

Why Do Some Homeless Succeed While Others Falter? A Network Science Perspective

Charalampos Chelmiss¹, Member, IEEE, and Khandker Sadia Rahman²

Abstract—Homelessness, a long-standing societal problem, appears to be on the rise, fueled in part by the Covid-19 pandemic. Looking at the homelessness system as a network of interconnected services which individuals traverse over time, we seek to shed light on their progression toward securing stable housing. We formalize the concept of stability upon exit and show that regardless of starting conditions, the ultimate goal is either reached quickly or not at all, indicating the importance of addressing the homeless' needs early on to avoid them "giving up." To better understand the causes that may contribute to positive outcomes for certain individuals versus others, we computationally analyze their pathways through the network of homeless services. We confirm the intuition that some individuals face more challenges than others based on their initial living conditions and initial placement to homelessness services. At the same time, we discover that simple signals can act as good indicators of individuals at risk of "falling through the cracks." Being able to predict such outcomes is critical to design assistive technology that can retain individuals who would otherwise falter.

Index Terms—Applied network science, computational social science, human factors, human services.

I. INTRODUCTION

MANY cities face significant rates of homelessness and a growing number of at-risk individuals fueled by, among other factors, increasing housing costs, limited affordable housing options, and more recently the Covid-19 pandemic. In the USA alone, more than 582 000 people were homeless on a single night in 2022, among which 30% experienced repeated homelessness [10]. In this work, we computationally analyze how chronically homeless individuals (i.e., individuals that have experienced repeated homelessness for two or more times) [9] navigate through the homelessness system over time to shed light into why certain individuals succeed in securing stable housing, while others exit the system without achieving this objective.

To study the progress (or lack thereof) of individuals through the homelessness system over time, we begin by representing individual-level longitudinal homelessness data collected by homeless service providers as a rich network of interconnected homeless services which the homeless individuals traverse over time. Based on this network representation, we computationally analyze the "trajectories" homeless

individuals take over the network once they *enter* the system, as well as examine the factors that affect their status (e.g., stable housing, rental, homeless) upon *exit*. In addition, since the goal of securing stable housing can be achieved by reaching the exit, *abandonment* (i.e., exiting the system without recording a destination) can be interpreted as "giving up." We do acknowledge that unobserved processes not captured in the administrative data could offer an alternative explanation for abandoned trajectories. Nevertheless, our setting enables a reasonable computational analysis of the reasons causing individuals to "give up," which would otherwise be impossible without tracking down and interviewing each and every one of those individuals. In particular, this article makes the following main contributions.

- 1) We model homeless services as a heterogeneous network and study the trails of individuals in need of such services over this network.
- 2) We propose three alternative definitions of stability based on individuals' destination after exiting the system.
- 3) We computationally analyze differences and commonalities between successful, unsuccessful, and abandoned trajectories using a one-of-a-kind longitudinal dataset.
- 4) We show that sometimes individuals follow trajectories that take them further away from the target, and in such cases, the likelihood of *backward transitions* increases dramatically, particularly so for individuals who are eventually unsuccessful in securing stable housing.
- 5) To ensure the reproducibility of our work, we make our source code available on GitHub at <https://github.com/IDIASLab/Homelessness>.

The rest of this article is organized as follows. Section II summarizes related work. Section III outlines the data used. Section IV provides a detailed description of the proposed network models of the homelessness system and their corresponding properties. Section V defines the notion of stable exit. Section VI investigates how key quantities change as individuals progress through the homelessness system, whereas Section VII investigates the factors that differentiate successful and unsuccessful individuals and those that exit the system prematurely. Finally, Section VIII concludes with a discussion of limitations and potential future directions.

II. RELATED WORK

Homelessness is a multifaceted societal issue, exacerbated by regulatory, political, and technological obstacles [6]. The work related to our study can be separated into three main themes, each of which we briefly discuss next.

Manuscript received 30 March 2022; revised 5 April 2023 and 30 June 2023; accepted 1 July 2023. This work was supported by the National Science Foundation under Grant ECCS-1737443. (Corresponding author: Khandker Sadia Rahman.)

This work involved human subjects or animals in its research. Approval of all ethical and experimental procedures and protocols was granted by the University at Albany Institutional Review Board.

The authors are with the Department of Computer Science, University at Albany, Albany, NY 12222 USA (e-mail: cchelmiss@albany.edu; srahman2@albany.edu).

Digital Object Identifier 10.1109/TCSS.2023.3296376

Structural determinants suspected to be contributing or directly leading to homelessness (see [2], [8], [12], [16], [19], [21], [22], [26]) have been well-studied. Individual factors contributing or directly leading to homelessness have also been extensively explored [23]. On the other hand, factors associated with outcomes (e.g., probability of an individual being in need of homeless services once again after exiting the system) after an individual has entered the homelessness system have been of interest [11], [13], [24], [28], [29]. This work complements prior art by offering three progressively exclusive definitions of stability, which are then used to computationally analyze the trails of individuals in need of homeless services.

The next theme of related work examines factors associated with outcomes (e.g., stable housing [24], or inversely, readmission probability [29]) once a homeless individual has entered the homelessness system. Aubry et al. [3] categorize the homeless individuals into typologies based on their frequency and length of stay within stable housing, whereas [5], [7], [15] group individuals based on frequency and length of stay in the homeless system. Subcategories of typologies are described in [18]. The majority of these works focuses on determining demographic and clinical differences between the typologies. In contrast, this study groups trajectories with respect to their exit information and studies the characteristics of the pathways leading to each specific exit (e.g., length of pathways, similarity between services, and stable exit at each step).

The last theme of related work can be traced back to the problem of algorithmic homeless service delivery [4], [14], [25]. This theme focuses on designing fair and interpretable algorithms for allocating scarce housing resources with diverse characteristics and eligibility criteria (e.g., emergency shelter versus permanent supportive housing) to individuals on a waiting list so as to increase the expected number of stably housed individuals upon exit of the homelessness system. Our focus is beyond accurate screening (i.e., allocation decisions for arriving homeless individuals to one of different types of housing resources). Specifically, our study is unique in that it is the first to computationally investigate and model the observed pathways of individuals through the homelessness system over time. We believe that the insights obtained from such analysis can be used to develop accurate predictive models, which in turn will be used to inform better machine learning methods for more effective homeless services' allocation.

III. DATA

A. Setting

Most communities in the USA rely on the so-called Coordinated Entry System, according to which homeless people who sign up for housing support (e.g., emergency shelter) are accessed for eligibility and vulnerability and are subsequently prioritized for housing based on such assessments and availability of resources. Services, organized by *project type*, are offered by shelters [27] as part of the continuum of care (CoC), i.e., the ecosystem for the homeless. Table I provides a description of project types [27].

Homeless service providers must routinely collect data about their clients, mainly for reporting to both the local

TABLE I
DESCRIPTION OF PROJECT TYPES [27]

Number	Project Type	Description
1	Emergency shelter	A project that offers temporary shelter for the homeless in general or for specific homeless populations. Requirements and limitations may vary by project and are specified by the funder.
2	Transitional housing	A project that provides temporary lodging and is designed to facilitate the movement of homeless individuals and families into permanent housing within a specified period of time, but no longer than 24 months. Requirements and limitations may vary by project and will be specified by the funder.
3	Permanent supportive housing	A project that offers permanent housing and supportive services to assist homeless persons with a disability (individuals with disabilities or families in which one adult or child has a disability) to live independently.
4	Street outreach	A project that offers services necessary to reach out to unsheltered homeless, connect them with emergency shelter, housing, or critical services, and provide urgent, non-facility-based care to unsheltered individuals who are unwilling or unable to access emergency shelter, housing, or an appropriate health facility.
6	Services only	A project that offers only stand-alone supportive services (other than outreach or coordinated entry) to address the special needs of participants (such as child care, employment assistance, and transportation services) and has associated housing outcomes.
11	Day shelter	A project that offers daytime facilities and services (no lodging) for persons who are homeless.
12	Homelessness prevention	A project that offers services and/or financial assistance necessary to prevent a person from moving into an emergency shelter or place not meant for human habitation.
13	Rapid rehousing	A permanent housing project that provides housing relocation and stabilization services and short- and/or medium-term rental assistance as necessary to help an individual or family move as quickly as possible into permanent housing and achieve stability.
14	Coordinated entry	A project that administers the continuum's centralized or coordinated process to coordinate assessment and referral of individuals and families seeking housing or services, including use of a comprehensive and standardized assessment tool.

authorities and national agencies, such as the USA Department of Housing and Urban Development (HUD). According to federal mandates, service providers collect data for all their clients into the Homeless Management Information System (HMIS), an information technology system founded by HUD and locally administered by a lead organization. Each record comprises individual-level data, including personally identifying information, socioeconomic characteristics and educational background, and health information, as well as dates and types of services received, and fields specifying the type of exit from homelessness (e.g., interim housing or hotel) and whether an individual has reentered the system multiple times [27].

B. Data Description

Our analysis is based on the data provided by CARES of NY Inc., a nonprofit organization that locally manages the HMIS for the Capital Region of the state of New York (officially defined as the Albany–Troy–Schenectady Metropolitan Statistical Area with a population size of $\sim 800\,000$ people). New York has experienced a 40% increase in the homeless population between 2007 and 2022 with a 17.7% increase among sheltered individuals experiencing chronic patterns of homelessness. Beyond New York, chronically homeless individuals in the USA increased by 16% between 2020 and 2022 [10]. In total, the dataset comprises 50 469 records of all services provided to 38 954 individuals by organizations in

the region between 2012 and 2018. Despite the narrow spatial focus of the dataset, our methodology is not restrained to this specific geographic boundary; it is applicable to any location across the USA, and potentially across the world, where administrative records of homeless services are systematically collected.

A complete description of the data elements to be stored within HMIS and their function and specific use are available at [27]. For reference, the most relevant data to our study include, but are not limited to: 1) the start and end dates of each service, transitions between services, exits, and reentries (i.e., receipt of services after exiting the system); 2) housing outcomes for those who exited the system; 3) duration between exit and return to homelessness (if applicable); 4) demographics (e.g., age and gender), education history (e.g., last grade completed), and disabling condition (if any); and 5) time-variant properties, such as monthly income and health (e.g., mental state). Individuals are identified by a unique and anonymous identifier. We focus on 6011 individuals (15.4% of the total number of individuals in our dataset) experiencing chronic patterns of homelessness, i.e., entering the homeless system multiple times, and the complete history of services they received (18818 records in total) from 125 project IDs,¹ which are in turn split into nine broad categories (i.e., project types), summarized in Table I.

IV. MODELING THE HOMELESSNESS SYSTEM AS A NETWORK

To shed light into the progress (or lack thereof) of individuals through the homelessness system, we begin by constructing two transition graphs from our historical data, namely: 1) a small network of project types, $\mathcal{G}_P$ and 2) a larger, and sparser, network of project IDs, $\mathcal{G}_I$. In the first graph, a node corresponds to a project type, whereas in the second to a project ID. A directed edge from node i to j captures the number of transitions from i to j recorded in the dataset. Self-loops from a node to itself (if present) are indicative of a transition to the same project (either type or id depending on the corresponding graph).

The key steps to construct either transition graph are illustrated in Fig. 1. We begin by constructing a trajectory for each individual i based on his/her enrolment records starting with the earliest entry date. The first node in this trajectory is an “entry” node, followed by the project type (similarly for project ID) associated with the first enrolment record. For every enrolment record for i , an edge is created between the previous node and the record with the next earliest entry date. In the event of overlapping services (e.g., when an individual is assigned to a project prior to exiting from the previous one), we update the exit date of the first enrolment to the start date of the subsequent enrolment. This process continues for every enrolment associated with an individual in order of entry date, until all records have been processed. Finally, an edge to node “exit” is added from i ’s last enrolment record, and the exit type (e.g., “permanent housing for formerly homeless persons”) is recorded. To construct each trajectory of a given

¹A unique identifier is automatically generated by the HMIS at the time the project is created in the HMIS.

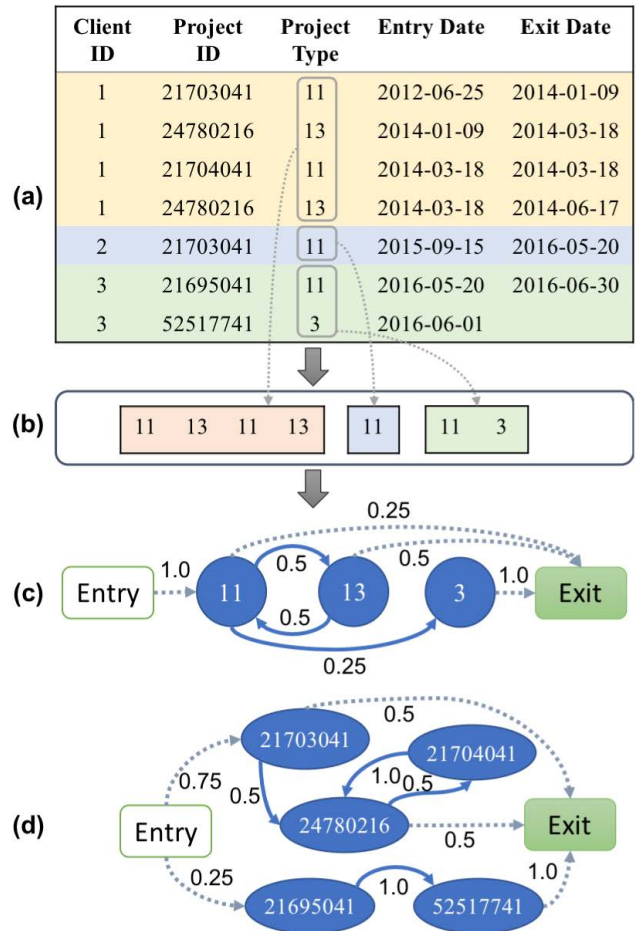


Fig. 1. Illustration of the transition graph construction process. (a) Sample enrolment data for three clients. (b) Project-type trajectories for those clients. (c) Obtained project-type transition graph. (d) Transition graph for project IDs is obtained in a similar manner.

individual, we search through all the records for their first service, last service, and overlapping services. Accordingly, the computational complexity of constructing trajectories is $O(mn)$, where n is the number of records and m is the number of individuals. Once the list of trajectories has been compiled, a transition graph between the nine project types (similarly for project IDs) and the two special nodes denoting entry and exit is constructed. In the transition graph, a directed, weighted edge e_{p_i, p_j} between nodes p_i and p_j encodes the number of trajectories in the historical data containing a transition from node p_i to p_j . Finally, to ensure that edge weights represent the probability of moving from one project type (similarly for IDs) to another, we normalize the outgoing edges from a node to add up to 1. For each edge, $e_{(p_i, p_j)}$, we search for p_i followed by p_j within all records of trajectories and count their co-occurrence to compute the edge weights between nodes. Consequently, computing the transition graph takes $O(|V|^2 \times n)$, where n and V denote the number of records and the set of nodes (i.e., project types for $\mathcal{G}_P$ and IDs for $\mathcal{G}_I$) accordingly.

As expected, the transition graph of project types is well-connected, whereas the transition graph of project IDs comprises 55 strongly connected components, the largest of which encompasses just 10% of all the nodes (see Table II).

TABLE II
HIGH-LEVEL NETWORK STATISTICS. NODES “ENTRY” AND
“EXIT” ARE NOT CONSIDERED IN THIS ANALYSIS

Metric	$\mathcal{G}_P$	$\mathcal{G}_I$
Number of nodes	9	125
Number of edges	61	428
Density	0.847	0.028
Average degree	6.778	3.424
Min/Max In-degree	6/8	1/11
Min/Max Out-degree	6/8	1/12
Diameter	2	4
Average path length	1.264	1.524
Number of WCCs	1	22
Size of largest WCC	100%	11.93%
Number of SCC	1	55
Size of largest SCC	100%	9.6%

The small diameter, when contrasted with the empirical length of trajectories in our dataset (see Section VI) already hints toward “inefficient” paths.

V. STABLE EXIT DEFINITION

Upon entering the homelessness system, individuals can follow very different paths. Pairing an individual with a specific project may depend on factors including personal characteristics, which determine eligibility (e.g., age or veteran status) and/or impose constraints (e.g., substance abuse), as well as availability of beds at a particular shelter at any given time. Although capacity constraints are not available to us, we set forth to explore the relationship between housing outcomes for those who orderly exit the homelessness system, and those that “fall through the cracks” (i.e., not receiving the services or care they need to reach stable housing), and the corresponding path (i.e., trajectory) they take over the transition graphs.

A definition of *stable exit* is lacking in the literature. For this reason, we propose, along with Dr. Wonhyung Lee, a trained *social scientist*, three alternative definitions of stable exit based on individuals’ destination after exiting the system, as shown in Tables III and IV. Specifically, Scheme 1 (S_1) categorizes destination types categorized as stable or unstable. Scheme 2 (S_2) is based on HUD’s destination types [1] that are in turn used to characterize what is considered positive or negative outcomes for street outreach, and permanent or temporary destinations for all other project types. Scheme 3 (S_3) differentiates the degrees of progress toward reaching the “ultimate goal.” For example, individuals in transitional housing can be considered to be homeless since their tenure is temporary. When available information is not adequate to determine stability, the term “unknown,” “other,” or “hard to judge” is used. In summary, our definition of stable exit progresses from inclusive in Scheme 1 to strict in Scheme 3.

VI. EFFICIENCY OF THE HOMELESSNESS SYSTEM

Our model of the homelessness system (see Section IV) provides an empirical estimate of the transition probabilities between services. A natural question to ask is how efficient is the homelessness system in terms of promoting effective and efficient trajectories, and how often do the homeless actually navigate such trajectories?

TABLE III
COLOR CODING OF EXIT DESTINATIONS BY SCHEME. HEREAFTER,
SCHEMES ARE ABBREVIATED AS S_1 , S_2 , AND S_3 , RESPECTIVELY

Scheme 1	Scheme 2	Scheme 3
Stable	Permanent	Ultimate goal
Unstable	Temporary	Closer to exit
Unknown	Institutional	Transitional phase
	Other	No progress or worse
		Hard to judge

TABLE IV
EXIT DESTINATIONS CATEGORIZED BY SCHEME, EACH
OF WHICH IS DEFINED IN TABLE III

S_1	S_2	S_3	Exit destination description
			Rental by client, no ongoing housing subsidy
			Owned by client, no ongoing housing subsidy
			Permanent housing (other than RRH) for formerly homeless persons
			Rental by client, with GPD TIP housing subsidy
			Rental by client, with VASH housing subsidy
			Rental by client, with other ongoing housing
			Owned by client, with ongoing housing subsidy
			Staying or living with family, permanent tenure
			Staying or living with friends, permanent tenure
			Moved from one HOPWA funded project to HOPWA Permanent Housing
			Rental by client, with RRH or equivalent subsidy
			Staying or living with friends, temporary tenure (e.g., room apartment or house)
			Hotel or motel paid for without emergency shelter voucher
			Staying or living with family, temporary tenure (e.g., room, apartment, or house)
			Transitional housing for homeless persons (including homeless youth)
			Moved from one HOPWA funded project to HOPWA Transitional Housing
			Long-term care facility or nursing home
			Foster care home or foster care group home
			Safe Haven
			Emergency shelter, including hotel or motel paid for with emergency shelter voucher, or RHY-funded Host Home shelter
			Place not meant for habitation (e.g., a vehicle, an abandoned building, us/train/subway station/airport or anywhere outside)
			Psychiatric hospital or other psychiatric facility
			Substance abuse treatment facility or detox center
			Hospital or other residential non-psychiatric medical facility
			Residential project or halfway house with no homeless criteria
			Jail, prison or juvenile detention facility
			Clients does not know
			Client refused
			Other
			Deceased
			No exit interview completed
			Data not collected

Fig. 2 shows how the length of actual paths taken by the homeless (green) compared to *effective* (a.k.a. “always moving forward”) trajectories (red), i.e., paths that ignore transitions to services individuals had received in the past.² For example, the effective trajectory of client 1 in Fig. 1 is {11, 13}. As expected, effective trajectories are short (two steps on average), whereas the length of actual paths shows great variance. Nevertheless, because of the distribution nicely fitting a power law (dashed black line), the majority of actual paths is not significantly longer. This finding in turn leads to question the reason for the large variance in actual paths

²We call trajectories including such transitions *backward* trajectories.

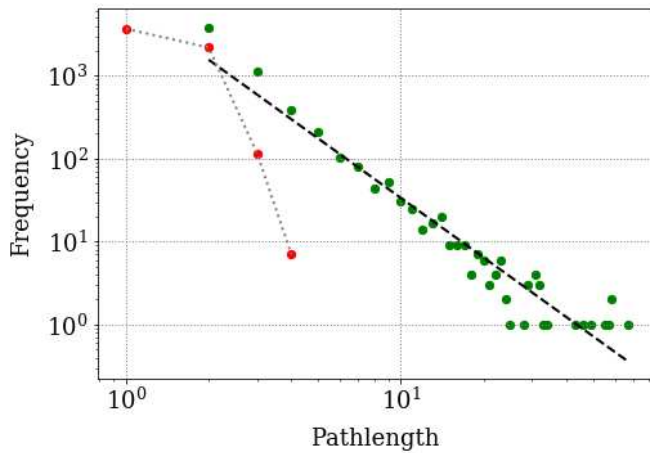


Fig. 2. Distribution of complete (green) and effective (red) trajectory lengths. Least-square power-law fit with slope of -2.38 is shown in black.

taken by the homeless. One possibility is that some individuals are better positioned than others (e.g., do not suffer from disabilities or substance abuse issues). Another possibility is that some types of exit are inherently harder than others (e.g., stable exit versus returning to homelessness).

To better understand the trajectory of individuals based on their status upon entry, we use the three schemes introduced in Section V. This is only possible because HMIS uses the same categorization for living conditions upon entry into the system and exit destination. Fig. 3(a) shows the cumulative distribution of path lengths for those who enter the system from an unstable (left) or stable (center) situation, or unknown status (right) according to Scheme 1, to exit with a corresponding destination in k steps. In all the cases, the distributions are heavy tailed, with unstable exits being reached when longer trajectories are taken, regardless of the starting point. Moreover, individuals starting from unstable conditions take longer trajectories before exiting the system to a stable destination. Finally, individuals exiting the system with an unknown or unstable destination follow longer paths on average when entering the system from stable or unknown conditions. We believe this to be indicative of individuals experiencing chronic patterns of homelessness. According to Scheme 2 [see Fig. 3(b)], paths to permanent housing require relatively few transitions between project types in most cases, with the exception of temporary starting conditions. Similarly, an institutional exit destination is often reached quickly. After a more careful examination of the trajectories of those entering the system from a temporary starting point, it becomes clear that a temporary entry status leads to longer trajectories through the system than any other starting point. Finally, the path length of trajectories leading to transitional destinations in Scheme 3 is comparatively longer than hard to judge exit points [see Fig. 3(c)]. Similarly, those starting from a closer to exit position are better posed to reach the ultimate goal quickly. For those entering the system from a transitional phase or no progress or worse entry point, paths are longer overall. In most cases, paths leading to a no progress or worse exit are comparatively shorter than paths leading to other types of exit points. Moreover, paths starting from “hard

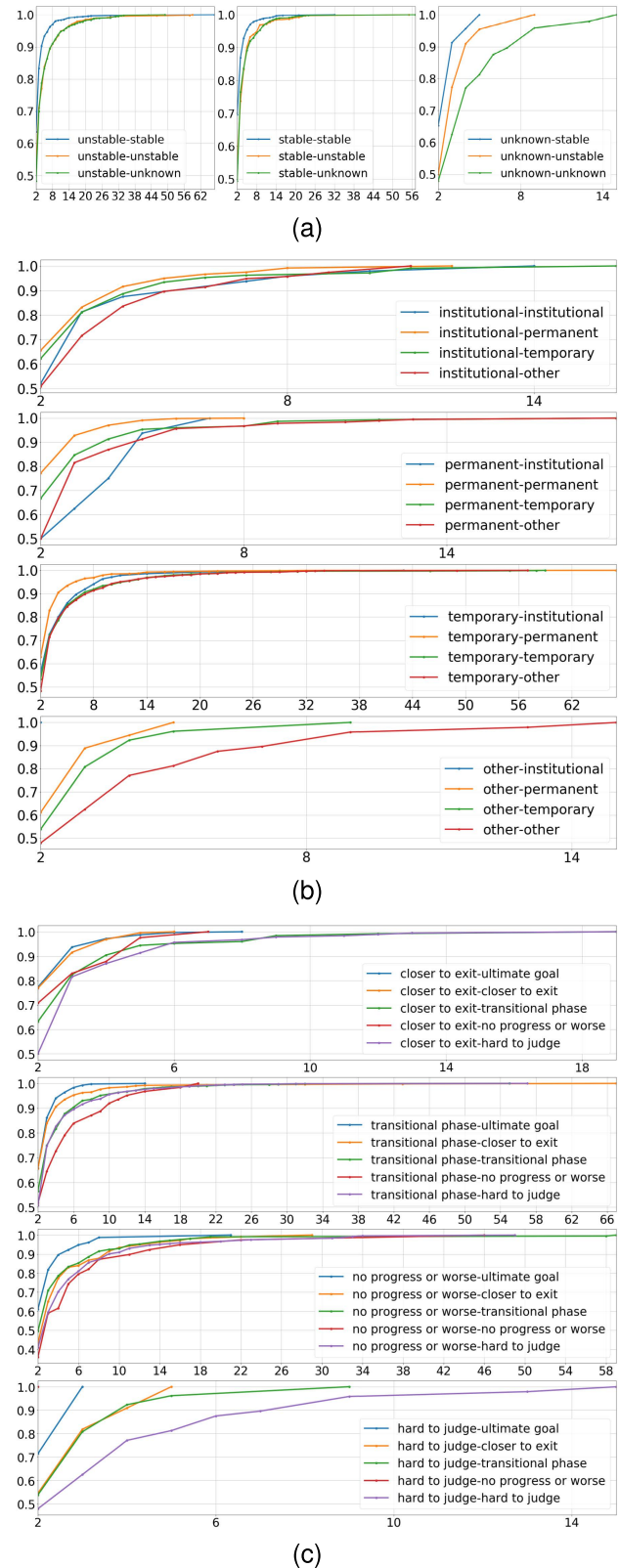


Fig. 3. Cumulative distribution function for path length as a function of entry and exit statuses. Entry and exit statuses are defined with respect to each of the three schemes presented in Tables III and IV. (a) Scheme 1. (b) Scheme 2. (c) Scheme 3.

to judge” are significantly shorter than those starting from “transitional phase” and “no progress or worse,” indicating that those individuals may have given up. The key insight

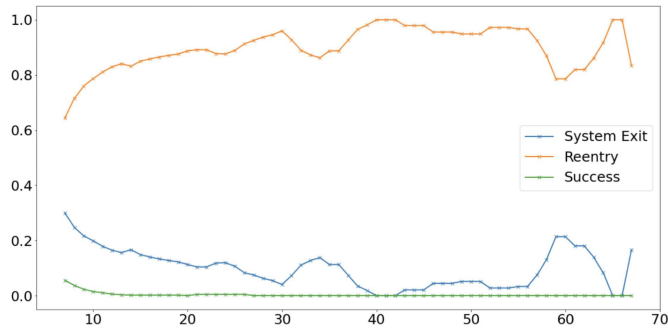


Fig. 4. Success, reentry, and abandonment probability (y-axis) according to Scheme 3, as a function of path position (x-axis).

is that the ultimate goal is either reached quickly or not at all, regardless of starting conditions. Therefore, it is crucial to minimize the number of transitions an individual has to make within the homelessness system.

A. Network Structure Versus Relevance

What about individuals exiting the system before reaching a stable exit? To examine whether there is any correlation between the length of trajectories and outcomes, we plotted the empirical probability of an individual reaching her goal (i.e., stable exit), reentering the system in the future, and exiting the system with a status other than stable exit, as a function of path length. Fig. 4 shows the result for Scheme 3. Nevertheless, this result is consistent for all the schemes. Intuitively, the probability of stable exit is higher for shorter paths and drops to zero by path length 14. At the same time, the probability of reentry increases sharply for paths involving more than ten transitions. Conversely, the probability that an individual “gives up” slowly decreases with path length. We believe this finding to be indicative of both persevering individuals, and the challenge to address the needs of homeless individuals as early as possible.

Finally, we wish to study the effect (if any) of the transition graph in promoting trajectories toward achieving a positive outcome. Prior work on real-world social networks has shown that the probability of a node linking to its r th closest node is inversely proportional to r [17]. The so-called “small-world” property of such networks is particularly desirable, since short chains can be discovered in every network exhibiting this property [17]. If this property were to be true for the transition graph, it would be indicative of an inefficient system in terms of helping an individual reach a stable exit. Specifically, transitioning to node j from i , with j being highly correlated to i (e.g., both being emergency shelters), would offer little help in achieving a positive outcome.

Since similarity between project types (similarly for project IDs) is not directly recorded in homelessness data, we proceed to estimate such similarity from trajectories. Specifically, trajectories associate individuals (I) and their corresponding characteristics (F) and project types (P) based on the services each individual receives at a given time as shown in Fig. 5. We use these associations to construct a heterogeneous graph of individuals, their features, with feature values treated as

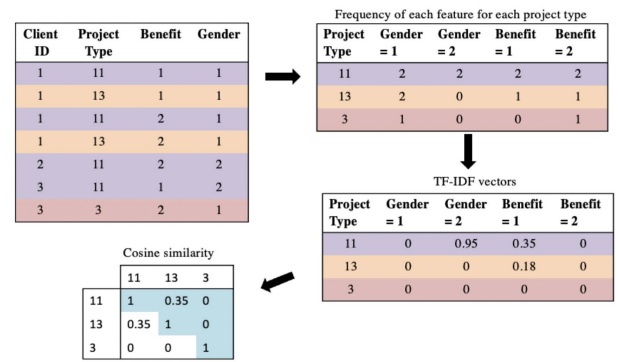


Fig. 5. Tripartite network representation of the homeless, their corresponding features, and project types (similarly for project IDs).

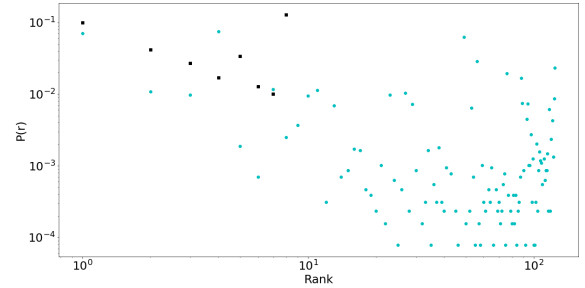


Fig. 6. Link probability as a function of rank for the network of project types (black squares) and project IDs (blue circles) accordingly.

discrete, and project types, with edges between the nodes representing different types of relations.

We represent this graph as an $(m + n + w) \times (m + n + w)$ symmetric adjacency matrix, with vertices of type P in one layer ($|P| = m$), vertices of type I in the second layer ($|I| = n$), and feature values F in the third layer ($|F| = w$)

$$A = \begin{bmatrix} PP_{m \times m} & PI_{m \times n} & PF_{m \times w} \\ IP_{n \times m} & II_{n \times n} & IF_{n \times w} \\ PF_{w \times m} & FI_{w \times n} & FF_{w \times w} \end{bmatrix}.$$

Here, $A_{i,j+m} = A_{i+m,j} = PI_{ij} = IP_{ij}$ encodes the fact that individual j has been assigned to project i , $A_{i+m,j+m+n} = A_{i+m+n,j+m} = IF_{ij} = FI_{ij}$ encodes the connection between individual i and feature node j , and diagonal submatrices PP , II , and FF are zero since our tripartite representation describes only interlayer edges.

To obtain a vector representation of project types over the feature space, we compute the second power of the adjacency matrix (i.e., A^2) as shown in Fig. 5. We then calculate the similarity between two projects as the cosine of their TF-IDF vectors [20] using the now filled $PF_{m \times w}$ submatrix. Applying the same methodology [17] used to compute similarity of nodes in real-world social networks, we compute $P(r)$, $\forall r$ as the fraction of nodes j that i links to out of all the project pairs (i, j) , such that j is the r th most similar to i among all the projects.

Fig. 6 shows the result. The overall downward trend both for project types and IDs indicates that the probability of a node being connected with an outward edge to another node is higher the more similar the two nodes are. As mentioned above, such scenario is undesirable since

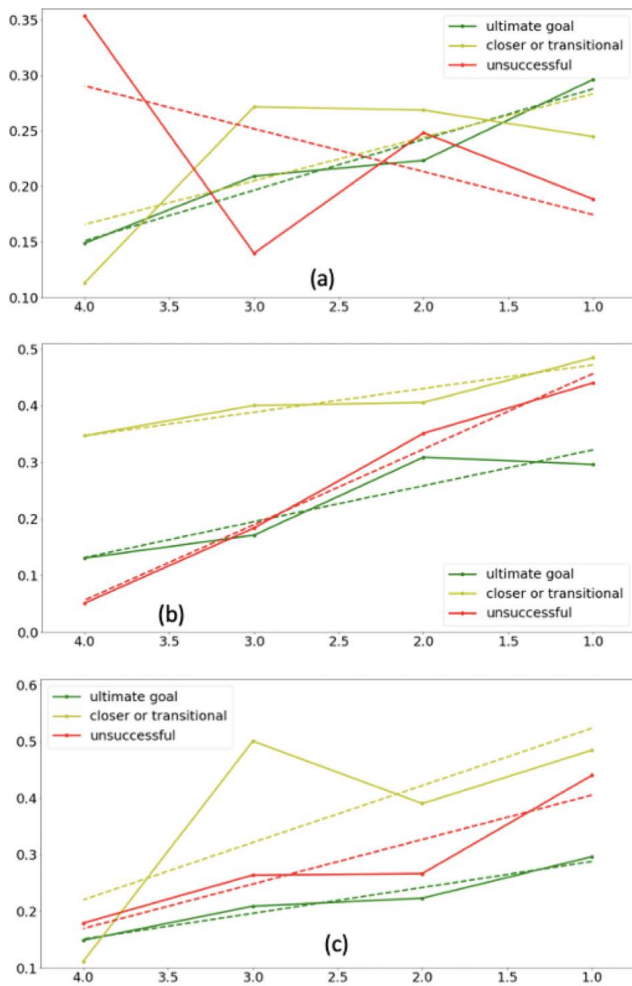


Fig. 7. Evolution of similarity as individuals progress toward different exits. The x - and y - axes show, respectively, the distance to exit and the cosine of the TF-IDF vectors. Dotted lines indicate best fit. (a) Similarity to the ultimate goal. (b) Similarity to the next step. (c) Similarity to the actual exit.

“short-range” links between project types or IDs of similar function (e.g., emergency shelters) may not bring individuals closer to their ultimate goal. At the same time, “long-range” edges between less similar nodes are less frequent.

B. Progressing Toward Stable Exit

Next, we examine similarity between project types as individuals progress toward “ultimate goal,” “transitional,” or “unsuccessful” states according to Scheme 3—transitional here also includes the individuals leading to “closer to exit” and unsuccessful includes “no progress or worse” and “hard to judge.” Specifically, we compute the cosine similarity of the TF-IDF vectors of the current project type and “ultimate goal” [Fig. 7(a)], or the following project type [Fig. 7(b)]. Conceptual similarity with the ultimate goal increases for those individuals that in the end reach, or come close to, the ultimate goal. On the other hand, similarity decreases for “unsuccessful” individuals. We interpret this result as indicative of lack of improvement. At the same time, individuals exiting as “closer or transitional” or “unsuccessful” exhibit much higher similarity to their actual exit (i.e., the

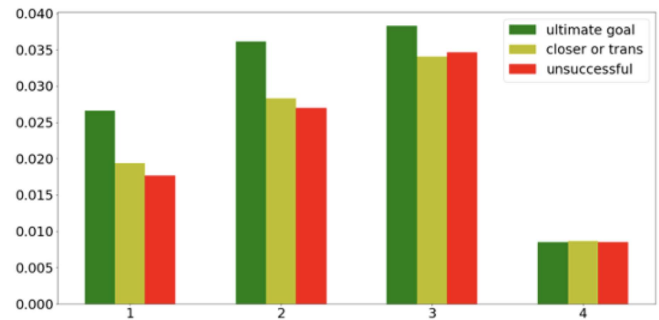


Fig. 8. Similarity between initial living condition with “ultimate goal” as a function of pathway length. Note that only effective paths are included in this analysis.

exit information recorded in the data) and next step [Fig. 7(b) and (c), respectively]. We interpret this result as indicative of individuals being assigned to services that do not necessarily facilitate reaching the “ultimate goal.” Finally, Fig. 8 shows how the initial living conditions may contribute to the ability of an individual to progress toward the “ultimate goal.” Individuals exiting to “ultimate goal” exhibit higher similarity to the ultimate goal at the initial stages (further confirmed in the last row of Table V). At the same time, the similarity score for unsuccessful individuals is low. Interestingly, longer trajectories (i.e., those with effective length of 4) exhibit low similarity for both successful and unsuccessful trajectories. We suspect this to be an artifact of aggregation of data collected over time as individuals receive more and more services. In summary, we conclude that the initial conditions of individuals upon entering the system are crucial into how they end up navigating the network of homelessness services.

Beyond the initial living conditions of an individual, it is unclear, up to this point, what additional factors may increase (or decrease) her ability to reach the ultimate goal. To elicit factors independent of individuals’ living conditions, we compare properties of the project that was visited first and last, and the ultimate goal for successful, unsuccessful, and abandoned pathways. The results are summarized in Table V. We make the following observations. First, the out-degree (out-degree), in-degree (in-degree), and PageRank of projects visited in abandoned pathways are lower than in successful ones, particularly so for projects visited last. Since these properties capture how easy it is to reach the particular project, we can conclude that trajectories visiting less connected projects are more likely to be abandoned. It is important to note that PageRank is based on a random-walker model, whereas homeless individuals are not expected to navigate the homelessness system at random. The fact that PageRank is a good indicator of whether an individual may be at risk of leaving the system is nontrivial and of practical importance (e.g., may be useful as a feature for predicting abandonment). Second, both the out-degree and in-degree of projects visited first and last are similar in successful and unsuccessful pathways. Therefore, these properties alone may not be indicative of success, and other confounding factors must be considered.

Fig. 9 shows the similarity between project types (1–14) and exit types (i.e., ultimate goal, transitional or closer, or unsuccessful). Initially, i.e., after receiving the first service upon entering the homelessness system, it is unclear whether an

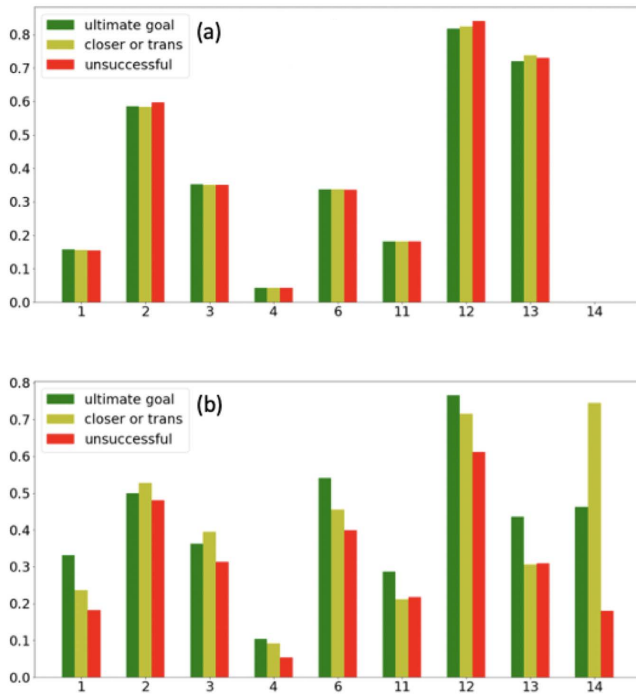


Fig. 9. Average similarity at (a) first step upon entry and (b) last step before exiting.

TABLE V

AVERAGE VALUES OF PROPERTIES FOR SUCCESSFUL, UNSUCCESSFUL, AND ABANDONED PATHWAYS

Property	Successful	Unsuccessful	Abandoned
Out-degree of first	8.12	8.14	7.97
In-degree of last	8.66	8.15	7.46
PageRank of first	0.12	0.15	0.10
PageRank of last	0.10	0.14	0.07
Cosine of first with ultimate goal	0.31	0.23	0.25

individual is going to succeed or not [Fig. 9(a)]. In contrast, similarity to ultimate goal drops significantly for unsuccessful individuals in the last step [Fig. 9(b)]. Interestingly, individuals being assigned to projects 1, 6, and 12–14 in their last step before exiting the system appear to have a significant advantage over unsuccessful individuals toward reaching the ultimate goal. At the same time, projects 2, 3, 6, 12, and 13 exhibit consistently high similarity with the ultimate goal both initially and at the last step. Conversely, project types 4, 1, and 11 have lower similarity with the ultimate goal, both initially and at the last step, signifying that transiting through these services is less likely to lead to the ultimate goal. For reference, the description of project types corresponding to the numbers reported above is provided in Table I.

VII. WHAT DIFFERENTIATES SUCCESSFUL, UNSUCCESSFUL, AND ABANDONED PATHWAYS?

Fig. 4 suggests that at every step there is a nonnegligible probability of an individual either exiting the system only to return later on or “giving up” altogether. At the same time, Figs. 7 and 8 provide some hints about why some individuals may give up without reaching the goal of securing stable

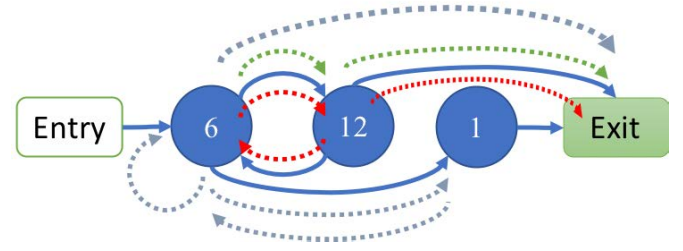


Fig. 10. Sample pathways over a network of three homelessness services. Green arrows illustrate an efficient and successful trajectory, whereas the red arrows show an inefficient and unsuccessful trajectory. Gray arrows show an inefficient, yet successful trajectory.

TABLE VI

STATISTICS OF BACKWARD TRANSITION WITHIN SUCCESSFUL, UNSUCCESSFUL, AND ABANDONED PATHWAYS

	Successful	Unsuccessful	Abandoned
Percentage of at least one backward transition	73.4	82.6	65.4
Average number of backward transitions	1.45	2.52	1.36
Average cycle length	1.11	1.15	1.07

housing. To further investigate the factors that contribute to unsuccessful exits, and more importantly, individuals giving up, we include in our analysis pathways that contain backward transitions and look into how successfully finished pathways differ from both unsuccessful and prematurely abandoned trajectories. We also examine the importance of efficiency, defined as the least number of *backward transitions* (i.e., transitions to services that have already been used in the past). Such transitions result in *cycles* within trajectories. Fig. 10 shows example trajectories of varying degrees of efficiency and end result (i.e., successful or unsuccessful). Abandoned trajectories correspond to those individuals that exit the system without reaching the exit node (i.e., their destination upon departure is unknown).

A priori, it is unclear whether backward transitions are a sign of successful or unsuccessful pathway. After all, backward transitions could equally indicate either unavailability of resources (e.g., an individual must be placed in an emergency shelter because all beds in long-term shelters are occupied) or placement in more appropriate services. Overall, out of 83% of the pathways containing at least one cycle, we found that backward transitions are more common in unsuccessful trajectories, as shown in Table VI. We additionally observed that unsuccessful pathways have higher number of backward transitions, on average, in comparison to successful and abandoned pathways, indicating that individuals with unsuccessful trajectories tend to reenter the same project type numerous times. This hints toward the need to assign individuals to less related project types so as to prevent leading them to unsuccessful exits. At the same time, the average cycle length per trajectory is almost invariant, regardless of the end result. However, Fig. 11 shows that unsuccessful trajectories tend to have longer cycles, indicating that even after visiting less related project types, the chance of backward transitions is not zero. We deduce that in hindsight, backward transitions play a

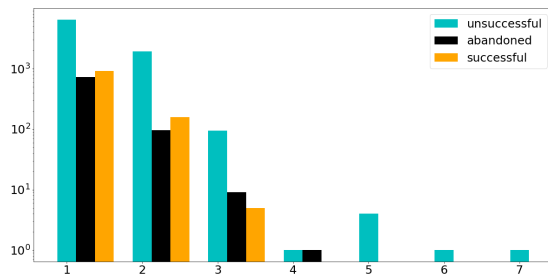


Fig. 11. Frequency of cycle lengths for successful, unsuccessful, and abandoned paths.

TABLE VII

PROBABILITY OF A BACKWARD TRANSITION CONDITIONAL ON INCREASE, DECREASE, OR NO CHANGE IN COSINE SIMILARITY FOR SUCCESSFUL, UNSUCCESSFUL, AND ABANDONED TRAJECTORIES

	Increase	Decrease	No change
Successful	0.88	0.04	0.997
Unsuccessful	0.90	0.29	1
Abandoned	0.62	0.92	1

vital role in distinguishing unsuccessful trajectories from other trajectory types.

We conclude our analysis by computing the probability of an individual making a backward transition depending on whether he/she made progress toward the target in the last step. Table VII summarizes the results. Note that for every trajectory, a backward transition is most likely when cosine similarity remains unchanged in the previous step. For successful and unsuccessful trajectories, the probability of a backward transition is high when cosine similarity increased in the last step and low when decreased. On the contrary, an opposite trend is observed for the abandoned trajectories indicating that the probability of backward transition conditional on increase or decrease in cosine similarity could potentially differentiate abandoned trajectories from the other categories of trajectories. Furthermore, the probability of a backward transition when cosine similarity decreases is considerably lower in successful than in unsuccessful trajectories, and even more so compared with abandoned trajectories.

Two insights can be gained from these observations. First, even though the underlying network structure is unknown to homeless individuals, both successful and unsuccessful individuals seem to be able to appreciate transitions that bring them closer to a desirable outcome, as measured by a decrease in cosine similarity (see Section VI-B). Second, successful individuals are better at forming this intuition, whereas unsuccessful individuals are seven times more likely to go backward after inching closer to the target. This is even more pronounced in the case of individuals who ultimately abandon their goal of stable housing.

VIII. CONCLUSION AND FUTURE WORK

A. Contributions

In this work, we used a one-of-a-kind dataset of administrative records collected by homeless service providers to shed light into the progress of individuals once they enter the homelessness system toward securing stable housing. Apart

from modeling the homelessness system as a network, which individuals traverse over time, we operationalized the notion of stable exit, which we subsequently used to computationally analyze differences and commonalities between successful, unsuccessful, and abandoned trajectories. Our analysis confirmed the intuition that some individuals face more challenges than others based on their initial living conditions and initial placement to homelessness services. However, simple signals, such as the PageRank of a project, can potentially act as good indicators for predicting abandonment, and therefore take action to assist the individuals at risk of “falling through the cracks.” We additionally showed that sometimes individuals follow trajectories that take them further away from the target, and in such cases, the likelihood of *cycles* increases dramatically, particularly so for individuals who are eventually unsuccessful in securing stable housing.

B. Limitations

Next, we would like to note the limitations of this work, which at the same time point to interesting future research directions. First, we acknowledge that our observations are based on a single dataset, geographically bounded to the Capital Region of the state of New York. Second, we rely on administrative data for receipt of services, which although indicate need, do not necessarily capture complete (un)availability of services at any given time. Both these limitations may have resulted in a biased sample. Because of this last limitation, causal effects of positive outcomes or other confounding variables, such as capacity constraints, need to be further explored.

C. Future Directions

By modeling the homelessness system as a network, and identifying simple, yet good predictors of “at-risk” individuals, we anticipate new functionality to be developed toward solving this long-standing societal problem. We anticipate for instance the development of predictive models to identify individuals at risk of abandonment or “backward” transitions. Being able to predict such outcomes is critical to design assistive technology that can retain individuals who would otherwise have given up. On the other hand, comparing the demographic and clinical characteristics of individuals following different trajectories (both upon entering the system for the first time, and after existing for the least time, as well as in between transitions) has the potential to correlate additional contributing factors with different outcomes.

GLOSSARY

diameter

The shortest distance between the nodes furthest apart from each other. 4

in-degree

The number of incoming edges to a node of a directed graph is called in-degree of the node. 4, 8

out-degree

The number of outgoing edges from a node of a directed graph is called out-degree of the node. 4, 8

PageRank

PageRank measures the importance a node based on the number and quality of the incoming edges. 8, 9
 project ID
 Project types are further categorized into project IDs. 3, 4, 6, 7
 project type
 Project type is the different categories of services provided by the homeless service providers. 2–9
 scheme
 Scheme is defined as a categorization of exit destinations. 4–7

REFERENCES

- [1] (2016). *System Performance Measure 7: Destination Classification*. Accessed: May 26, 2021. [Online]. Available: <https://www.hudexchange.info/resource/4966/system-performance-measure-7-destination-classification/>
- [2] I. Anderson and J. Christian, “Causes of homelessness in the U.K.: A dynamic analysis,” *J. Community Appl. Social Psychol.*, vol. 13, no. 2, pp. 105–118, 2003.
- [3] T. Aubry et al., “Housing trajectories, risk factors, and resources among individuals who are homeless or precariously housed,” *Ann. Amer. Acad. Political Social Sci.*, vol. 693, no. 1, pp. 102–122, Jan. 2021.
- [4] M. J. Azizi, P. Vayanos, B. Wilder, E. Rice, and M. Tambe, “Designing fair, efficient, and interpretable policies for prioritizing homeless youth for housing resources,” in *Proc. Int. Conf. Integr. Constraint Program., Artif. Intell., Oper. Res.* Cham, Switzerland: Springer, 2018, pp. 35–51.
- [5] L. Benjaminsen and S. B. Andrade, “Testing a typology of homelessness across welfare regimes: Shelter use in Denmark and the USA,” *Housing Stud.*, vol. 30, no. 6, pp. 858–876, Aug. 2015.
- [6] D. Culhane, “The potential of linked administrative data for advancing homelessness research and policy,” *Eur. J. Homelessness Volume*, vol. 10, no. 3, pp. 109–126, 2016.
- [7] D. P. Culhane, S. Metraux, J. M. Park, M. Schretzman, and J. Valente, “Testing a typology of family homelessness based on patterns of public shelter utilization in four us jurisdictions: Implications for policy and program planning,” *Housing Policy Debate*, vol. 18, no. 1, pp. 1–28, 2007.
- [8] M. Elliott and L. J. Krivo, “Structural determinants of homelessness in the United States,” *Social Problems*, vol. 38, no. 1, pp. 113–131, Feb. 1991.
- [9] M.-J. Fleury, G. Grenier, J. Sabetti, K. Bertrand, M. Clément, and S. Brochu, “Met and unmet needs of homeless individuals at different stages of housing reintegration: A mixed-method investigation,” *PLoS ONE*, vol. 16, no. 1, Jan. 2021, Art. no. e0245088.
- [10] M. Henry, T. de Sousa, C. Roddey, S. Gayen, and T. J. Bednar, “The 2022 annual homeless assessment report (AHAR) to congress. Part 1: Point-in-time estimates of homelessness,” U.S. Dept. Housing Urban Develop., Washington, DC, USA, Tech. Rep., 2022. [Online]. Available: <https://www.huduser.gov/portal/sites/default/files/pdf/2022-AHAR-Part-1.pdf>
- [11] B. Hong, A. Malik, J. Lundquist, I. Bellach, and C. E. Kontokosta, “Applications of machine learning methods to predict readmission and length-of-stay for homeless families: The case of win shelters in New York City,” *J. Technol. Human Services*, vol. 36, no. 1, pp. 89–104, Jan. 2018.
- [12] J. Jarvis, “Individual determinants of homelessness: A descriptive approach,” *J. Housing Econ.*, vol. 30, pp. 23–32, Dec. 2015.
- [13] M. Johnstone, C. Parsell, J. Jetten, G. Dingle, and Z. Walter, “Breaking the cycle of homelessness: Housing stability and social support as predictors of long-term well-being,” *Housing Stud.*, vol. 31, no. 4, pp. 410–426, May 2016.
- [14] A. Kube, S. Das, and P. J. Fowler, “Allocating interventions based on predicted outcomes: A case study on homelessness services,” in *Proc. AAAI Conf. Artif. Intell.*, 2019, vol. 33, no. 1, pp. 622–629.
- [15] R. Kuhn and D. P. Culhane, “Applying cluster analysis to test a typology of homelessness by pattern of shelter utilization: Results from the analysis of administrative data,” *Amer. J. Community Psychol.*, vol. 26, no. 2, pp. 207–232, Apr. 1998.
- [16] B. A. Lee, T. Price-Spratlen, and J. W. Kanan, “Determinants of homelessness in metropolitan areas,” *J. Urban Affairs*, vol. 25, no. 3, pp. 335–356, Aug. 2003.
- [17] D. Liben-Nowell, J. Novak, R. Kumar, P. Raghavan, and A. Tomkins, “Geographic routing in social networks,” *Proc. Nat. Acad. Sci. USA*, vol. 102, no. 33, p. 11623–11628, 2005.
- [18] W. McAllister, M. C. Lennon, and L. Kuang, “Rethinking research on forming typologies of homelessness,” *Amer. J. Public Health*, vol. 101, no. 4, pp. 596–601, Apr. 2011.
- [19] P. H. Rossi and J. D. Wright, “The determinants of homelessness,” *Health Affairs*, vol. 6, no. 1, pp. 19–32, Jan. 1987.
- [20] H. Schütze, C. D. Manning, and P. Raghavan, *Introduction to Information Retrieval*, vol. 39. Cambridge, U.K.: Cambridge Univ. Press, 2008.
- [21] K. H. Shelton, P. J. Taylor, A. Bonner, and M. van den Bree, “Risk factors for homelessness: Evidence from a population-based study,” *Psychiatric Services*, vol. 60, no. 4, pp. 465–472, Apr. 2009.
- [22] M. Shinn, J. Gottlieb, J. L. Wett, A. Bahl, A. Cohen, and D. B. Ellis, “Predictors of homelessness among older adults in New York City: Disability, economic, human and social capital and stressful events,” *J. Health Psychol.*, vol. 12, no. 5, pp. 696–708, Sep. 2007.
- [23] M. Shinn and J. Khadduri, *In the Midst of Plenty: Homelessness and What to do About it*. Hoboken, NJ, USA: Wiley, 2020.
- [24] M. Shinn et al., “Predictors of homelessness among families in New York City: From shelter request to housing stability,” *Amer. J. Public Health*, vol. 88, no. 11, pp. 1651–1657, Nov. 1998.
- [25] H. Toros and D. Flaming, “Prioritizing homeless assistance using predictive algorithms: An evidence-based approach,” *Cityscape*, vol. 20, no. 1, pp. 117–146, 2018.
- [26] J. Tsai and R. A. Rosenheck, “Risk factors for homelessness among U.S. Veterans,” *Epidemiologic Rev.*, vol. 37, no. 1, pp. 177–195, Jan. 2015.
- [27] United States Department of Housing and Urban Development. (2020). *HMIS Data Standards Manual*. Accessed: May 26, 2021. [Online]. Available: <https://www.hudexchange.info/resource/3824/hmis-data-dictionary/>
- [28] B. Van Straaten, J. Van der Laan, G. Rodenburg, S. N. Boersma, J. R. L. M. Wolf, and D. Van de Mheen, “Dutch homeless people 2.5 years after shelter admission: What are predictors of housing stability and housing satisfaction?” *Health Social Care Community*, vol. 25, no. 2, pp. 710–722, Mar. 2017.
- [29] Y.-L.-I. Wong, D. P. Culhane, and R. Kuhn, “Predictors of exit and reentry among family shelter users in New York City,” *Social Service Rev.*, vol. 71, no. 3, pp. 441–462, Sep. 1997.



Charalampos Chelmis (Member, IEEE) received the B.S. degree in computer engineering and informatics from the University of Patras, Patras, Greece, in 2007, and the M.Sc. and Ph.D. degrees in computer science from the University of Southern California at Los Angeles, Los Angeles, CA, USA, in 2010 and 2013, respectively.

He is currently an Assistant Professor of computer science with the University at Albany, State University of New York, Albany, NY, USA, and the Director of the Intelligent Big Data Analytics,

Applications, and Systems Laboratory, Albany. His research interests include high-dimensional and/or interrelated data and social good applications.

Dr. Chelmis has served/serving as the Co-Chair and a TPC Member in international conferences including AAAI and TheWebConf. He is also an Associate Editor of the *Social Network Analysis and Mining* journal.



Khandker Sadia Rahman received the B.Sc. degree in electrical and electronic engineering from the Bangladesh University of Engineering and Technology, Dhaka, Bangladesh, in 2017. She is currently pursuing the Ph.D. degree in computer science with the University at Albany, State University of New York, Albany, NY, USA.

She was an AI Engineer at Hiperdyne Corporation, Tokyo, Japan. She works as a Project Research Assistant under the supervision of Dr. Charalampos Chelmis. She is also working explicitly with providing AI solutions for homelessness. Her research interests include network science and using AI for social good.