PATHS: Agent-Based Modeling of Homelessness Pathways

Nowshin Tasnim, Khandker Sadia Rahman, and Charalampos Chelmis

Department of Computer Science,
University at Albany, NY 12222, USA
ntasnim@albany.edu, srahman2@albany.edu, cchelmis@albany.edu

Abstract. Homelessness is a complex and persistent societal challenge that requires innovative thinking in order to design solutions that not only provide immediate relief but also address the problem holistically. In this work, we introduce a probabilistic agent-based model of the homeless service system (PATHS) in our effort to develop a virtual laboratory that will eventually serve as a testbed for policy interventions and a tool for the ethical study of homelessness. PATHS is designed to simulate the movement of individuals through the unobserved network of homeless services while at the same time capturing system-level population dynamics by incorporating practical constraints such as capacity limits, time-dependent entry rates, and transition bottlenecks. Our experimental evaluation shows that PATHS closely matches the characteristics of service navigation in a real-world dataset, despite its simplicity, and outperforms baseline models, often by a considerable margin.

Keywords: Agent-based model, complex system, socio-technical system

1 Introduction

Homelessness has been a persistent, multifaceted issue that spans housing insecurity, economic hardship, social inequality, mental health challenges, and fundamental human rights, with profound consequences for both individuals and society at large [1–6]. In 2024, more than 770, 000 individuals in the United States were without stable housing on a single night, marking the highest number ever recorded [7]. This surge in homelessness has placed unprecedented strain on local service systems responsible for providing housing support and related services [8]. Effectively addressing the needs of individuals experiencing homelessness requires strategic allocation of limited resources—an increasingly difficult task amid rising demand and constrained capacity [9]. Moreover, the effectiveness of existing decision-making processes in allocating these limited resources remains unclear, as allocation strategies have been relatively understudied [10, 11]. The situation is further complicated by regulatory restrictions, inadequate infrastructure, and technological limitations, which often hinder the timely and effective implementation of solutions [12].

Technologically, homeless service providers rely on accurate and timely data to make informed decisions. Enhancing the ability to predict future trends based on existing patterns would further enable more efficient and effective resource allocation, ultimately improving service delivery and outcomes [13]. People experiencing homelessness often move through different services over time—like shelters, job programs, or health support. We can think of this system as a network, where each service is a node, and the connections between them show how people move from one to another. Using network science helps us understand how these services are connected, how likely certain transitions are, and where there might be gaps or bottlenecks in the system [14].

Within this context, simulation provides a powerful tool for exploring system behavior over time by constructing computational representations of real-world dynamics [15]. Among simulation techniques, Agent-Based Modeling (ABM) is particularly well-suited for capturing the complexities of such systems by simulating the behavior of heterogeneous, interacting agents to obtain emergent system-level dynamics [16–18]. ABMs are especially effective for systems characterized by feedback loops (such as homelessness), spatial or temporal locality, and agent heterogeneity [19].

In this work, we propose a probabilistic, discrete-time ABM over a directed network of homeless services, in which nodes represent service types (projects) and weighted edges model the empirically derived transition probabilities between them. Specifically, our probabilistic ABM replicates individual movement through the system using real-world administrative data, capturing both micro-level trajectories (i.e., sequences of service engagements) and macro-level dynamics (i.e., service occupancy over time). The model incorporates practical constraints such as service capacity limits, time-dependent entry rates, and transition bottlenecks to more accurately reflect real-world conditions. Thus, the proposed ABM offers a data-driven, yet principled virtual lab for simulating individual-level transitions within the homelessness system, which in turn can facilitate the ethical study of homelessness by testing, for instance, the (un)anticipated outcomes of interventions such as service prioritization and AI-powered referral systems.

2 Related Work

2.1 Homelessness

Homelessness has been recognized as a multifaceted social phenomenon shaped by demographic disparities and influenced by structural inequalities, systemic discrimination, and health-related vulnerabilities [1–4, 6, 12, 20–22]. Unlike prior studies that primarily emphasize demographic disparities and structural inequalities, this work focuses on replicating the individual trajectories of homeless individuals across the unobserved network of homeless services. To the best of our knowledge, the proposed model is the first to capture individual pathways at this level of granularity, offering for the first time a high-resolution view of how personal circumstances, behavioral decision-making, and service network structures interact over time.

2.2 Predictive Modeling

Several predictive models have been developed to address different challenges within the homelessness domain. One line of research focuses on binary classification tasks—such as predicting whether an individual is at risk of reentering the homelessness service system (HSS) [23–25], or whether they are likely to experience chronic homelessness [26, 27]. Another line of work explores multiclass service assignment, aiming to predict the next service that should be provided to an individual [28–30]. For instance, [29] infers a network of services based on observed transitions and uses it to predict the most likely next service. [28] takes a probabilistic approach, designing a Bayesian network to model service dependencies. In contrast, [30] employs a representation learning framework to directly learn service pathways for next-step prediction. Finally, [24] develops a service assignment model that not only predicts next services but also aims to minimize the risk of reentry using a counterfactual prediction approach. The latter can serve as valuable tools for targeted interventions but typically generate static, outcome-based predictions and often function as black-box models with limited interpretability. On the other hand, network science-based studies [31,32] offer structural insights into service transitions. However, such studies neither provide methods to simulate individual-level behaviors over time nor model dynamic interactions within the HSS. In contrast, the proposed agent-based model explicitly simulates the temporal trajectories of individuals while allowing for the modeling of resource constraints, decision processes, and system evolution. These properties make ABM better suited for exploring individual pathways, testing intervention strategies, and informing policy-level decisions in this complex and resource-constrained system.

2.3 Agent-based Modeling

ABM has been widely used to simulate decentralized systems with heterogeneous agents and dynamic interactions, enabling researchers to study how macro-level phenomena emerge from localized behaviors [17–19,33,34]. Prior work is too vast to enumerate; representative studies include models of online communities [35], artificial societies [36,37], and logistical networks, such as crowd-shipping for food rescue [38] and freight transportation systems [39]. While related, the domains studied in the literature differ both in objectives and constraints from the homelessness system, where transitions are decided based on factors, including but not limited to, capacity limitations and individual socioeconomic characteristics, and the network of services is completely unobserved. ABMs for homelessness have explored spatial victimization risks [40], individual capability restoration through policy interventions [41], and macro-scale encampment evolution [42]. To the best of our knowledge, none of these works model homeless individuals and are therefore incapable of simulating individual-level pathways over time.

This is in sharp contrast with the proposed ABM, which focuses specifically on reproducing individual trajectories through the unobserved network of homeless services under real-world constraints.

3 Background

Homeless service providers offer a range of services, such as emergency shelters and permanent housing, that are organized as distinct project types (as categorized in [43]). These project types are encoded as numeric identifiers ranging from 1 to 14. Individuals experiencing homelessness often cycle in and out of the HSS. Administrative records, maintained by homeless service providers in accordance with U.S. federal regulations, record the assignment of each individual to services, which in turn can be used to establish detailed timelines of each person's service history, including start and end dates of each service, transitions between service types, and patterns of exiting and reentering the system. Thinking of HSS as an unobserved network of interconnected services that individuals traverse over time [31] enables the study of individual movements through this complex sociotechnical network over time, with the broader aim of promoting long-term housing stability. Specifically, we think of HSS as a network of services connected by directional edges that capture transitions between them and additionally incorporate two special nodes that denote system *entry* (an individual's first recorded interaction with the HSS) and *exit* (the final recorded departure of an individual), respectively. The advantage of including these special nodes is the ability to record individuals' full trajectory of involvement in the system–from initial entry into any service to eventual exit–as opposed to merely recording transitions between services. Here, we use the inferred HSS network [31] to drive the simulation of individual trajectories.

4 Proposed Method

We propose a Probabilistic discrete-time Agent-based Model of the Homeless Services system (PATHS), grounded in empirical data and designed to replicate the individual trajectories of homeless individuals across a network of services over time. Unlike prior models that operate at the service level or on aggregate, PATHS explicitly models individual trajectories, constrained by empirical timelines, capacity limitations, and service-specific transition patterns. Specifically, each individual within the homeless system is modeled as an autonomous agent that, at any given time, is either assigned to a service or remains unassigned. The latter case captures an individual's status when first entering the system (i.e., *entry* state) or after no longer being active in the system (i.e., *exit* state).

Mathematically, let the set of agents navigating the homeless service network be denoted as $A = \{a_1, a_2, \ldots, a_{|A|}\}$. The state of each agent $a_i \in A$ at time step $t \in [0, N]$ is represented as $s_i^t \in \mathcal{S}$, where $\mathcal{S}$ is the set of all possible nodes in the network: $\mathcal{S} = \{\mathit{entry}, \mathit{exit}, P_1, P_2, \ldots, P_K\}$, P_j denotes the j-th service (project type), and K is the total number of services. The *entry* node marks the agent's

Table 1: Parameters of the proposed model, PATHS

Parameter	Value	Description
A	$\{a_1, a_2, \ldots, a_{\lvert A\rvert}\}$	Set of agents
$\mathcal{S}$	$\{entry, exit, P_1, P_2, \ldots, P_K\}$	Set of states (i.e., services, *entry*, and *exit*)
C	$C \in \mathbb{N}^{\lvert\mathcal{S}\rvert\times 1}$	Service-specific capacity constraint
t	$[0, N]$	Time step ($N = 72$ in our experiments)
T	$T \in \mathbb{R}^{\lvert\mathcal{S}\rvert\times\lvert\mathcal{S}\rvert}$	Transition probability matrix
τ	$\tau \in \mathbb{N}^{N\times 1}$	Temporal Constraint
E	$E \in \mathbb{N}^{N\times 1}$	Entry Constraint

first recorded interaction with the system (i.e., when an individual is assigned to a service for the first time), while the *exit* node indicates their final recorded interaction with the HSS.

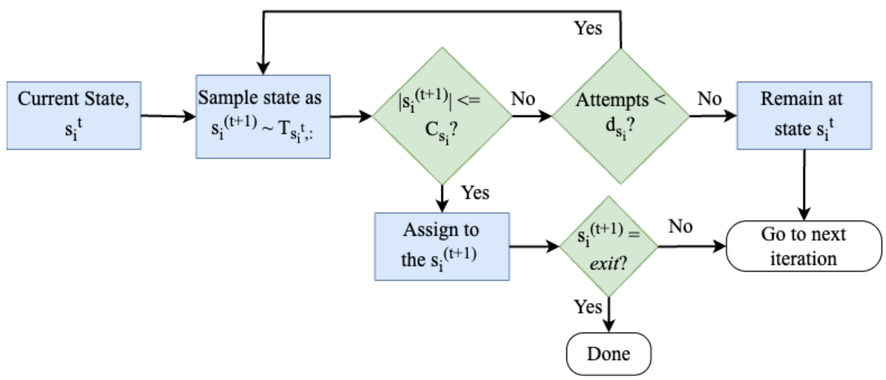

Fig. 1: Flowchart of agent transitions between states in each time step.

The simulation begins by determining whether an individual is entering the system for the first time (i.e., there exists no prior record of this individual, and thus a corresponding agent). If so, a new agent a_i is instantiated and initialized with state, $s_i^t = entry$. From s_i^t, PATHS selects a service for a_i based on the transition matrix $T_{s_i^t,:}$ and subject to capacity constraints. Specifically, if the selected service is at capacity, an alternative is chosen from the remaining reachable services, up to $(d_{s_i^t} - 1)$ additional attempts, where $d_{s_i^t}$ is the number of reachable services from s_i^t, as shown in Fig. 1. At each time step t, an agent is eligible to transition to a new service if: (1) there is at least one valid transition from the current state (i.e., $T_{s_i^t,:} > 0$), and (2) the current state is not an *exit* state, i.e., $s_i^t \neq exit$. If the next service is at capacity, the model attempts to find an alternative (up to $d_{s_i^t} - 1$ times) by selecting in turn each of the reachable services in descending order of their transition probability (i.e., $T_{s_i^t,:}$) from the current service. When an agent transitions to the *exit* state (which indicates that the individual is no longer active in the system), simulation for that agent ends.

4.1 Model Inputs

Table 1 summarizes the parameters that are provided as inputs to PATHS. Next, we describe the methodology used to compute them. Fig. 2a shows sample administrative data, which records the services received by each individual along with their corresponding entry and exit dates. To better reflect real-world limitations, we introduce a set of constraints. Since these constraints are not directly available in the data, we derive them empirically as follows.

(a)

ID	Service	EntryDate	ExitDate
12	11	2012-01-04	2012-02-03
12	13	2012-02-03	2012-04-18
12	11	2012-04-18	2012-05-20
13	3	2012-01-10	2012-02-12
13	6	2012-02-12	2012-03-15
14	6	2012-01-15	2012-04-27
15	11	2012-01-22	2012-03-13
16	3	2012-02-05	2012-02-07
14	1	2017-12-09	2017-12-25
15	1	2017-12-01	-----------
16	1	2017-12- 15	-----------
14	1	2017-12-29	-----------

(b)

ClientID	Trajectory
12	*entry*, 11, 13, 11, *exit*
13	*entry*, 3, 6, *exit*
14	*entry*, 6, 1, 1
15	*entry*, 11,1
16	*entry*, 3, 1

(d)

Service	Capacity
1	4
3	2
6	2
11	3
13	1

(c)

Timestep	Entry Constraint	Temporal Constraints
1	4	0
2	1	3
3	0	2
4	0	2
5	0	1
72	3	1

Fig. 2: Overview of the model inputs construction used in the proposed agent-based simulation: (**a**) sample administrative records of individuals, (**b**) service trajectories for each individual including *entry* and *exit*, (**c**) entry and temporal constraints (E and τ) at each time step t, and (**d**) service-specific capacity constraints C.

First, we model the homeless service system as a directed network of services denoted by $\mathcal{G} = (\mathcal{S}, \mathcal{E})$, from which we derive a transition probability matrix T, where each entry $T_{i,j}$ represents the empirical probability of transitioning from state s_i to state s_j based on observed service trajectories. Here, $\mathcal{S}$ represents the set of all states, including services as well as the special *entry* and *exit* states. The set of directed edges $\mathcal{E}$ captures the transitions between states, with edge weights encoding transition probabilities that are computed following the methodology in [31]. Fig. 3 illustrates the resulting transition graph constructed from the sample data shown in Fig. 2a. To construct this graph, we first extract the trajectory of each individual from the administrative data, as illustrated in Fig. 2b. For example, the trajectory for the individual 12 is: *entry* $\rightarrow$ 11 $\rightarrow$ 13 $\rightarrow$ 11 $\rightarrow$ *exit*.

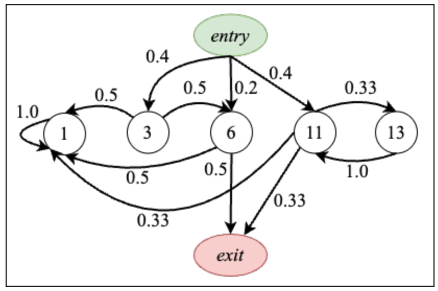

Fig. 3: Transition graph of states for the sample data in Fig. 2b

We then count the frequency of each unique transition across all individual trajectories—for instance, the transition *entry* $\rightarrow$ 11 appears twice. Finally, to compute transition probabilities, we normalize the outgoing edges from each node so that their weights sum to 1. For example, there are three outgoing transitions from node 11: 11 $\rightarrow$ 13, 11 $\rightarrow$ 1, and 11 $\rightarrow$ *exit*, each appearing once in the data. After normalization to ensure that transition probabilities sum up to 1, each of these transitions is assigned a probability of 1/3.

Next, we define a service-specific capacity constraint, which represents the maximum number of individuals that service P_i can accommodate at any given time. This constraint ensures that the model respects realistic capacity limits when assigning agents to services. We acknowledge that capacity constraints may fluctuate over time, however, we consider the simple case where such constraints are static for modeling simplicity. To estimate the capacity C_{P_i} of each P_i, we compute the number of individuals assigned to P_i at each timestep and take the maximum across all timesteps. Formally, $C_{P_i} = \max_t C^t_{P_i}$, where $C^t_{P_i}$ denotes the number of individuals assigned to P_i at timestep t. For example, in Fig. 2a, service 1 has four individuals assigned at timestep $t = 72$. Since this is the maximum observed across all timesteps, $C_1 = 4$, as shown in Fig. 2d. This process is repeated for every service.

We additionally define the entry constraint (E) to record the number of new agents entering the system at each timestep. This constraint ensures that agents are not all created at the initial timestep, but instead reflect the actual timing of individuals entering the homeless service system. By doing so, the model more accurately mirrors real-world service entry patterns over time. Formally, the entry constraint at timestep t is defined as $E^t = \sum_{i=1}^{|A|} \mathbf{1}_{a_i^t \notin A^*}$, where A^* denotes the set of agents already present in the system prior to timestep t, and $\mathbf{1}_{a_i^t \notin A^*}$ is an indicator function that evaluates to 1 if agent a_i is entering the system for the first time. Thus, E^t represents the number of new agents initialized at timestep t. For example, in Fig. 2a, at timestep $t = 1$, individuals 12, 13, 14, and 15 enter the system for the first time, resulting in the creation of four new agents. This is represented as $E^1 = 4$.

Finally, we introduce the temporal constraint (τ), which limits the total number of state transitions allowed across all agents at each timestep. This constraint ensures that transitions are distributed over time, rather than allowing the model to progress through trajectories rapidly. Without this constraint, the simulation may prioritize reaching the end of trajectories prematurely, thereby underrepresenting longer service trajectories. For example, in Fig. 2a, at timestep $t = 2$, each of the individuals 12, 13, and 16 transitions from one state to another. Specifically, individual 12 transitions from $11 \rightarrow 13$, individual 13 from $3 \rightarrow 6$, and individual 16 from $3 \rightarrow 1$. This results in a temporal constraint value of $\tau^2 = 3$, indicating that three transitions occur during that timestep.

4.2 Validation Approach

Model validation is a critical step in ensuring the credibility, utility, and generalizability of any simulation framework. In the context of homelessness, where interventions can have a significant real-world impact on humans as well as policy implications, it is essential to ensure that the model not only replicates observed data patterns in empirical data but also preserves meaningful behavioral and systemic dynamics. A validated model builds trust with stakeholders, supports decision-making, and increases the potential for actionable insights.

To evaluate the validity of PATHS, we compare simulated outputs against key trends observed in a real-world dataset. Specifically, we assess PATHS' ability to replicate system-level temporal dynamics by examining population distributions across services over time. We additionally evaluate PATHS at the individual level by evaluating two behavioral indicators: (i) trajectory length (i.e., the number of service episodes per person) and (ii) duration of stay (i.e., the total number of months spent in the system). Evaluating PATHS both at the system-wide level and at the level of individuals allows us to assess the model's representational and predictive capacity by evaluating whether the emergent properties of simulations obtained using PATHS align with the empirical observations.

5 Experimental setup

5.1 Data description

We evaluate our model using a real-world dataset provided by CARES of NY Inc. The dataset comprises $18,567$ administrative records of service episodes associated with $5,993$ individuals, spanning the period from 2012 and 2017. Each record corresponds to a unique service episode and includes both demographic and service-related information, such as entry and exit dates, project types, and a range of socioeconomic and health-related attributes. We specifically use each record's unique ID, service type (or project type), and the corresponding entry and exit dates to construct individual-level trajectories through the homelessness system. The dataset encompasses a total of 9 distinct services representing the range of services offered across the homeless support network. To capture system entry and exit points explicitly, we augment the dataset with two additional states: state 0 represents *entry* to HSS, whereas -3 denotes *exit* (i.e., an individual has left HSS without reentry).

As part of our evaluation, we use individuals' trajectory length (i.e., the number of recorded service episodes per individual) and duration of stay (i.e., number of months one used a given service). We therefore report descriptive statistics for these two quantities in our dataset. Trajectory length has a median of 4 (mean of 5), indicating that while most individuals engage briefly with the system, a smaller subset has substantially longer service histories. Similarly, length of stay has a median of 11 (mean of 16), reflecting a right-skewed distribution in which most individuals receive support for a relatively short duration, whereas a few remain in the system for extended periods.

5.2 Metrics

We evaluate PATHS with respect to **Root mean squared error (RMSE)**, which measures the average of squared differences between predicted and actual values, and **Mean Absolute Error (MAE)**, which measures the average magnitude of prediction errors. Specifically, we use RMSE to penalize large errors and MAE as an interpretable metric that is less sensitive to large errors.

5.3 Baselines

We compare PATHS against four variants of random transition strategies (i.e., a service is selected uniformly at random from the set of all possible states $\mathcal{S}$, rather than relying on the transition probabilities in matrix T) in which we impose the following entry constraints: (**Random**) the entry constraint remains the same as PATHS; (**RM_Dbl**) doubling the entry constraint per timestep; (**RM_MaxE**) replacing the entry constraint with the maximum new entry observed across all timesteps; and (**RM_MedE**) using the median value for the entry constraint for each timestep.

5.4 Implementation Setup

PATHS is implemented in Python using the AgentPy library. The simulation runs in monthly timesteps (t) over $N = 72$ months. Due to the model's probabilistic nature, each configuration is run ten times, and the results are averaged. To assess the robustness of the model under varying levels of data availability, we evaluate it under two scenarios: (1) using the complete service trajectories from 2012 to 2017 and (2) using only the last four recorded service episodes per individual, in line with [28] that recommended this strategy to improve predictive accuracy. All experiments were conducted on macOS with Apple M2 chip and 16 GB RAM. Our source code is publicly available at https://github.com/IDIASLab/PATHS.

6 Result Analysis

To provide a comprehensive assessment of PATHS, we evaluate its performance using both quantitative metrics and qualitative analysis.

Table 2: Comparative evaluation of PATHS and the baselines. The best-performing model is highlighted in bold.

Model	Random	RM_Dbl	RM_MaxE	RM_MedE	PATHS	PATHS ($n = 4$)
MAE	27.49± 6.41	26.19± 5.91	122.1± 17.06	29.12± 6.46	**11.84± 2.51**	12.53± 2.10
RMSE	28.0± 6.4	26.79± 5.9	123.41± 17.04	29.48± 6.45	**12.42± 2.51**	13.1± 2.11

6.1 Quantitative Analysis

Table 2 shows that PATHS consistently achieves the lowest error across all evaluation metrics and outperforms the baselines, even when limited to only the last four service episodes per individual (i.e., PATHS ($n = 4$)). The high errors (up to three times higher than PATHS) reported for all baselines suggest that, despite its simplicity, PATHS is a reasonable model of HSS.

6.2 Qualitative Analysis

Fig. 4 shows that PATHS effectively captures key structural patterns and temporal fluctuations observed in the empirical data. It replicates periodic surges in P1 and the gradual buildup in Exit and maintains stable low occupancy in smaller services (e.g., P3, P4, P6, and P14) as well as moderate levels in P11 and P12. These results suggest the model retains major seasonal and structural dynamics across both full and reduced input settings.

When operating with limited historical data (i.e., when using only the last four occurrences per individual), as shown in Figs. 4c and 4d, a modest decline in fidelity is observed. While core trends remain consistent, the model underestimates late-stage surges in more dynamic services and shows diminished sensitivity to subtle seasonal variations. Moreover, as seen in Fig. 4b, the simulated trajectories tend to smooth out high-frequency fluctuations present in the empirical data (Fig. 4a), especially for services like P1 and P11—an effect likely due to averaging across multiple simulation runs. Overall, the proposed model demonstrates strong generalizability and robustness, accurately reflecting population dynamics under both rich and sparse data conditions. These findings underscore its utility as a decision-support tool for resource planning and policy development in HSS.

In Fig. 5a, using the full historical dataset, the proposed model closely mirrors the ground truth across a wide range of trajectory lengths. Both simulated and ground truth distributions exhibit a heavy-tailed pattern, with most individuals having short trajectories and a smaller proportion experiencing extended stays. The alignment is particularly strong in the mid-range lengths (5–20 transitions), with slight deviations at the extremes due to the sparsity of long-stay individuals. These results demonstrate the model's ability to replicate real-world trajectory distributions when sufficient data is available. On the other hand, in Fig. 5b,

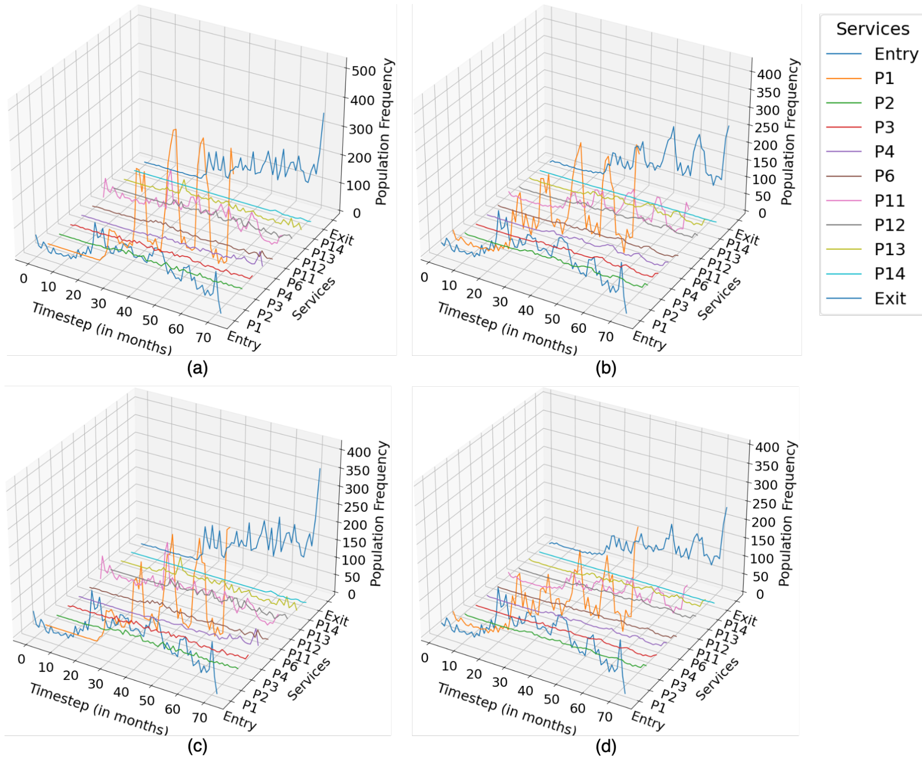

Fig. 4: Comparison of temporal population dynamics across services (project types), including *entry* and *exit* over N timesteps. Empirical data (left) and PATHS output (right) for the full dataset ((a) and (b)), and using only the last four occurrences per individual ((c) and (d)), accordingly.

while the overall trend remains consistent, the distribution becomes truncated due to limited historical depth. Shorter trajectories dominate; the model's ability to represent rare, longer stays is diminished. Nonetheless, the simulation still reflects the empirical pattern reasonably well for trajectory lengths between 2 and 6, indicating that the model retains robustness under constrained input conditions. Fig. 6a shows that the proposed model successfully captures the exponential-like decay observed in the ground truth using the full historical dataset, with strong alignment for shorter stays. However, it slightly underestimates the proportion of individuals with stays exceeding 40 months, likely due to smoothing of high-variance, long-duration patterns. When using only the last four occurrences (Fig. 6b), the model still captures the overall decay structure, though with slightly increased variance. Agreement remains strong up to approximately 40 months, but the limited historical input leads to under-representation in the tail, where long stays are more difficult to recover. Overall, these findings highlight the model's resilience in reproducing population-level behavioral dis-

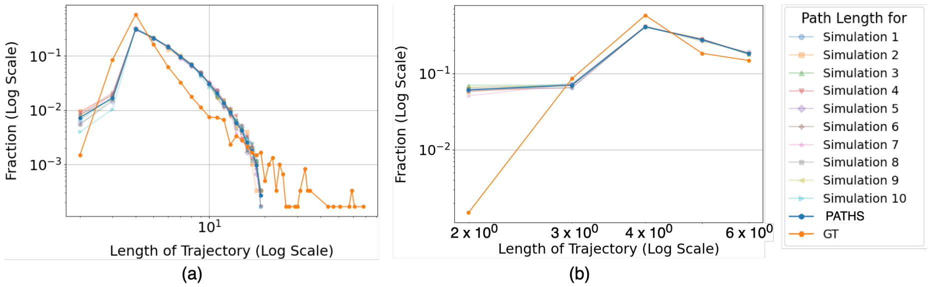

Fig. 5: Log-scaled distribution of path length for each of 10 simulations of PATHS, their average (PATHS), and ground truth (GT) using (a) the full dataset and (b) only the last 4 occurrences per individual.

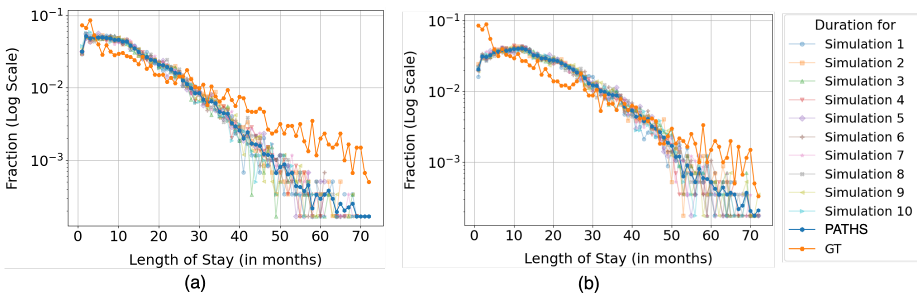

Fig. 6: Log-scaled distribution of clients' length of stay (LOS) in the homelessness system, measured in months. The figure compares the ground truth (GT), the result of each of 10 simulations of PATHS, and their average (PATHS), using (a) the full dataset and (b) only the last four occurrences per individual.

tributions, even with reduced historical data, while also emphasizing the added fidelity gained from richer longitudinal records, particularly in capturing long-tail behaviors.

7 Conclusion

In this study, we introduced PATHS, a probabilistic agent-based model designed to simulate individual transitions within the homeless services system using real-world administrative data. The model effectively reconstructs early-stage population dynamics and individual service trajectories, offering a high-resolution view of how the homeless interact with a complex network of support services. Our quantitative evaluation showed that PATHS outperforms baselines, demonstrating its robustness and practical relevance. Our qualitative analysis of individual agent trajectories revealed that PATHS generates realistic service sequences and captures key behavioral patterns observed in empirical data, providing supporting evidence in favor of its interpretability.

Limitations. PATHS maintains strong performance in early and mid-term horizons, effectively capturing core system dynamics and individual trajectories. However, long-term forecasting remains challenging due to the sparsity of long-term data and the limited incorporation of individual-level features. Furthermore, the current implementation is purely statistical and does not account for contextual or causal factors that may influence service assignments. Finally, the use of a single dataset cannot guarantee PATHS' generalizability.

Future Directions. By enabling the simulation of homelessness service assignments at the individual level, PATHS provides a foundation for ethical policy evaluation and system design. We aim to enhance PATHS' fidelity by learning more expressive representations of the service network, integrating feature-driven counterfactual reasoning, and combining probabilistic modeling with rule-based mechanisms. These improvements will enable PATHS to serve as a robust decision-support tool that will be capable of identifying system bottlenecks, evaluating policy interventions, and informing the design of more responsive, data-driven strategies to address homelessness. Importantly, this line of work serves as a stepping stone toward developing fair and unbiased models for service assignment, enabling equitable allocation of limited resources across diverse demographic and needs-based subgroups within the homeless population. In the future, we will also run simulations using diverse datasets to further validate PATHS' generalizability and adaptability across varying service environments.

References

1. Mago, V.K., Morden, H.K., Fritz, C., Wu, T., Namazi, S., Geranmayeh, P., Chattopadhyay, R., Dabbaghian, V.: Analyzing the impact of social factors on homelessness: a fuzzy cognitive map approach. BMC medical informatics and decision making 13, 1–19 (2013)
2. Padgett, D.K.: Homelessness, housing instability and mental health: making the connections. BJPsych bulletin 44(5), 197–201 (2020)
3. Fowler, P.J., Hovmand, P.S., Marcal, K.E., Das, S.: Solving homelessness from a complex systems perspective: insights for prevention responses. Annual review of public health 40(1), 465–486 (2019)
4. Foscarinis, M.: Homelessness and human rights: Towards and integrated strategy. . Louis U. Pub. L. Rev. 19, 327 (2000)
5. Lynch, P., Cole, J.: Homelessness and huamn rights: Regarding and responding to homelessness as a human rights violation. Melb. J. Int'l L. 4, 139 (2003)
6. Bose, A., Roy, S., Majumder, S., Chowdhury, I.R.: Lost in the cityscape: Exploring urban homelessness, its societal imprints and policy suggestions. Societal Impacts 1(1-2), 100026 (2023)
7. de Sousa, T., Henry, M.: The 2024 annual homelessness assessment report (ahar) to congress. part 1 – pit estimates of homelessness in the u.s. (2024), https://www.huduser.gov/portal/sites/default/files/pdf/2024-AHAR-Part-1.pdf, available online
8. Meyer, B.D., Wyse, A., Williams, D.: Asylum seekers and the rise in homelessness. Tech. rep., National Bureau of Economic Research (2025)

9. Fowler, P.J., Wright, K., Marcal, K.E., Ballard, E., Hovmand, P.S.: Capability traps impeding homeless services: A community-based system dynamics evaluation. Journal of social service research 45(3), 348–359 (2019)
10. Shinn, M., Greer, A.L., Bainbridge, J., Kwon, J., Zuiderveen, S.: Efficient targeting of homelessness prevention services for families. American journal of public health 103(S2), S324–S330 (2013)
11. Kube, A.R., Das, S., Fowler, P.J.: Fair and efficient allocation of scarce resources based on predicted outcomes: implications for homeless service delivery. Journal of Artificial Intelligence Research 76, 1219–1245 (2023)
12. Culhane, D.: The potential of linked administrative data for advancing homelessness research and policy. European Journal of Homelessness 10(3) (2016)
13. Singham, D.I., Lucky, J., Reinauer, S.: Discrete-event simulation modeling for housing of homeless populations. Plos one 18(4), e0284336 (2023)
14. Borgatti, S.P., Mehra, A., Brass, D.J., Labianca, G.: Network analysis in the social sciences. science 323(5916), 892–895 (2009)
15. Zeigler, B.P., Praehofer, H., Kim, T.G.: Theory of modeling and simulation. Academic press (2000)
16. Davidsson, P.: Multi agent based simulation: beyond social simulation. In: International workshop on multi-agent systems and agent-based simulation. pp. 97–107. Springer (2000)
17. Helbing, D.: Agent-based modeling. In: Social self-organization: Agent-based simulations and experiments to study emergent social behavior, pp. 25–70. Springer (2012)
18. Balke, T., Gilbert, N.: How do agents make decisions? a survey. Journal of Artificial Societies and Social Simulation 17(4), 13 (2014)
19. Bonabeau, E.: Agent-based modeling: Methods and techniques for simulating human systems. Proceedings of the national academy of sciences 99(suppl_3), 7280–7287 (2002)
20. Lee, B.A., Tyler, K.A., Wright, J.D.: The new homelessness revisited. Annual review of sociology 36(1), 501–521 (2010)
21. Sikich, K.W.: Global female homelessness: A multi-faceted problem. Gender Issues 25, 147–156 (2008)
22. Fazel, S., Geddes, J.R., Kushel, M.: The health of homeless people in high-income countries: descriptive epidemiology, health consequences, and clinical and policy recommendations. The Lancet 384(9953), 1529–1540 (2014)
23. Gao, Y., Das, S., Fowler, P.J.: Homelessness service provision: A data science perspective. In: AAAI Workshops (2017)
24. Kube, A., Das, S., Fowler, P.J.: Allocating interventions based on predicted outcomes: A case study on homelessness services. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 33, pp. 622–629 (2019)
25. Chelmis, C., Rahman, K.S.: Why do some homeless succeed while others falter? a network science perspective. IEEE Transactions on Computational Social Systems 11(2), 2703–2712 (2023)
26. VanBerlo, B., Ross, M.A., Rivard, J., Booker, R.: Interpretable machine learning approaches to prediction of chronic homelessness. Engineering Applications of Artificial Intelligence 102, 104243 (2021)
27. Messier, G., John, C., Malik, A.: Predicting chronic homelessness: The importance of comparing algorithms using client histories. Journal of Technology in human ServiceS 40(2), 122–133 (2022)

28. Rahman, K.S., Zois, D.S., Chelmis, C.: Bayesian network modeling and prediction of transitions within the homelessness system. In: ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). pp. 1–5. IEEE (2023)
29. Rahman, K.S., Chelmis, C.: Modeling and predicting individual transitions within the homelessness system. Social Network Analysis and Mining 13(1), 77 (2023)
30. Rahman, K.S., Chelmis, C.: Predictive modeling of homeless service assignment: A representation learning approach. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 39, pp. 28312–28320 (2025)
31. Chelmis, C., Rahman, K.S.: Peeking through the homelessness system with a network science lens. In: Proceedings of the 2021 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining. pp. 69–73 (2021)
32. Chelmis, C., Qi, W., Lee, W., Duncan, S.: Smart homelessness service provision with machine learning. Procedia Computer Science 185, 9–18 (2021)
33. Macal, C.M., North, M.J.: Tutorial on agent-based modeling and simulation. In: Proceedings of the Winter Simulation Conference, 2005. pp. 14–pp. IEEE (2005)
34. De Wolf, T., Holvoet, T.: Towards autonomic computing: agent-based modelling, dynamical systems analysis, and decentralised control. In: IEEE International Conference on Industrial Informatics, 2003. INDIN 2003. Proceedings. pp. 470–479. IEEE (2003)
35. Murdock, I., Carley, K.M., Yağan, O.: An agent-based model of reddit interactions and moderation. In: Proceedings of the International Conference on Advances in Social Networks Analysis and Mining. pp. 195–202 (2023)
36. Davidsson, P.: Agent based social simulation: A computer science view. Journal of artificial societies and social simulation 5(1) (2002)
37. Gilbert, N.: Agent-based social simulation: dealing with complexity. The Complex Systems Network of Excellence 9(25), 1–14 (2004)
38. Mittal, A., Gibson, N.O., Krejci, C.C.: An agent-based model of surplus food rescue using crowd-shipping. In: 2019 Winter Simulation Conference (WSC). pp. 854–865. IEEE (2019)
39. Roorda, M.J., Cavalcante, R., McCabe, S., Kwan, H.: A conceptual framework for agent-based modelling of logistics services. Transportation Research Part E: Logistics and Transportation Review 46(1), 18–31 (2010)
40. Acevedo, V., Chowdhury, L.S.: Modeling sexual violence within the homelessness population through an agent-based approach. The PUMP Journal of Undergraduate Research 2, 58–72 (2019)
41. Aguilera, A., Osman, N., Curto, G.: Agent-based modeling meets the capability approach for human development: Simulating homelessness policy-making. arXiv preprint arXiv:2503.18389 (2025)
42. Lindstrom, M.R., Bertozzi, A.L.: Qualitative features of a nonlinear, nonlocal, agent-based pde model with applications to homelessness. Mathematical Models and Methods in Applied Sciences 30(10), 1863–1891 (2020)
43. Hmis project type descriptions (2020), https://homelessassistanceprograms.com/hmis-project-type-descriptions/, available online