Learning to Predict Transitions within the Homelessness System from Network Trajectories

Khandker Sadia Rahman
Department of Computer Science
University at Albany, SUNY
Albany, NY, USA
srahman2@albany.edu

Charalampos Chelmis†
Department of Computer Science
University at Albany, SUNY
Albany, NY, USA
cchelmis@albany.edu

***Abstract*—This study infers the unobserved underlying network of homeless services from administrative data collected by homeless service providers. Both the structure of the inferred network, and historical observations, are used to identify individuals with similar trajectories so that their next assignments can be predicted. Experimental evaluation shows that the proposed approach performs well not only on predicting exit from the system, or simply guessing high frequency services (as most baselines), but is also successful in less frequent scenarios.**

***Index Terms*—Complex systems, network inference, similarity**

I. Introduction

Homelessness poses a long–standing problem to the society. The number of individuals in the U.S. alone experiencing prolonged homelessness for at least 12 months, or repeated homelessness over a period of three years (i.e., chronic homeless [5]) increased by 20% [6] between 2020 and 2021. Numerous methods have been proposed to predict reentry (e.g., [8]), as well as the risk of an individual to become chronically homeless (e.g., [11]). Unlike such methods, this work focuses on individuals experiencing chronic homelessness, broadly defined here as individuals entering the homeless system two or more times. Viewing the history of each individual as a sequence of services and time of stay within each service, the goal is to learn a model that can be used to accurately estimate the next service an individual will be assigned to within the homeless system in the future. Our proposed approach represents the homeless system as a network of interconnected services which individuals traverse over time. Given the history of an individual, it identifies individuals with similar histories, and predicts the next service the given individual will be assigned to based on such histories and the underlying network structure. Our key contributions can be summarized as follows:

- We infer the network of homeless services from administrative data collected by homeless service providers.
- We define a similarity score between ordered sequences of services that the homeless are assigned to as they traverse the network of services.
- We propose a method that, given the history of an individual, can predict the most likely service she will be assigned to next.
- For reproducibility, we make our source code available at https://github.com/IDIASLab/TRACE.

This material is based upon work supported by the National Science Foundation under Grant No. ECCS–1737443.

†Both authors contributed equally.

IEEE/ACM ASONAM 2022, November 10-13, 2022
http://dx.doi.org/10.1145/XXXXXXX.XXXXXXX
978-1-6654-5661-6/22/$31.00 © 2022 IEEE

II. Related work

Unlike prior art for reentry and chronic homelessness prediction (e.g., [7], [9], [11]), this work addresses the more challenging problem of determining whether an individual will exit the homeless system, and otherwise, the exact program she will be assigned to next. The proposed similarity score differs from existing similarity measures for time–series data [12], in that our study involves trails of timestamped categorical data as opposed to numerical data such measures are applicable to. Compared to prior art on sequence analysis, which is often used to determine a common subsequence between two categorical sequences (e.g., [4]) and assumes i.i.d. data points in a sequence, the approach described here utilizes the network that generates the observed sequences to compute similarity between two categorical trajectories. Finally, methods for network inference from data [2] either assume that diffusion traces are directed acyclic graphs and that the probability of transitioning from one node to another is fixed and same for all edges, or infer interactions for pairs of nodes expected to be directly connected via an edge. Instead, in this study, the temporal chain of events is observed (leading to an easier inference task) while at the same time, observed trajectories may contain cycles (resulting in a harder inference problem).

III. Problem Statement

Homeless service providers offer services that are organized in *project types* (e.g., emergency shelters, transitional housing) [10]. We denote the set of project types as $P = \{p_1, p_2, \ldots, p_n\}$, and the set of individuals requesting services multiple times as $C = \{c_1, c_2, \ldots, c_m\}$. Reentries can be viewed as temporally ordered sequences of tuples $(p_i, t_i = [s_i, e_i])$, where $p_i \in P$, and s_i

and e_i are the times at which individual $c \in C$ enters to, and exits from p_i, accordingly. Such a *trajectory*, $T_c = (p_1, [s_1, e_1]), (p_2, [s_2, e_2]), \dots, (p_N, [s_N, e_N]))$ for each individual $c \in C$, where for each two consecutive tuples $s_{i+1} \geq e_i$, records her transitions from p_i to p_{i+1}. Given a set of trajectories $\mathcal{T}$, and query trajectory $T_q = ((q_1, [s_1, e_1]), \dots, (q_N, [s_N, e_N])$ of an individual $q \in C$, up to time e_N, we aim at predicting the project type $\hat{q}_{N+1} \in P$ that she will be assigned to immediately after q_N (or exit if she is likely to exit the system).

IV. Inferring the graph of homeless services

The underlying connectivity of homeless services (i.e., potential paths an individual can take once she is admitted into the homeless system) is neither directly observable nor known. We thus set forth to uncover the aggregate dynamics of the homeless system from the observed sequences of services that it generates for each individual.

We begin by modelling the network as a directed graph $\mathcal{G} = (P, E)$, where P is the set of nodes representing services visited by individuals in C, and E is the set of edges between nodes, such that a directed edge appears between p_i and p_j if at least one trajectory in $\mathcal{T}$ exists, in which p_j appears after p_i. We determine the weight of each edge $(p_i, p_j) \in E$ based on the *number of steps* taken before reaching p_j from p_i, and the position where p_i appears in the trajectory (i.e., *offset*). Specifically, a path from p_i to p_j in a trajectory T involves $j - i + 1$ steps starting from offset i. Therefore:

$$w_{ij} = \sum_{k=1}^{N} \sum_{l=0}^{M-1} \alpha^{k-1+l} f_{ij,kl}, \tag{1}$$

where N is the maximum number of steps, and M the largest offset, $\alpha \in [0, 1]$ is some attenuation factor, and $f_{ij,kl}$ denotes the number of times a transition from p_i to p_j appears in any path across all trajectories in $\mathcal{T}$ with k steps at offset l. Finally, we normalize the weights of outgoing edges at each node to sum to 1. Edge weights satisfy the following properties: $w_{ij} > 0, \forall p_i, p_j \in P$, w_{ij} is undefined if no path $\exists$ from p_i to p_j, $w_{ij} \approx 0$ if path from p_i to p_j is long, and $\sum_{j \in V} w_{ij} = 1, \forall p_i \in P$.

In summary, inference proceeds as follows. First, for each trajectory, all possible unique paths are extracted. The value of $f_{ij,kl}$ is then computed by counting the frequency of each path across all trajectories. To avoid double counting, only unique paths at each number of steps are included in the computation of $f_{ij,kl}$. Next, edge weights are calculated using Eq.(1), number of steps, and offset of each path appearing in the trajectories.

V. Trajectory Similarity

To compare a given individual's trajectory with historical trajectories, we wish to measure similarity σ, between a query trajectory T_q, and a historical trajectory T_c given graph $\mathcal{G}$, while taking into account the distance between nodes in $\mathcal{G}$ appearing in T_q and T_c respectively, the temporal overlap between the trajectories, and the time intervals individuals stayed on each node in the corresponding trajectories.

Let the distance between two nodes $q_i, p_j \in P$ that appear in T_q and T respectively, be the weighted shortest path distance, $d(q_i, p_j)$, between them in $\mathcal{G}$. The distance between $q_i \in T_q$ and T within time interval t_q can then be computed as:

$$d(q_i, T, t_q) = \frac{\min_{(p_j, t_j) \in T[t_q]} \max_{|t_j \cap t_q|} d(q_i, p_j)}{D_{\mathcal{G}}}, \tag{2}$$

where $D_{\mathcal{G}}$ is the maximum weighted distance between any two nodes in $\mathcal{G}$, and $T[t_q]$ denotes the sequence of nodes in T visited during t_q. Intuitively, this distance is minimized for $(p_j, t_j) \in T[t_q]$ that can be reached quickly from q_i in $\mathcal{G}$, and for which the interval t_j significantly overlaps with t_q. By definition, $0 \leq d(q_i, T, t_q) \leq 1$. The similarity between trajectories T_q and T for time interval t, is therefore:

$$\sigma(T_q, T, t) = \frac{\sum_{(q, t_q) \in T_q[t]} |t_q \cap t_j| \times e^{-d(q, T, t_q)}}{\sum |t_q|}, \tag{3}$$

where $|.|$ denotes the length of a time interval, and t_j is the time interval in T corresponding to the node that minimizes $d(q, T, t_q)$. By definition, $0 \leq \sigma(T_q, T, t) \leq 1$, with $\sigma(T_q, T_q, t) = 1$ for any time interval t, and $\sigma(T_q, T, t) = 0$ for any two trajectories T_q and T with no temporal overlap.

VI. Trajectory Similarity Estimation and Probabilistic Prediction

In this section, we describe *TRACE*, a novel approach for Trajectory SimilaRity EstimAtion and ProbabilistiC PrEdiction. Specifically, given the history $T_q = ((q_1, [s_1, e_1]), \dots, (q_N, [s_N, e_N])$ of an individual, and time interval $t = [s_{N-\gamma}, e_N]$, where γ is the number of prior services received in the past, as well as the set $\mathcal{T}$ of trajectories of other individuals, TRACE begins by calculating the effective length λ of T_q. Then, TRACE identifies the most similar trajectory $T_c \in \mathcal{T}$ to T_q within t, using graphs G_λ and $G_{\lambda+1}$. The rationale for this design choice is that the next node may either be a node already visited in the past (in which case the effective length of T_q will remain unchanged) or a new node (in which case the effective length of T_q will increase by 1). The project type $\hat{q}_{N+1}$ that q is expected to be assigned to next is therefore estimated to be the one that maximizes the transition probability from p_c, the last matching project type in trajectory T_c that maximizes $\sigma(T_q, T_c, t)$ over either G_λ or $G_{\lambda+1}$. Therefore, $\hat{q}_{N+1}$ is obtained by maximizing the following objective:

$$\mathbb{1}_{(q_N, p_i) \in E} \times \mathbb{1}_{(p_c, p_i) \in E} \times P(p_i \in \text{NB}(q_N) \cap \text{NB}(p_c) | p_c), \tag{4}$$

where $\mathbb{1}$ is the indicator function. To ensure the predicted node is reachable from q_N, p_i is constrained to be in the out–neighborhood of both p_c and q_N. If no such node can be found within T_c, the search over trajectories continues until a trajectory is identified that satisfies this constraint, or if no further trajectories are left to be examined. Furthermore, only those trajectories temporally overlapping with t are considered.

Algorithm 1 TRACE

Input: $\mathcal{H}, T_q, \gamma$

Output: Node $\hat{q}_{N+1}$ individual c is going to visit after q_N

Identify subset $\mathcal{H}_t \subseteq \mathcal{H}$ of trajectories temporally overlapping $t = (s_{N-\gamma}, e_N)$

Compute effective length, λ, of query trajectory T_q

for each $T_c \in \mathcal{H}_t$ **do**

  $\forall q_i \in T_q[t]$ compute distance to $T_c[t]$ using Eq.(2) on $\mathcal{G}_\lambda$ (similarly for $\mathcal{G}_{\lambda+1}$)

  Compute similarity between $T_q[t]$ and T_c using Eq.(3)

  $T_{max} \leftarrow T_c$ with highest similarity score

end for

Using T_{max}, find $\hat{q}_{N+1}$ that maximizes Eq. (4)

return $\hat{q}_{N+1}$

VII. Experimental Evaluation

A. Data

Experiments are performed using an anonymized set of records for $6,011$ individuals that received homeless services multiple times by providers in the Capital Region of the state of New York, between 2012 and 2018. The data was provided by CARES of NY. A subset of individuals $(3,475)$ exited to stable (i.e., rental or owned housing without any housing subsidy) or fairly stable (these include permanent housing, rental or owned housing with housing subsidy, and long term facility) exits [3].

We split the data into $\mathcal{T}_{train}$ and $\mathcal{T}_{test}$ as follows. Trajectories with entry date up to the end of 2016 are included in training set, $\mathcal{T}_{train}$, whereas trajectories with entry dates from the beginning of 2017 onward are included in test set, $\mathcal{T}_{test}$. This split ensures that predictive models are trained on data that does not contain future information, and therefore avoids data leakage. The test set $\mathcal{T}_{test}$ is further split into two disjoint sets, namely historical set, $\mathcal{H}$, and query set, Q. For each query in Q, the most similar trajectory in $\mathcal{H}$ is obtained using Eq.(3). Graphs $\mathcal{G}_\lambda$ and $\mathcal{G}_{\lambda+1}$ are computed over $\mathcal{T}_{train}$.

B. Metrics

- **Precision@k**: measures how many times the predicted service is correct using the top k predictions.
- **Recall@k**: measures how many times each project type is identified correctly at k^{th} prediction.

We differentiate our analysis between individuals that exit the system (i.e., exit point) and those that transition to a new service (i.e, interim point) for a granular assessment of the predicting capabilities of TRACE and the baselines.

C. Baselines

We consider two variants of TRACE, namely TRACE_1, that uses only G_λ, and TRACE_2 that uses both G_λ and $G_{\lambda+1}$. We also compare TRACE with the baselines described below. Given the most similar trajectory T_c to query T_q, $\hat{q}_{N+1}$ is predicted to be:

- $\mathbf{p_c - Next}$: p_{c+1} (i.e., the service following p_c in T_c).
- $\mathbf{p_c - NB}$: the highest transition probability out–neighbor of $p_c \in P$. Note that we found no difference in performance when requiring the out–neighbor of p_c to also be an out–neighbor of q_N ($\mathbf{p_c q_N - NB}$).
- $\mathbf{q_N - NB}$: the out–neighbor of $q_N \in V$ with the highest transition probability from q_N.
- **RN**: a random node $p \in P$. We consider two variants, namely selecting a node (**UR**) uniformly at random, and (**PA**) with a probability that is proportional to a node's in–degree (i.e., preferential attachment [1]).
- **Sim-attenuate**: the node identified using Eqs.(3) and (4), with the difference that an attenuation factor β^k, where $0 \leq k \leq K$ is the number of nodes in $T[t]$, is used in Eq.(3) to penalize intervals which are further in the past.

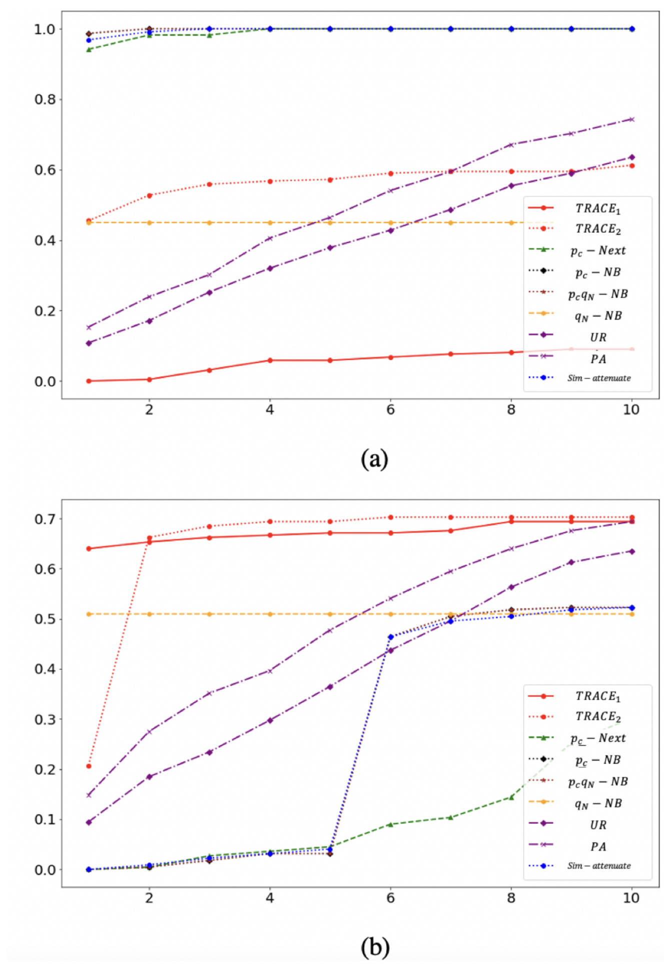

Fig. 1. Precision@k plot for TRACE and baselines predicting (a) exit points and (b) interim points (Precision at x–axis and k at y–axis).

VIII. Results and Analysis

Fig. 1 shows that while most methods perform well at predicting exit points, only $\mathbf{TRACE_1}$ and $\mathbf{TRACE_2}$ excel at predicting interim points. In fact, $\mathbf{q_N - NB}$ consistently predicts the most frequent project type as the next transition and is therefore meaningless. Similarly, both $\mathbf{p_c - NB}$ and **Sim-attenuate** predict project type 1 with high probability

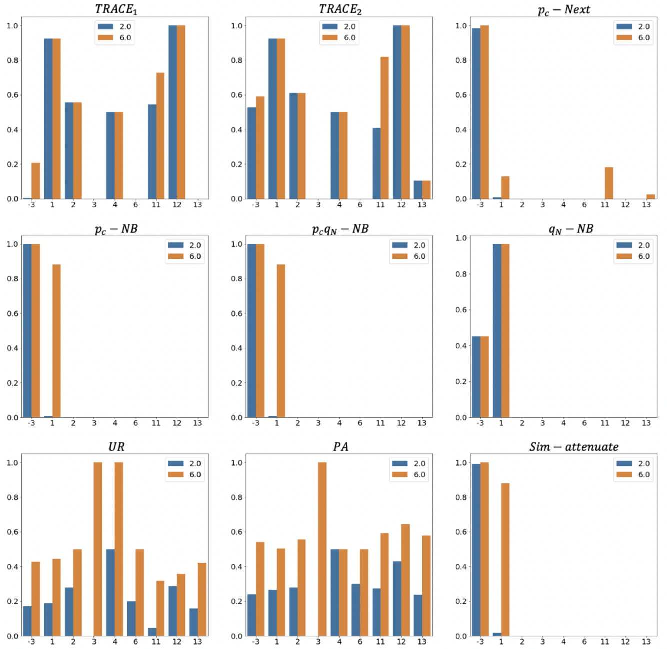

Fig. 2. Recall (y–axis) achieved by TRACE and the baselines at $k = 2$ (blue) and $k = 6$ (orange) for each project type (x–axis) separately.

(80%) resulting in good performance with respect to exit prediction, but meaningless prediction of interim points. Interestingly, the performance of $\mathbf{p_c - NB}$ and **Sim-attenuate** improves dramatically for $k \geq 5$. However, Fig. 2 suggests that this is an artifact of their high recall for project type 1, which comprises 50% of the ground truth. $\mathbf{p_c - Next}$ focuses on high frequency project types (i.e., 1, 11, 13). Finally, $\mathbf{TRACE_1}$ performs poorly for exit points mainly due to its low recall. Instead, $\mathbf{TRACE_2}$ seems to predict exit at first (i.e., $k = 1$), explaining its reduced performance for interim points. However, $\mathbf{TRACE_2}$'s recall improves dramatically as k increases, as shown in Fig. 2. This result suggests that a tiered system to predict whether an individual is more likely to exit or not, followed by a prediction of the next step using $\mathbf{TRACE_2}$ is likely to further improve prediction accuracy.

Next, we experiment with different values of α to determine its effect on TRACE's recall, as shown in Fig. 3. When $\alpha = 0.5$, the TRACE$_1$'s exit performance improves dramatically, whereas $\alpha = 0.8$ only slightly improves performance for interim points. On the other hand, $\alpha = 0.2$ achieves the best results for TRACE$_2$ for both exit and interim points, but the improvement over $\alpha = 0.5$ is not significant. We therefore report results for $\alpha = 0.5$ for both TRACE$_1$ and TRACE$_2$ when comparing their performance against the baselines.

Finally, we explore the effect of hyperparameter γ, and find that its value becomes critical for lengthy trajectories, which require more past information to improve the chances of accurately predicting the next interim point. To demonstrate this, we focus on queries of length greater than 4. Fig. 4(b) implies that the higher the value of γ, the higher the precision. In our comparisons, our results are therefore for $\gamma = 6$.

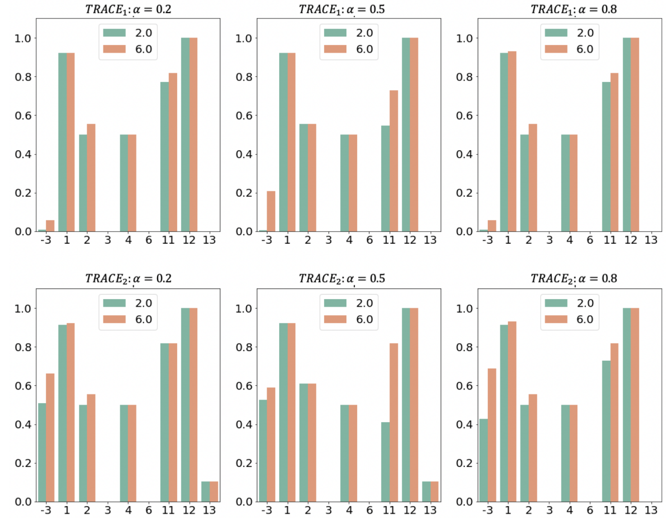

Fig. 3. Recall@k (y−axis) of TRACE with varying α.

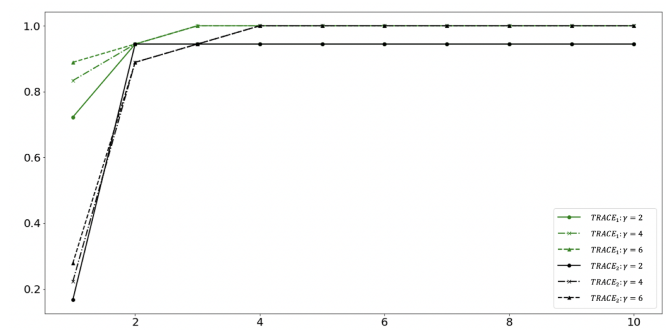

Fig. 4. Precision@k for interim points of $TRACE_1$ and $TRACE_2$ with varying $\gamma = \{2, 4, 6\}$.

IX. Conclusion

In this study, we proposed an approach that begins by inferring the network of services that the homeless visit as they strive to secure permanent housing. We subsequently defined a score to assess the similarity of their trajectories. Based on these two contributions we proposed a method to predict the most likely service an individual will be assigned to next given her history. Our experimental evaluation showed the ability of the proposed approach to better match the observed sequences as opposed to baselines. Our approach can used as a building block for more complex applications, such as recommending service assignment. However, in replicating the observed data, biases in the service assignment process may be replicated. In order to avoid this, and to additionally evaluate the potential ability of "wrong" predictions to lead to better outcomes, we will explore counterfactual predictions in future work. We additionally plan to address limitations, such as accounting for imbalances among project types.

References

[1] R. Albert and A.-L. Barabási, "Statistical mechanics of complex networks," *Reviews of modern physics*, vol. 74, no. 1, p. 47, 2002.

[2] I. Brugere, B. Gallagher, and T. Y. Berger-Wolf, "Network structure inference, a survey: Motivations, methods, and applications," *ACM Computing Surveys (CSUR)*, vol. 51, no. 2, pp. 1–39, 2018.

[3] C. Chelmis and K. S. Rahman, "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*, 2021, pp. 69–73.

[4] C. H. Elzinga, "Sequence analysis: Metric representations of categorical time series," *Sociological methods and research*, 2006.

[5] 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, p. e0245088, 2021.

[6] M. Henry, T. de Sousa, C. Roddey, S. Gayen, T. Joe Bednar, and A. Associates, "The 2021 Annual Homeless Assessment Report (AHAR) to Congress. Part 1: Point-in-time estimates of homelessness," *The US Department of Housing and Urban Development*, 2021.

[7] 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," *Journal of Technology in Human Services*, vol. 36, no. 1, pp. 89–104, 2018.

[8] A. Kube, S. Das, and P. J. Fowler, "Allocating interventions based on predicted outcomes: A case study on homelessness services," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, no. 01, 2019, pp. 622–629.

[9] G. Messier, C. John, and A. Malik, "Predicting chronic homelessness: The importance of comparing algorithms using client histories," *Journal of Technology in Human Services*, pp. 1–12, 2021.

[10] United States Department of Housing and Urban Development, "HMIS Data Standards Manual," *Retrieved May 26, 2021, from https://www.hudexchange.info/resource/3824/hmis-data-dictionary/*, 2020.

[11] B. VanBerlo, M. A. Ross, J. Rivard, and R. Booker, "Interpretable machine learning approaches to prediction of chronic homelessness," *Engineering Applications of Artificial Intelligence*, vol. 102, p. 104243, 2021.

[12] H. Wang, H. Su, K. Zheng, S. Sadiq, and X. Zhou, "An effectiveness study on trajectory similarity measures," in *Proceedings of the Twenty-Fourth Australasian Database Conference-Volume 137*, 2013, pp. 13–22.