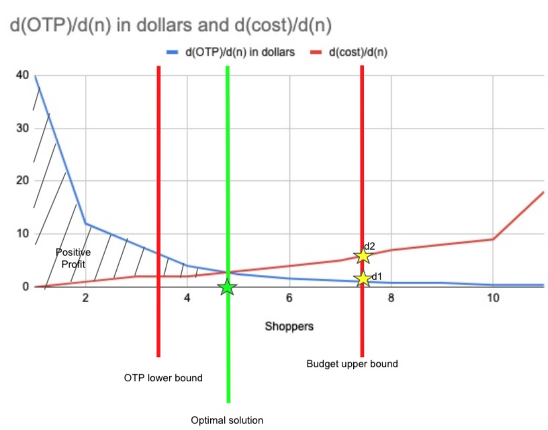

Fig. 9: Find the Optimal Value of α

VI. Discussion and Conclusion

The computational efficiency of the framework enables its scalability to large platforms. Since platform activities are region-specific, optimization can be performed independently for each region, where the number of customers and service providers remains manageable. This decentralized approach allows for large-scale parallel computing in a cloud-based infrastructure, where data is partitioned by region, and optimization processes run in parallel to generate optimal reward policies.

By integrating game theory, dynamic programming, and behavioral insights, the proposed framework ensures that incentive mechanisms are not only effective in enhancing user engagement but also fair, scalable, and adaptable to changing market dynamics in a manner that optimizes platform efficiency and profitability. By dynamically adjusting to real-time fluctuations in user behavior and market conditions, the framework helps maintain equilibrium between supply and demand.

While the model demonstrates strong effectiveness within a grocery delivery platform, its adaptability extends to various sharing economy applications. Leveraging dynamic Nash equilibrium and machine learning-driven predictions, the framework remains scalable and applicable across diverse ecosystems within the sharing economy.

The findings highlight the significance of dynamic incentive allocation and the role of behavioral considerations in designing reward strategies. Future research should focus on real-world implementation, deeper machine learning integration, and long-term sustainability to further refine and expand the framework's applicability and impact.

References

[1] Sundararajan, A. (2016). The Sharing Economy: The End of Employment and the Rise of Crowd-Based Capitalism. MIT Press.
[2] Dinara Davlembayeva, Savvas Papagiannidis, Platform-provider relationship dynamics in the sharing economy: Challenges and implications, Industrial Marketing Management, Volume 111, 2023, Pages 242-256
[3] Acquier, A., Daudigeos, T., & Pinkse, J. (2017). Promises and Paradoxes of the Sharing Economy: An Organizing Framework. Technological Forecasting and Social Change, 125, 1–10.
[4] Frenken, K., & Schor, J. (2017). Putting the Sharing Economy into Perspective. Environmental Innovation and Societal Transitions, 23, 3–10.
[5] Mi, Z., Coffman, D. The sharing economy promotes sustainable societies. Nat Commun 10, 1214 (2019). https://doi.org/10.1038/s41467-019-09260-4
[6] Marco Guerriero, Veronika Megler, Oliver Boom, and Sokratis Kartakis, Preventing customer churn by optimizing incentive programs using stochastic programming, 2020, AWS Machine Learning Blog
[7] Zervas, G., Proserpio, D., & Byers, J. W. (2017). The Rise of the Sharing Economy: Estimating the Impact of Airbnb on the Hotel Industry. Journal of Marketing Research, 54, 687-705. https://doi.org/10.1509/jmr.15.0204
[8] Gneezy, U., Meier, S., & Rey-Biel, P. (2011). When and Why Incentives (Don't) Work to Modify Behavior. Journal of Economic Perspectives, 25(4), 191–210.
[9] Chang, S.-L.; Lee, K.-C.; Huang, R.-R.; Liao, Y.-H. Resource-Allocation Mechanism: Game-Theory Analysis. Symmetry 2021, 13, 799. https://doi.org/10.3390/sym13050799
[10] Mnih, V., Kavukcuoglu, K., Silver, D., et al. (2015). Human-Level Control Through Deep Reinforcement Learning. Nature, 518(7540), 529–533.
[11] Kuleshov, V., & Precup, D. (2014). Algorithms for Multi-Armed Bandit Problems. arXiv preprint arXiv:1402.6028.
[12] Weibull, J. W. (1995). Evolutionary Game Theory. MIT Press.