MEAL: Model of Empathy Augmented Logistics for Food Security

Seoyeong Park
North Carolina State University
Raleigh, NC, United States
spark43@ncsu.edu

Munindar P. Singh
North Carolina State University
Raleigh, NC, United States
mpsingh@ncsu.edu

ABSTRACT

Millions globally lack access to nutritious food, experiencing food insecurity. Despite growing recognition and efforts, equitable distribution remains hard due to limited resources and the diverse preferences of users. The problem of equitable food distribution requires a robust, socially intelligent infrastructure. We propose Meal (Model of Empathy Augmented Logistics for Food Security), a novel approach based on a multistakeholder recommendation system for food allocation, balancing user needs and society's sustainability. Unlike existing systems that focus on either user satisfaction or logistical operations, our model considers both users and food banks, dynamically adapting to changing preferences and resource availability. Our simulated experiments demonstrate improvement in metrics such as over single-stakeholder models, suggesting significant potential for improved food access and resource utilization in addressing food insecurity.

KEYWORDS

food security, multiagent system, reinforcement learning, simulation

ACM Reference Format:
Seoyeong Park and Munindar P. Singh. 2024. MEAL: Model of Empathy Augmented Logistics for Food Security. In *Proc. of the 23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2024), Auckland, New Zealand, May 6 – 10, 2024*, IFAAMAS, 8 pages.

1 INTRODUCTION

Food insecurity remains a critical global concern, affecting millions of people across the world who struggle to access nutritious food on a daily basis [15]. USDA [15] specifies that 12.8% of households (equivalent to 17 million) in the US experience food insecurity.

The food bank system in the US is an established approach to distributing limited food to individuals and households in need. Yet, ensuring efficient and equitable distribution is difficult when supplies are limited, and preferences are diverse. Though the main goal of food banks is to alleviate food insecurity, recipients deserve to choose what they eat from available options, considering their dietary constraints. However, food banks today provide only limited personalization.

Previous research in solving food-related problems has focused on either logistical operations to make efficient distribution [8, 9] or user-centric perspectives recommending items of users' tastes [18], but not tackled both aspects simultaneously. We propose a novel approach for recommending food items that consider users and the food bank. This multistakeholder approach aligns more closely with the intricate nature of food insecurity. It holds the potential to relieve food insecurity and unfairness by leading users to adjust their preferences if doing so is better for efficient distribution.

Our approach, Meal (Model of Empathy Augmented Logistics for Food Security), encourages flexibility in preferences and optimizes for model of distribution by considering users' needs along with the inability of the food allocation. This dual focus yields more users access to the food they need while avoiding imbalanced distribution and waste. Meal accounts for changing preferences by users and resource availability by local food banks (providers), unlike many state-of-the-art personalized recommendation systems that assume the environment is static. This dynamic approach benefits both the providers and users, thus improving the societal impact.

1.1 Motivation

The provider's objective is to ensure the efficient allocation of available food resources. This includes minimizing food waste, optimizing food distribution, and supporting the broader community's needs while supplying food items that align with user preferences. It would be best if everyone could get the food items that they most prefer. However, resources are limited, and it is impossible to always allocate what perfectly matches preferences. Meal employs social and psychological factors to elicit empathy from users for them to accept allocated items.

We introduce the idea of *prosociality*. Prosociality fosters social responsibility and encourages users to promote social welfare even though their individual goals are not completely satisfied. In our setting, prosocial behavior is seen in conceding preferred foods to others by voluntarily accepting less preferred foods in limited supply. This means recommending choices that may not be ideal for users or providers but promote fairness and overall well-being for both. To achieve this, Meal incorporates insights from cognitive architecture and social norms, mimicking the realistic decision-making processes of humans and creating a more prosocial environment.

Social welfare. Social welfare aggregates the users' and providers' benefits from a resource allocation. Efficient allocation, considering societal gains while meeting user needs effectively, leads to enhanced societal productivity, increased satisfaction of stakeholders, and potentially reduced waste and operational costs.

Fairness. We model prosocial decision-making by incorporating *fairness*, referring to societal outcomes based on food distribution. In human societies, people can figure out each other's valuations to develop a sense of prosociality, which posits that people may be self-interested, but their decisions are affected by how relatively poorly others fare [6]. Fairness captures the intuition that all users have an equal opportunity to receive satisfying food allocations. It prevents

Proc. of the 23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2024), N. Alechina, V. Dignum, M. Dastani, J.S. Sichman (eds.), May 6 – 10, 2024, Auckland, New Zealand.

certain users from being consistently favored over others. Some users who need a particular item more may not get the essential item when the item is limited in quantity. Fairness contributes to the overall well-being of the community by promoting the distribution of food. This is crucial in scenarios where the goal is promoting not just individual well-being but also social well-being.

Trust. Food preferences are not static but change depending on one's personal situation. Thus, recommending food items that match preferences is important to promote trust in the food bank system. Our model should be transparent in its decision-making for the recommendations it generates. It also needs to be aware of the long-term consequences of recommendations.

1.2 The Food Distribution Setting

Food banks are nonprofit organizations that seek to alleviate food insecurity and promote well-being in communities. Around the world, food banks are dedicated to collecting, storing, and distributing food to people in need.

Figure 1 summarizes the heart of a complex ecosystem realized in the US involving government, donors, local food banks, volunteers, and the households they serve. Government support includes funding, food supplies, and capabilities to procure, store, transport, and distribute food [15]. Food banks partially rely on donations from donors, organizations, and retailers. Local food banks (providers) serve as intermediaries between regional food banks and the communities they support, ensuring that food reaches the right hands. They receive food supplies from food banks or directly from donors, sort, and distribute them to users. Each provider may reflect the needs of its community, adjusting the quantity and types of food provided to match local preferences and dietary requirements. Users are individuals and households experiencing food insecurity. They seek assistance from food banks, often traveling to designated locations to receive food under the current system.

Maximizing satisfying users while allocating resources is challenging: a solution must satisfy user needs and preferences while minimizing food waste and satisfying capacity and other constraints.

1.3 Contribution and Approach

Meal benefits stakeholders by learning user preferences and making usable food available for the community. Our experimental results show that the proposed method achieves reduced waste and increased satisfaction in distributing food items compared to models that consider only one side, either users or providers. The contributions of this paper are as follows:

- We propose a novel value-aware multistakeholder framework that maximizes user satisfaction and social benefits using Q-learning. We formulate the task as a multiobjective optimization problem of empathetic personalized recommendations.
- Our framework incorporates a dynamic multistakeholder context (users and providers), unlike many studies that assume a static context or a single stakeholder.

We conduct an experiment with a simulated dataset to verify that our model finds a way of recommending preferred options and enhancing overall prosociality.

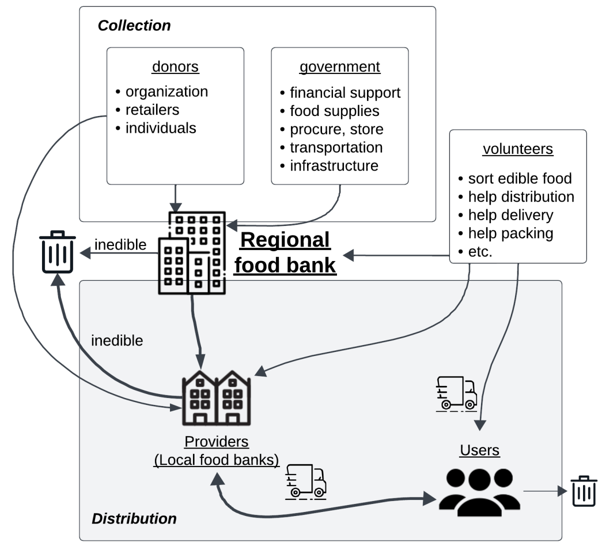


Figure 1: Food distribution ecosystem, based on the US setting

2 RESEARCH ON MULTISTAKEHOLDER SYSTEMS

Unlike traditional recommendation systems, multistakeholder recommendation systems address the objectives and needs of more than one party. These systems can help overcome limitations in incorporating system objectives that apply across stakeholders. For example, Sürer et al. [14] seek not only to maximize user preferences but also to consider providers in multisided platforms. Other studies apply this concept to consider social values and business operations, such as fairness and profitability [1, 4, 11]. Similarly, in an educational setting, Zheng et al. [19] propose a utility-based multistakeholder recommender system to address balancing the needs of students and instructors. Moreover, Ghanem et al. [7] propose an agent-based simulation for an e-commerce recommendation system that considers business profit and consumers with a linear model.

One prevalent approach to implementing recommendation systems is to treat them as Markov Decision Problems (MDPs) and employ reinforcement learning (RL) algorithms [13]. Recent studies leverage RL to handle multiple objectives of recommendation systems [2, 5, 12]. For example, Kwak and Huettel [10] apply RL to establish a decision-making paradigm for oneself and charity and understand differences in prosocial tendencies. However, despite the diverse applications of multistakeholder RL in recommendation systems, it has not been applied to address the problem of food insecurity.

3 PROBLEM FORMULATION

For simplicity, we choose to focus on the interplay between the core stakeholders of users and providers.

In the scenario we consider, the food distribution system is similar to a personalized recommendation system. For example, a user living with a child interacts with a food bank system where the user expresses her preferences for different food items and her profile. Let's say the user prefers to have fresh fruits and vegetables, milk, and whole grains. An app, which contains all inventory information provided by the food bank, recommends items from such categories that she might like, for example, a bag of apples, fat-free milk, and a box of granola based on the user's profile and past selections. The user then determines whether to accept the recommended items and indicates her satisfaction with the accepted items. This feedback helps the app adapt to improve its suggestions. However, pleasing the user is not the only goal of the food bank. Since food banks need to allocate food and maintain their inventory while reducing operation costs and waste, the app does not always recommend items that are the user's favorites. For instance, if apples are in short supply or highly requested by others, the app might suggest oranges to ensure a fair distribution of popular items, even though oranges are not the user's top choice. Doing so helps ensure as many people as possible get what they need and keeps the food bank running smoothly.

Thus, users and providers have different perspectives on food items when it comes to product recommendations. Meal recognizes complexity by placing two primary stakeholders with distinct objectives: the users, who have their own needs and preferences, and the providers, who have community desires. The notation we used in this paper is summarized in Table 1.

3.1 Stakeholders and Their Objectives

We now describe our two main types of stakeholders.

3.1.1 Users. Users are individuals served through our recommendation system. Their goal is to acquire food items that align with their preferences and needs. This user-centric perspective emphasizes the importance of enhancing user satisfaction and personalized experiences for food allocation.

User preferences. As users interact with the system, their preferences for food items are constantly captured and refined. These preferences evolve over time and are shaped by factors such as age, health status, dietary constraints, household status, and willingness to make prosocial choices. The agent learns these dynamics by reflecting user feedback toward recommended food items. This learning process allows the agent to provide recommendations matching a user's tastes and current needs. A user u's preferences are represented by a list of ratings ranging from 0 (no preference or experience) to 5 (extremely like) toward a food item f. That is, a user u provides satisfaction h toward recommended food items. Therefore, we define cumulative user satisfaction H as below:

$$H = \sum h_{u,f}$$

3.1.2 Providers. Ensuring the effective distribution of available food resources is the main goal of providers. This entails reducing food waste, maximizing the distribution of food, and meeting the needs of their community while providing food items that suit user preferences. The provider prioritizes not merely using in-stock items but also fulfilling user requests as closely as possible. However, they might propose less-preferred alternatives when necessary. The provider intends to trigger empathy and gently nudge users to accept alternatives that are allocated by utilizing social and psychological factors that influence decision-making.

Provider's benefit. The provider has a measure of goodness for allocation influenced by points such as:

- C_1 Perishable items have short shelf lives. In particular, items that demand refrigeration affect storage capacity. Conversely, nonperishable (canned or dry) goods can be held in stock at a lower cost.
- C_2 Food items with limited quantities may need to be kept for users with priority needs, such as for infants.
- C_3 Food with high quantities may need to be consumed faster under lower demand. Food items with excessive quantities must be distributed quickly before they expire and to avoid taking up space.

Providers earn more benefits if they consume food items with higher scores. The provider's benefit c is determined by the aggregate score of the food items when the allocation to u is accepted. These scores are updated in real-time as allocations are made.

$$C = \sum c_{u,f}$$

3.1.3 Objective for Decision Making. Our primary objectives are simultaneously maximizing user satisfaction and maximizing the provider's benefit. The agent understands the values of stakeholders, the future state of the world for each action it can perform, and the social experience its user will derive for each action it can perform. Then, since we cannot maximize both objectives, the agent moderates to achieve an optimal trade-off between two stakeholders. To balance these objectives, a weighted sum of user satisfaction H and provider benefit C is used with a weighting factor denoted as ω ($0 \le \omega \le 1$). We choose the optimal value of ω^* that minimizes the difference between H and C (Equation 2). Therefore, the agent's overall reward for the decision-making objective is a weighted combination of satisfaction and provider benefit as Equation 1.

$$\sum R = \omega * H + (1 - \omega) * C \quad (1)$$

$$\omega^* = \arg\max_{\omega}(H, C) \quad (2)$$

The parameter ω ranges between completely provider-focused valuation ($\omega = 0$) and completely user-focused ($\omega = 1$). H and C are updated each time a particular recommendation is taken.

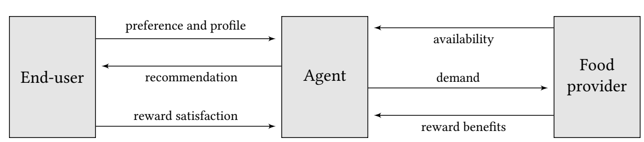


Figure 2: Model architecture

3.2 Model Design

We formally define our problem setup in this section. We have a set of users U and a set of food items F, where each user in U has profile information and unique food preferences toward each food item in F, captured in a matrix of numerical ratings, $P : U \times F \longrightarrow [0, 5]$. Each item in F carries attributes that reflect its importance in consumption priority and benefits to the provider. These attributes include multiple factors, such as inventory capacity, expiration date, and perishability, shaping the provider benefits $c_{u,f}$ associated with each recommendation happening at time step t. Within this dynamic framework, $d_u \in D$ represents a recommendation for user u at a specific time step t. It contains two attributes: a recommended food item and a binary indicator of whether it is accepted. Subsequently, we assume that user satisfaction $h_{u,f}$ comes as ratings. The problem involves finding the optimal way to distribute the available food to users over time while considering their preferences and impact on the community, in other words, managing the trade-off between these two objectives. Table 1 includes the notation we used in this paper.

Table 1: Notation

Notation	Description
U	a set of users
F	a set of food items
$P : U \times F$	a matrix of ratings from user u to food f
$p_{u,f} \in P$	a numerical rating toward food f from user u; $p \in \mathbb{Z}, 0 \leq p \leq 5$
$D : F \rightarrow U$	a set of recommendation from F to U
$d_u \in D$	a recommendation for user u at time step t
$h_{u,f}$	user satisfaction from u at time step t; $h \in \mathbb{Z}, 0 \leq s \leq 5$
H	cumulative user satisfaction
$c_{u,f}$	provider benefits for a given recommendation at time step t
C	cumulative provider benefit
$\omega \in [0, 1]$	weighting factor between user satisfaction and provider benefit
e_u	estimated prosociality of u
α	learning rate; fixed to 0.1
β	weighting factor for the probability of acceptance; fixed to 0.9
γ	discount factor; fixed to 0.9
ϵ	exploration rate; fixed to 0.1

Formally, we define the above problem as a Partially Observable Markov Decision Process (POMDP) where an agent (recommender) interacts with the environments (users and food provider) over time to maximize cumulative rewards of combined benefits. $\langle S, A, T, R, O, \Omega, \gamma \rangle$, where $s \in S$ is a finite set of states (i.e., user preferences and profiles, inventory status), $a \in A$ is a finite set of actions (i.e., the possible recommendations), T is a set of transition probabilities between states (i.e., the probability of acceptance), O is a set of observations (i.e., whether the recommendation is taken or not, users' satisfaction feedback), Ω is a set of conditional observation probabilities of receiving an observation $o \in O$ after taking action $a \in A$ at state s, R is a reward function (i.e., a combination of user satisfaction and provider's benefit from accepted recommendations controlled by the weighting factor ω, as defined in Equation 2), and $\gamma \in [0, 1)$ is the discount factor.

By using Q-learning [17], our model effectively adapts to dynamic changes in users' needs, food availability, and other factors and incorporates long-term interaction into their decision-making process.

Figure 2 and Algorithm 1 describe how our model operates.

4 EXPERIMENTAL SETTING

We evaluate our model through simulations to understand how prosocial decisions are made throughout interactions. The simulated environment comprises data consisting of three sets: user profiles, preference ratings, and food inventory. Since it is hard to acquire real-world food preference data and food bank availability, we arbitrarily approximated the values of food items in our simulation by seeding the survey results of food pantry needs [3]. Our approximation method provides a realistic representation of the effectiveness of our model and mimics the interactive environment and dynamic user behaviors. The simulation runs 1,000 episodes and each episode of the simulation begins with initialized data and agents. During the episode, the model recommends food items, which the user accepts or rejects. The model recommends one item at a time, potentially resulting in the same recommendations multiple times based on availability. Each episode terminates when the inventory becomes empty or reaches a predefined number of steps.

4.1 User Profile and Prosociality

The main agents in our model are the users. We have crafted a user community with unique profiles. For simplicity, each user's profile includes age, whether they have dietary restrictions or disease, family size, and ratings towards food items). We set 33% of users as aged over 65 and 45% of users as having a child. The distribution of family size followed a survey statistic: the mean is three, and the standard deviation is two [3]. Table 2 shows two examples of user personas. User 1 is a single-person household with no allergies or any types of dietary restrictions and is in the healthiest condition. User 2 is a family of five members, at least one of whom is a child. However, nobody in the family has any allergies or any types of dietary restrictions.

Table 2: Sample user profiles

	User 1	User 2
age	22	47
family size	1	5
has allergies or restrictions	no	no
has child(ren)	no	yes

A user may accept or reject a recommendation. The probability of acceptance hinges on two factors: how much the recommended item matches preference and the user's inherent willingness to yield (Equation 3). Users don't know how much the provider gains when they accept or reject recommendations. Ratings for particular items may be undefined. We estimate their potential satisfaction level with the most similar user's preferences using cosine similarity if no value is assigned for a particular item.

$$\text{Probability of acceptance} = \beta * p_{u,d_u} + (1 - \beta) * e_u \tag{3}$$

Willingness to yield can be affected by random factors. To quantify user prosociality, we break user profiles into three age groups, three family sizes, a binary indicator of whether they have dietary restrictions, and a binary indicator of whether they have one or more children (Table 3). We set assumptions to estimate prosociality.

- A_1 Younger adults are more likely to make prosocial choices.
- A_2 Single-person households are more altruistic and flexible than multi-person households.

Algorithm 1 Q-Learning Based Food Recommendation Agent

Initialize Q-table with zeros
Parameters: learning rate, discount factor, exploration rate, weight
Inputs: $s \in S : U, F, P$
for episode **do**
  Retrieve U and P
  while Inventory is not empty or defined time step **do**
    Select a random action with the exploration rate ϵ
    Otherwise select d_u, $d_u = \arg\max_a Q(s, a)$ from F
    u determines to accept or reject, probability decided by Equation 3
    if Accepted **then**
      Get user satisfaction h and provider benefit c
    else
      $h, c = 0$
    end if
    Reward $r = \omega * h + (1 - \omega) * c$
    Update Q-value for the current state-action pair $Q(s, a) \leftarrow Q(s, a) + \alpha * (r + \gamma * \max_{a'} Q(s', a') - Q(s, a))$
    Transition to the next state (new recommendation) and update inventory
  end while
end for

Estimated prosociality scores (e) range from 0 to 1 and influence the probability of accepting recommended food not in their preference. According to these assumptions, in Table 2, User 1 is more prosocial than User 2.

Table 3: User profile parameters and computing user prosociality score

age e_1	18-: 0.5, 65+: 0.5, else: 0.9
family size e_2	1: 0.9, 2+: 0.5, 4+: 0.25
has allergies or restrictions e_3	yes: 0.5, no: 0.9
has child(ren) e_4	yes: 0.5, no: 0.9
estimated prosociality $e = e_1 * e_2 * e_3 * e_4$	

4.2 Food Inventory

Our simulation necessitates a comprehensive and realistic dataset that encompasses not just the items but also their attributes. We obtained a food list from USDA [16] (169 different items) and classified it into six categories that people request every day, which are meat, fruits and vegetables, dairy, eggs, cooking items (like oils and seasoning), and others. However, since the USDA [16] data lacks the specific attributes we need, we augmented attributes with feasible assumptions as close to demands mentioned in Caspi et al. [3]. For simplicity, we limit to considering quantity, expiration date, and perishability as key components of setting urgency of allocation. To make the dataset sufficient for the simulation environment, the quantity is set to big enough for each item. Then, we arbitrarily formulate the values of each item representing the criteria mentioned in Section 3.1.2. If the quantity becomes zero, the associated value becomes zero, and the item should not be recommended. If the expiration date has passed while the item is still available, all the remaining becomes wasted.

5 RESULTS

This section presents the experimental results of our proposed recommendation model. We conduct simulations with 1,000 agents, each corresponding to a user, to verify our model. Our study considers three baselines: random recommendation, user-focused, and provider-focused approaches.

Random recommendation Random recommendation represents a naive approach that randomly recommends items that are in stock, regardless of user preferences or provider benefit. This baseline disregards fairness and trust.

User-focused This baseline solely considers users' preferences based on their past interactions and preferences. This model is equivalent to assigning a weighting factor ω of 1 to prioritize actions aligned with user desires and completely ignore provider benefit.

Provider-focused This model prioritizes the provider benefit and disregards user preferences. It is equivalent to assigning a weighting factor ω of 0 to support the benefit of the provider-side operation exclusively.

To evaluate our model's performance, we consider three distinct values for the weighing factor (ω^*) in our setting: 0.2 determined by Equation 2, and 0.5 which evenly considers both sides. The results consistently show that our model with the optimal value of the weighting factor outperforms in satisfying both stakeholders' objectives. The model is trained with a learning rate (α) of 0.1, a discount factor (γ) of 0.9, an exploration rate (ϵ) of 0.1, and a prosociality weight (β) of 0.1.

5.1 Comparison on Provider Benefit and User Satisfaction

As shown in Figure 3a, the provider-focused model delivers the highest cumulative provider benefit, and the user-focused model achieves the lowest provider benefit. The provider benefit decreases as the weight assigned to the provider decreases, in other words, it increases inversely related to ω.

However, Figure 3b shows high cumulative user satisfaction, underscoring the superiority of weighted models. User satisfaction visibly improves, unlike what we originally expected both stakeholders to sacrifice to some extent if we set a parameter for the reward. The balanced ($\omega = 0.5$) model and the weighted models, particularly the optimal ($\omega = 0.2$ in this setting) value, outperform the user-focused model in terms of getting higher user satisfaction. It indicates that Meal recommends items that users like more, as shown in Figure 3b and Figure 4a.

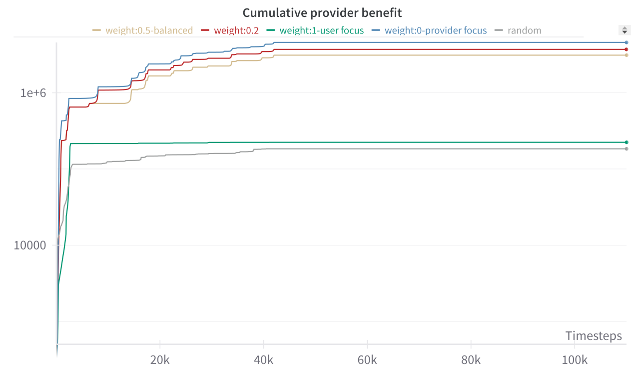
This suggests that users find greater satisfaction with recommendations that take into account both user preference and society's welfare. These models consistently demonstrate higher cumulative satisfaction than the user-focused model, as previously depicted in Figure 4a. This trade-off indicates that Meal fulfills the intended objectives even though it might sacrifice some provider benefits.

5.2 Acceptance Rate Tendency

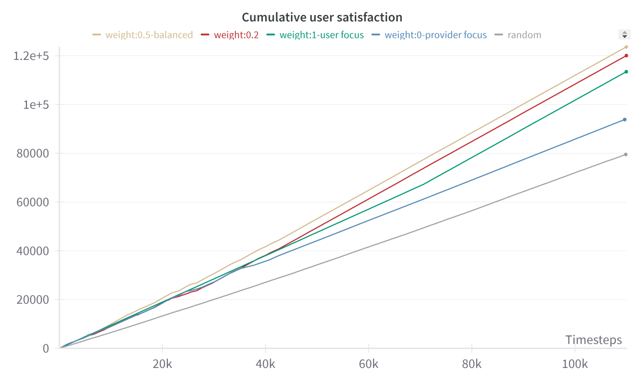
As shown in Figure 4b, the gap in the acceptance rate between the user-focused and provider-focused models differs notably. The user-focused model dominates all other models, especially the provider-focused and random recommendation. We could observe that the acceptance rate gradually drops in the provider-focused model unlike increasing in other models. These results imply that users find provider-focused recommendations unsatisfactory and are more likely to reject them. The acceptance rate is affected by how much the model skews to user satisfaction. The higher the weight on user preferences, the higher the acceptance rate. However, interestingly, while the user-focused model achieves the highest acceptance rate, other models with weighting converge around similar rates with minor variations, only less than 0.01%. This observation indicates that while the user-focused model has the strongest alignment with user preferences and needs, weighted models still achieve a fairly close acceptance rate.

5.3 Trade-Offs

In our exploration of the relationship between the two objectives, we evaluate various weightings to determine the optimal value of ω, as in Equation 2. Figure 4c summarizes the trade-offs between user satisfaction and provider benefit with different ω values under the same conditions, highlighting the important results. It shows that the cumulative satisfaction gain with weighted models surpasses that of the user-focused model, proving the effectiveness of our approach. It becomes evident that the balanced model appears to yield the highest user satisfaction. This observation might initially consider that the balanced model ($\omega = 0.5$) is optimal. However, with our multifaceted objectives, the model with $\omega = 0.2$ is closer to a preferable choice. It is substantiated by its ability to attain higher provider benefits with the lowest difference between user satisfaction and the provider benefit compared to the balanced model, thereby aligning more closely with our objective to maximize cumulative overall rewards (Figure 5). Simultaneously, it maintains higher user satisfaction than the user-focused model, which still upholds a substantial level of user satisfaction. It suggests that strategically balancing user preferences and provider benefits in consideration can unlock greater user satisfaction than solely prioritizing user needs. This discovers when users are willing to concede for the social good and shift their preferences, we can obtain a more equitable and sustainable food distribution system.



(a) Cumulative provider benefit



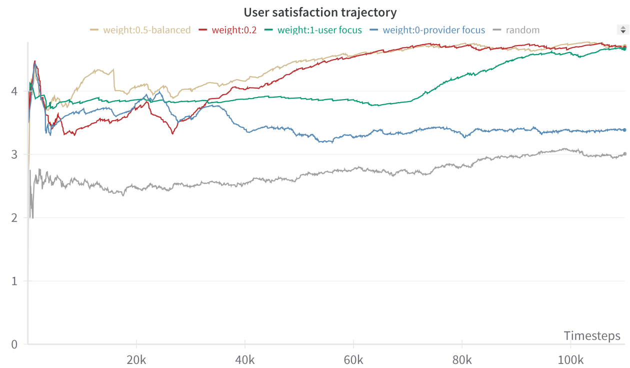
(b) Cumulative user satisfaction

Figure 3: Cumulative values for each stakeholder

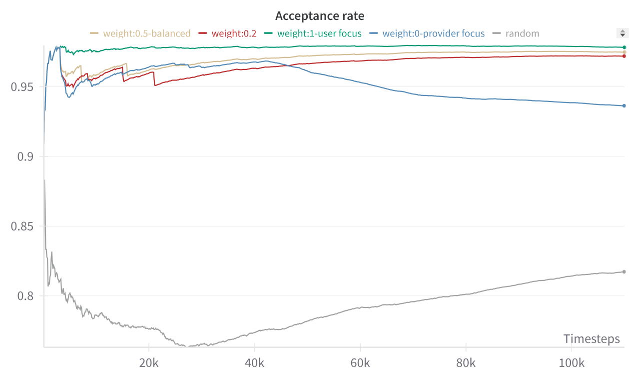
6 LIMITATIONS AND FUTURE WORK

Our proposed model faces several limitations. One of the main limitations of our proposed model lies in addressing nutritional factors and health considerations. The model recommends items solely relying on explicit preferences toward each food item given by users. However, a food suggestion can be made from various angles, not just based on preferred options; for example, considering nutritional information and substitutions accounting for specific dietary restrictions or medical conditions could enhance the overall well-being and promote the health of users within society. In the same context, more complicated settings arising from individual health statuses, socioeconomic backgrounds, culture, religious factors, regional disparities, and other diversity across communities remain challenging for scaling up the model's performance for optimization.

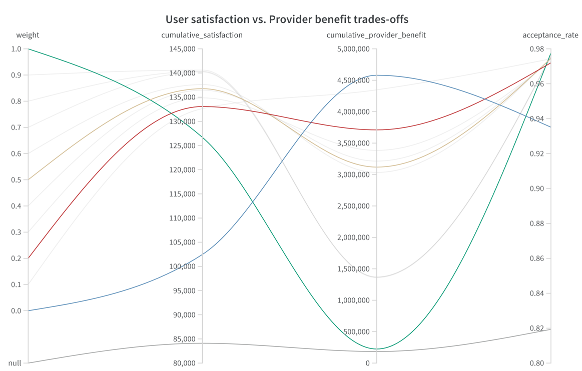
We only focus on two stakeholder types: users and food providers. Incorporating additional stakeholder types and their objectives



(a) User satisfaction trajectory of different ω



(b) Acceptance rates of different ω



(c) Trade-offs with varying ω. The learning rate is 0.1 for all experiments except the random recommendation model (gray line). The learning rate and weight are null in this case.

Figure 4: Simulation results in $\omega = 0.2$ and trade-offs

could provide a more holistic view of the food distribution problem. Potential involvement could encompass government agencies, volunteers, donors, and logistics. However, ensuring social well-being, fairness, and trust among these stakeholders while achieving their own goals poses a challenge.

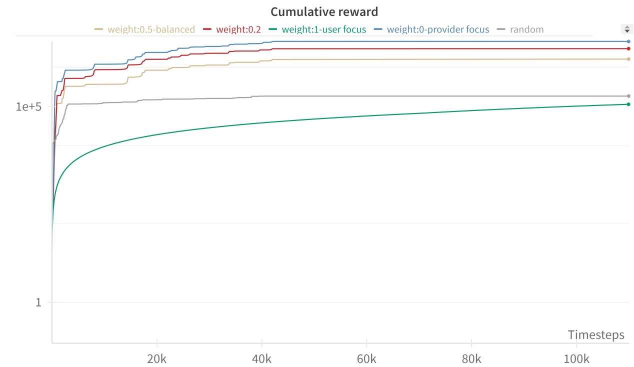


Figure 5: Cumulative combined reward

Additionally, our model omits the numerous internal and external factors that influence decision-making and cognition. We may explore further psychological theories and social science approaches to capture more accurate measurements of human cognitive architecture and design persuasive strategies to eventually nudge users to shift their preferences for the social good or to make healthier choices in more complex situations.

One last challenge we encounter pertains to obtaining real-world data. This limitation may affect the realism of our simulation, especially in assessing the provider's benefit.

7 CONCLUSION

Achieving equitable food distribution requires a multifaceted endeavor that meets various goals. Our approach involves introducing a value-aware recommendation system that accommodates two forefront stakeholders' needs and their interactions in allocating food resources. Meal seeks to optimize the allocation strategy toward maximizing the rewards for user satisfaction and provider benefit, employing Q-learning. Our findings highlight that the right balance for the reward function enhances user satisfaction while maximizing provider benefits. Our experiments simulate the society aligning with theoretical literature and other empirical findings in the relevant fields. Such alignment reinforces the robustness and applicability of our proposed method in real-world scenarios. Although the simulation is inconclusive on user preferences, it reveals relationships between variables that demonstrate the relevant trends. In future work, we will further investigate and incorporate more specific and realistic situations of food allocation systems by adding more complex and dynamic elements of stakeholders, as well as incorporating psychological architecture to mimic human minds.

ACKNOWLEDGMENTS

This work is supported by the NSF under grant 2125600.

REFERENCES

[1] Himan Abdollahpouri, Gediminas Adomavicius, Robin Burke, Ido Guy, Dietmar Jannach, Toshihiro Kamishima, Jan Krasnodebski, and Luiz Pizzato. 2020. Multi-stakeholder Recommendation: Survey and Research Directions. *User Modeling and User-Adapted Interaction* 30 (2020), 127–158.

[2] Mohammad Mehdi Afsar, Trafford Crump, and Behrouz H. Far. 2023. Reinforcement Learning Based Recommender Systems: A Survey. *Comput. Surveys* 55, 7 (2023), 1–38.

[3] Caitlin E. Caspi, Cynthia Davey, Christina Bliss Barsness, Nora Gordon, Laura Bohen, Marna Canterbury, Hikaru Peterson, and Rebekah Pratt. 2021. Peer Reviewed: Needs and Preferences Among Food Pantry Clients. *Preventing Chronic Disease* 18 (2021).

[4] Harshal A. Chaudhari, Sangdi Lin, and Ondrej Linda. 2020. A General Framework for Fairness in Multistakeholder Recommendations. *arXiv preprint* (2020). arXiv:2009.02423

[5] Charles Evans and Atoosa Kasirzadeh. 2023. User Tampering in Reinforcement Learning Recommender Systems. In *Proceedings of the 2023 AAAI/ACM Conference on AI, Ethics, and Society (AIES)*. ACM, Montréal, QC, Canada, 58–69. https://doi.org/10.1145/3600211.3604669

[6] Ernst Fehr and Klaus M. Schmidt. 1999. A Theory of Fairness, Competition, and Cooperation. *The Quarterly Journal of Economics* 114, 3 (1999), 817–868.

[7] Nada Ghanem, Stephan Leitner, and Dietmar Jannach. 2022. Balancing Consumer and Business Value of Recommender Systems: A Simulation-Based Analysis. *Electronic Commerce Research and Applications* 55 (2022), 101195.

[8] Tomoshi Iiyama, Daisuke Kitakoshi, and Masato Suzuki. 2021. An Approach for Creation of Logistics Management System for Food Banks Based on Reinforcement Learning. In *The 4th International Conference on Information Science and Systems (ICISS)*. ACM, Edinburgh, United Kingdom, 60–67.

[9] Tomoshi Iiyama, Daisuke Kitakoshi, and Masato Suzuki. 2022. Optimizing Food Allocation in Food Banks with Multi-agent Deep Reinforcement Learning. In *Proceedings of the International Conference on Technologies and Applications of Artificial Intelligence (TAAI)*. IEEE, Tainan, Taiwan, 203–208.

[10] Youngbin Kwak and Scott A. Huettel. 2016. Prosocial Reward Learning in Children and Adolescents. *Frontiers in Psychology* 7 (2016), 1539.

[11] Edward C. Malthouse, Khadija Ali Vakeel, Yasaman Kamyab Hessary, Robin Burke, and Morana Fuduric. 2019. A Multistakeholder Recommender Systems Algorithm for Allocating Sponsored Recommendations. In *Proceedings of the Workshop on Recommendation in Multi-stakeholder Environments*, Vol. 2440. CEUR-WS.org, Copenhagen, Denmark.

[12] Changhua Pei, Xinru Yang, Qing Cui, Xiao Lin, Fei Sun, Peng Jiang, Wenwu Ou, and Yongfeng Zhang. 2019. Value-Aware Recommendation Based on Reinforced Profit Maximization in E-commerce Systems. *arXiv preprint* (2019). arXiv:1902.00851

[13] Guy Shani, David Heckerman, Ronen I. Brafman, and Craig Boutilier. 2005. An MDP-Based Recommender System. *Journal of Machine Learning Research* 6, 9 (2005), 1265–1295.

[14] Özge Sürer, Robin Burke, and Edward C. Malthouse. 2018. Multistakeholder Recommendation with Provider Constraints. In *Proceedings of the 12th ACM Conference on Recommender Systems*. ACM, Vancouver, BC, Canada, 54–62.

[15] USDA. 2023. *Food Security and Nutrition Assistance*. USDA. https://www.ers.usda.gov/data-products/ag-and-food-statistics-charting-the-essentials/food-security-and-nutrition-assistance/ [Accessed 11-12-2023].

[16] USDA. 2023. What We Eat In America (WWEIA) Database | Ag Data Commons 2023. https://data.nal.usda.gov/dataset/what-we-eat-america-wweia-database [Accessed 12-13-2023].

[17] Christopher J.C.H. Watkins and Peter Dayan. 1992. Q-Learning. *Machine Learning* 8 (1992), 279–292.

[18] Longqi Yang, Cheng-Kang Hsieh, Hongjian Yang, John P. Pollak, Nicola Dell, Serge Belongie, Curtis Cole, and Deborah Estrin. 2017. Yum-Me: A Personalized Nutrient-Based Meal Recommender System. *Transactions on Information Systems (TOIS)* 36, 1 (2017), 1–31.

[19] Yong Zheng, Nastaran Ghane, and Milad Sabouri. 2019. Personalized Educational Learning with Multi-Stakeholder Optimizations. In *Adjunct Publication of the 27th Conference on User Modeling, Adaptation and Personalization (UMAP)*. ACM, Larnaca, Cyprus, 283–289.