Machine Learning (2026) 115:19
https://doi.org/10.1007/s10994-025-06904-z

The urban impact of AI: modelling feedback loops in location-based recommender systems

Giovanni Mauro[1,2] · Marco Minici[3] · Luca Pappalardo[1,2]

Received: 10 April 2025 / Revised: 13 August 2025 / Accepted: 23 September 2025

Abstract
Location-based recommender systems are increasingly embedded in location-based services, shaping individual mobility decisions in urban environments. While their predictive accuracy has been extensively studied, less attention has been paid to their systemic impact on urban dynamics. In this work, we introduce a simulation framework to model the human–AI feedback loop underpinning next-venue recommendation, capturing how algorithmic suggestions influence individual behaviour, which in turn reshapes the data used to retrain the models. Our simulations, grounded in real-world mobility data, systematically explore the effects of algorithmic adoption across a range of recommendation strategies. We find that while recommender systems consistently increase individual-level diversity in visited venues, they may simultaneously amplify collective inequality by concentrating visits on a limited subset of popular places. This divergence extends to the structure of social co-location networks, revealing broader implications for urban accessibility and spatial segregation. Our framework operationalizes the feedback loop in next-venue recommendation and offers a novel lens through which to assess the societal impact of AI-assisted mobility, providing a computational tool to anticipate future risks, evaluate regulatory interventions, and inform the design of ethic algorithmic systems.

Keywords Human-AI coevolution · Human-AI feedback loop · Next-venue recommendation · Human mobility · Diversity · Location-based recommender systems

Editors: Riccardo Guidotti, Anna Monreale, Dino Pedreschi.

✉ Giovanni Mauro
giovanni.mauro@sns.it

✉ Luca Pappalardo
luca.pappalardo@isti.cnr.it

Marco Minici
marco.minici@icar.cnr.it

[1] ISTI-CNR, Via G. Moruzzi 1, 56124 Pisa, Italy

[2] Scuola Normale Superiore, Piazza dei Cavalieri,7, 56126 Pisa, Italy

[3] ICAR-CNR, Via Pietro Bucci 8/9c, 87036 Rende, Italy

Published online: 08 January 2026

1 Introduction

Recommender systems have become increasingly pervasive in urban life, seamlessly integrated into the online platforms that millions of individuals rely on to make daily decisions such as navigating cities and searching for accommodation (Pedreschi et al., 2025; Quijano-Sánchez et al., 2020). Among the most influential are location-based recommenders (a.k.a. points-of-interest recommenders), which suggest places to visit—such as restaurants, cafés, parks, shops, or cultural landmarks—through widely used location-based services like Google Maps, Yelp, and Dianping. These platforms act as decision-making assistants, offering personalised suggestions in real time based on an individual's current location, mobility history, preferences, and contextual factors such as time of day or weather (Pappalardo et al., 2023).

The growing importance of location-based recommenders stems from both their ubiquity and impact. As mobile devices and GPS-enabled apps become increasingly pervasive tools in our daily lives, we increasingly delegate micro-decisions (where to eat, shop, or spend leisure time) to algorithmic systems. These recommendations are powered by machine learning techniques and deep learning architectures, which can model complex spatio-temporal patterns and predict human whereabouts with high accuracy (Luca et al., 2021). By shaping when and where people move in the city, location-based recommenders exert a subtle yet powerful influence on human movements, urban flows, and the popularity of venues. For example, a recent study shows that the recommender system implemented by Uber Eats led to significant improvements in consumer engagement and gross bookings, influencing the popularity of certain restaurants (Wang et al., 2025).[1]

Although much research has focused on developing accurate next-location prediction algorithms (Luca et al., 2021), and recent work has begun to question their ability to generalise to unknown situations (Luca et al., 2023), their broader implications for urban dynamics remain largely unexplored (Pedreschi et al., 2025; Pappalardo et al., 2023, 2024). This gap in understanding is particularly pressing in light of recent regulatory developments. The European Union's Digital Services Act introduces mandatory risk assessment requirements for very large online platforms and search engines, with a focus on assessing their impact on users and society.[2] Among these platforms is Google Maps, one of the most influential systems that recommend physical venues, which is now officially subject to this regulatory scrutiny.

Understanding the urban impact of location-based recommenders is challenging because their interaction with individuals creates an intricate feedback loop (Pedreschi et al., 2025): each individual decision contributes to the formation of the big data that is used to train the recommender system, which in turn influences future decisions about where to go. This recursive process generates a continuously evolving cycle that can reshape the urban landscape. Capturing this feedback loop in its entirety—encompassing algorithmic recommendations, user responses, the parameters of machine learning models, and the retraining frequency—is difficult, as it would require direct access to the online platforms delivering the recommender systems. At present, only the companies operating these platforms have the technical and legal means to observe and analyse the full feedback loop. The Delegated

[1] See also https://www.gsb.stanford.edu/insights/better-way-make-recommendations-power-popular-platforms.

[2] https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32022R2065

Act detailing how vetted researchers can gain access to platform data under Article 40 of the digital services act was only published in July 2025, and to date no scientific studies conducted by researchers outside the platforms have been able to experimentally analyse their impact. Therefore, it is not surprising that most research so far has focused on mathematical formalisation of the feedback loop, rather than its empirical measurement (Pappalardo et al., 2024; Chen et al., 2023). While some progress has been made in this direction for other human-AI ecosystems—such as social media, online retail, and generative AI (Lesota et al., 2024; Zhou et al., 2024; Shumailov et al., 2024; Mansoury et al., 2020; Jiang et al., 2019; Sun et al., 2019; Nguyen et al., 2014; Piao et al., 2023; Coppolillo et al., 2025)—no equivalent effort has yet been undertaken for feedback loops that involve location-based recommender systems despite their growing influence on urban dynamics.

In this article, we propose a computational framework that models the feedback loop between individuals and location-based recommender systems, with the goal of simulating its impact on individual and collective patterns of visits to public venues in a city. Our simulation framework integrates multiple types of recommender systems—from neighbourhood-based to deep learning methods—trained on real-world data describing venue visits in New York City and Tokyo. At the core of the framework is the venue selection mechanism, which governs how individuals decide where to go next: with a given probability, they follow a recommendation; otherwise, they make an autonomous choice, simulated using a mechanistic model of human mobility. We run the simulation over several weeks, allowing the feedback loop between individuals and the recommender system to unfold over time. As individuals interact with the system—receiving recommendations, making choices, and visiting venues—these interactions continually update the data used to retrain the recommender system, which in turn influences future recommendations. This dynamic process enables us to observe how different levels of algorithmic influence shape urban mobility patterns. Specifically, we study: (i) the diversity of visits at the individual level, i.e. how varied each individual's venue choices are; (ii) the variation in venue popularity at the collective level, i.e., how visits are distributed across all venues; and (iii) the structure of the individuals' co-location network, i.e., how often individuals visit the same venues within a given time window.

Our results reveal that the impact of recommender systems varies across the various recommendation methods investigated, yet a consistent pattern emerges: at the individual level, we find an increased diversification in venue visits, while at the collective level, diversity decreases reflecting a rise in inequality in venue popularity. We further examine the tension between personalization and concentration, and find that the increase in individual diversity is largely driven by visits to already popular venues, ultimately making individuals' visitation patterns more similar to one another. This mirrors patterns observed in other human–AI ecosystems, where recommender systems improve individual choice but simultaneously amplify system-wide popularity bias. Our study thus adds another piece to the broader mosaic of how recommender systems shape complex social systems, highlighting the structural consequences of algorithmic mediation across diverse domains.

The remainder of the paper is organised as follows. Section 2 reviews the literature on the impact of recommender systems on urban dynamics. Section 3 presents our modelling of the feedback loop and the simulation framework. Section 4 outlines the experimental setup and Sect. 5 reports the results of our simulations. Section 6 interprets our findings in the broader

context of human–AI coevolution and human mobility. Finally, Sect. 7 concludes the paper by discussing the limitations of our study and proposing directions for future research.

2 Related work

This section reviews key experimental studies that examine how recommendation systems impact urban dynamics.

A growing body of research investigates the impact of GPS-based navigation systems on the urban environment. Arora et al. (2021) simulate the effects of Google Maps navigation in Salt Lake City, reporting average reductions of 6.5% in travel time and 1.7% in CO_2 emissions. These benefits are even more pronounced for individuals whose routes were altered by the route recommendations. Cornacchia et al. (2022) use TomTom APIs to simulate navigation services in Milan, showing that adoption rates critically shape system-wide outcomes: emissions increase at very low or very high adoption levels but decrease when adoption stabilises around 50%. Perez-Prada et al. (2017) show that the widespread adoption of eco-routing during congestion can reduce CO_2 and NO_x emissions by 10% and 13%, respectively. However, these gains come with trade-offs: NO_x exposure increases by 20%, travel times rise by 28%, and vehicle concentration in downtown areas grows by 16%. Two studies simulate routing systems in Florence, Rome, and Milan to show that prioritising diversity in routing can lead to more balanced traffic distribution, improved road network coverage, and reductions in CO_2 emissions (Cornacchia et al., 2024, 2023).

Several studies examine how recommendation systems on house-renting platforms influence socio-economic dynamics, often reinforcing systemic inequalities. Edelman and Luca (2014) uncover racial disparities in New York City's Airbnb market, reporting a 12% revenue gap between Black and non-Black hosts, even after controlling for property characteristics and guest ratings. Zhang et al. (2021) expand on this finding, showing that the AirBnB smart pricing algorithm can reduce these price disparities and improve revenues for black hosts, although it does not completely eliminate systemic bias.

Additional research investigates how ride-hailing and ride-sharing platforms influence urban mobility through algorithmic matching and incentive design. Bokányi and Hannák (2020) find that a ride recommender system prioritising lower-earning drivers in New York City would reduce disparities and increase driver earnings compared to prioritizing the closest available vehicle. Afèche et al. (2023) analyse passenger-driver matching and centralised dispatch algorithms, showing that while centralized control improves efficiency, it can restrict service access in less densely populated areas, exacerbating spatial inequities.

2.1 Position of our work

A review of the literature reveals two major gaps (Pappalardo et al., 2024). First, there are no existing studies that have examined the actual or potential impact of location-based recommendation systems on urban dynamics. Second, although a growing body of research outside the urban domain has begun to formalise and model human–AI feedback loops (Sun et al., 2019; Mansoury et al., 2020; Jiang et al., 2019; Nguyen et al., 2014; Shumailov et al., 2024; Lesota et al., 2024; Zhou et al., 2024; Coppolillo et al., 2025), little attention has

been paid to how these feedback loops manifest in urban environments or how they can be formally modelled.

A notable exception is the work by Ensign et al. (2018), which shows how a feedback loop can distort the allocation of police resources, reinforce existing biases, and generate unrealistic crime distributions in a city. By introducing corrective mechanisms to prevent repeated deployments in the same areas, the authors demonstrate that even minor changes in algorithmic design can significantly influence urban outcomes. While this study represents a pioneering effort, the modelling of human–AI feedback loops and their effects on urban mobility and spatial behaviour remain largely unexplored.

Our work takes a first step toward addressing these gaps by explicitly modelling the human–AI feedback loop in the context of location-based recommendation systems. We introduce an open-source simulation framework that enables the systematic evaluation of how different recommendation algorithms influence key aspects of urban dynamics, such as venue popularity and individual visitation patterns.

3 Methodology

We design a simulation framework to model the human-AI feedback loop in the context of next-venue recommendation. In Sect. 3.1, we introduce the notation and definitions necessary to understand the simulation framework. Section 3.2 presents our modelling of the feedback loop and the simulation methodology. Section 3.3 describes the metrics used to evaluate the impact of the feedback loop on individual and collective venue visitation patterns.

3.1 Preliminaries

Let U be a set of individuals and V a set of public venues. We define a dataset $D \subseteq U \times V \times T$ as an interaction matrix, where each tuple $(u, v, t) \in D$ represents a visit by individual $u \in U$ to venue $v \in V$ at time $t \in T$. The dataset describes all visits up to a time $T_{\max}$, and tuples in it are sorted chronologically. For each individual $u \in U$, we define $D_u = \{v_1, \ldots, v_n\}$ as the time-ordered sequence of venues visited by u in D, sorted chronologically. The dataset D is partitioned into two disjoint subsets: a training set D_{train}, which contains all tuples $(u, v, t) \in D$ such that $t \leq T_{\text{train}}$, and a post-training set D_{post}, which contains all tuples such that $T_{\text{train}} < t \leq T_{\max}$.

Each venue $v \in V$ is associated with a categorical label provided by a function $\ell : V \to C$, where $C = \{c_1, \ldots, c_z\}$ is a predefined set of venue categories (e.g., *restaurant*, *museum*, *school*). Venues are geolocated via a function $\pi : V \to [-90, 90] \times [-180, 180]$, which maps a venue $v \in V$ to its latitude and longitude coordinates. The geographical distance between two venues $v_i, v_j \in V$ is computed as the great-circle distance between their coordinates on the Earth's surface.

We define a scoring function $\mathcal{R}_\theta : U \times V \times T \to [0, 1]$ with parameters θ and trained on D_{train}. For any individual $u \in U$, venue $v \in V$, and time $t \in T$, the score $\mathcal{R}_\theta(u, v, t)$ represents the estimated probability that u will visit v as next location at a future time $t' > t$. The recommender system, $\mathcal{A}$, is an operator that, for a given user u and time t, ranks all

candidate venues according to their scores $\mathcal{R}_\theta(u, v, t)$ and returns the top-k venues from this ranking.

Definition 3.1 (*Next-venue recommendation problem*) Given a set of individuals U, a set of venues V, and a time t, the next-venue recommendation problem consists in predicting the venue $v \in V$ that an individual $u \in U$ is most likely to visit next. The goal of a next-venue recommender system $\mathcal{A}$ is to learn a function $R_\theta(u, v, t)$ that estimates the likelihood of future interactions and to return the top-k venues with highest scores.

3.2 Modeling and simulating the feedback loop

Given the sets U, V, the datasets D, D_{train} and D_{post}, the scoring function $\mathcal{R}_\theta$ and the recommender system $\mathcal{A}$, we model the feedback loop to simulate the generation of visits to venues in V by individuals in U.

We assume a scenario where individuals have a general intent or agenda (e.g., visiting a restaurant or bar) and then choose a specific venue based on either personal knowledge or the algorithmic suggestions provided by recommender system $\mathcal{A}$. As the simulation of the feedback loop goes by, a dataset $\mathcal{S}$ is generated, which contains the simulated visits of the individuals to the venues. Figure 1 provides an overview of the framework that models the feedback loop, illustrating the key components and the flow of interactions between individuals and the recommender system.

The simulation for each individual $u \in U$ is outlined in Algorithm 1. The process begins with training the recommender system $\mathcal{A}$ on the dataset D_{train} (line 1). The simulation then

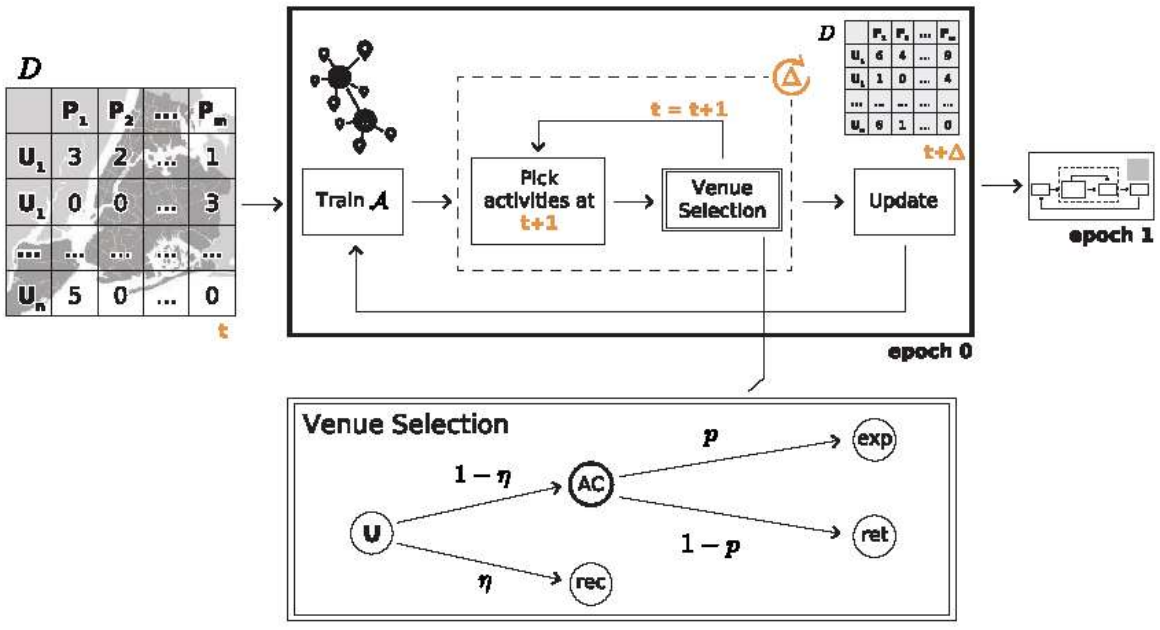


Fig. 1 Overview of the simulation framework modelling the feedback loop. At each step of the simulation, an individual chooses between relying on recommendations (with probability η) or an autonomous choice (AC) (with probability $1 - \eta$). AC is split between exploration (with probability p) and preferential return (with probability $1 - p$), where p is a fixed parameter equals to $\rho \cdot S^{-\gamma}$ where S is the number of distinct previously visited locations. Initially, a recommender system $\mathcal{A}$ is trained on the history of venue visits. At each timestep $t + 1$, individuals' next venues are selected, and the dataset is updated accordingly. After Δ timesteps, the recommender is retrained, allowing the feedback loop between individual decisions and algorithmic suggestions to evolve over time

iterates over each visit tuple $(u, v, t) \in D_{\text{post}}$ (line 5), simulating user mobility in three sequential steps: category selection (line 6), where the category of the next venue is taken from D_{post}; venue selection (line 7), where the specific venue is selected based on the chosen category and the individual's decision to use or not the recommender system; and algorithm retraining (line 10), which updates $\mathcal{A}$ every Δ time units to incorporate newly simulated data in $\mathcal{S}$.

Input Data: Dataset of visits D, Training dataset D_{train}, Observation dataset D_{post}. All datasets are sorted by time.
Hyperparameters: Adoption rate η, Retraining frequency Δ
Output Data: Dataset $\mathcal{S}$ of simulated visits

1: $\mathcal{A} \leftarrow \text{AlgorithmTraining}(D_{\text{train}})$ ▷ Train the recommender system
2: $(u, v, t) \leftarrow \text{last}(D_{\text{train}})$ ▷ Last tuple of the training dataset
3: $t_{\text{last\_training}} \leftarrow t$ ▷ Timestamp of the last training
4: $\mathcal{S} \leftarrow \varnothing$
5: **for** $(u, v_{\text{current}}, t)$ in D_{post} **do** ▷ Iterate over the visits
6: $c \leftarrow \text{CategorySelection}(v_{\text{current}})$ ▷ Select the category of the next venue
7: $v_{\text{next}} \leftarrow \text{VenueSelection}(u, v_{\text{current}}, t, \mathcal{A}, \eta)$ ▷ Select the next venue
8: $\mathcal{S} \leftarrow \mathcal{S} \cup \{(u, v_{\text{next}}, t)\}$ ▷ Add the simulated visit to the dataset
9: **if** $t - t_{\text{last\_training}} > \Delta$ **then** ▷ Check if it's time to retrain the model
10: $\mathcal{A} \leftarrow \text{AlgorithmTraining}(D_{\text{train}} \cup \mathcal{S})$ ▷ Retrain the recommender system
11: $t_{\text{last\_training}} \leftarrow t$ ▷ Update last training time
12: **end if**
13: **end for**

Algorithm 1 Simulation framework that models the feedback loop

3.2.1 Category selection

For each visit by an individual u in D_{post}, we observe the actual venue $v \in V$ visited by u and determine the corresponding category $c = \ell(v)$. Using the actual category of the visited venue, as observed in the data, serves as a proxy for the underlying intention or agenda of the individual u.

3.2.2 Venue selection

Given the target category c representing the type of venue to be visited, the individual $u \in U$ must select a specific venue of that category. The venue selection process is detailed in Algorithm 2.

The first step involves deciding whether to rely on the recommender system $\mathcal{A}$ or on personal knowledge (line 4 of Algorithm 2). This decision reflects the user's choice between following algorithmic suggestions or acting autonomously. It is modelled as a Bernoulli trial with success probability $\eta \in [0, 1]$, referred to as the adoption rate, which quantifies the likelihood that u accepts the recommendation. The parameter η plays a central role in the simulation, controlling the degree to which the user behaviour is shaped by the recommender. If the Bernoulli trial succeeds, u follows the recommendation (recommendation-based choice, line 5); otherwise, u selects a venue independently (autonomous choice, lines 7–10).

Input Data: Venue set V, Current venue $v_{\text{current}} \in V$, Target category c, User history D_u, Training dataset D_{train}, Simulated dataset S, Jump length distribution $\mathcal{P}_{\text{jump}}$
Hyperparameters: Adoption rate η, Exploration probability p
Output Data: Selected venue v_{next}

1: $r \leftarrow \text{Sample}(\mathcal{P}_{\text{jump}})$ ▷ Sample a jump length from the empirical distribution
2: $r^* \leftarrow \text{Median}(\mathcal{P}_{\text{jump}})$ ▷ Median of jump distribution
3: $V_{c,r}^{(v_{\text{current}})} \leftarrow \{v \in V \mid \ell(v) = c \wedge \text{dist}(v_{\text{current}}, v) \leq r\}$ ▷ Candidate venues
4: **if** Bernoulli(η) **then**
5: $v_{\text{next}} \leftarrow \mathcal{A}(u, V_{c,r}^{(v_{\text{current}})})$ ▷ Select venue using recommender system
6: **else**
7: **if** Bernoulli($1-p$) **then** ▷ Return to a previously visited venue
8: $v_{\text{next}} \leftarrow \text{PreferentialReturn}(D_u, c)$
9: **else** ▷ Explore a new venue
10: $v_{\text{next}} \leftarrow \text{Explore}(V_{c,r}^{(v_{\text{current}})}, D_u, r^*)$
11: **end if**
12: **end if**

Algorithm 2 Venue selection for individual u

3.2.3 Recommendation-based choice

If an individual u follows the suggestion of the recommender system $\mathcal{A}$ (line 5 of Algorithm 2), a set of candidate venues $V_{c,r}^{(v_{\text{current}})} \subseteq V$ is retrieved (line 3, Algorithm 2). This set includes all venues belonging to the target category c and located within a radius r of the current location of the individual v_{current}.

The radius r is sampled from the empirical distribution of jump lengths observed in D, reflecting the tendency of individuals to prefer short-range displacements (Gonzalez et al., 2008). This distribution is obtained by enumerating all pairs of consecutive venues (v_{i-1}, v_i) in the visit sequences D_u for each user $u \in U$, and computing the great-circle distance between v_{i-1} and v_i using their geographic coordinates.

Each venue $v \in V_{c,r}^{(v_{\text{current}})}$ is assigned a score $\mathcal{R}_\theta(u, v, t)$. The resulting score set S is normalized using min-max scaling, yielding the set of normalized scores $\hat{S}$. The recommender system $\mathcal{A}$ then returns the top-k venues in $L \in V_{c,r}$ based on $\hat{S}$. Finally, u selects a venue $v_{\text{next}} \in L$ with probability proportional to its normalised score $\hat{S}(v)$.

3.2.4 Autonomous choice

If the individual u moves independently of the recommender system $\mathcal{A}$ (line 7 of Algorithm 2), we model their choice of venue using a modified version of the exploration and preferential return (EPR) model, a well-established framework for human mobility introduced by Song et al. (2010). The EPR model captures the fundamental mechanism of human mobility as a trade-off between returning to previously visited locations and exploring new ones. It is a foundational model in the study of human mobility (Barbosa et al., 2018)and

its flexibility makes it a suitable basis for modelling autonomous mobility decisions in our framework.

Based on the EPR model, at each decision point, we let the individual u choose between:

- *Preferential return* (line 8, Algorithm 2): with probability $1 - p$, the individual selects a venue from their set of previously visited locations, with probability proportional to the number of times each venue has been visited by u.
- *Exploration* (line 10, Algorithm 2): with probability p, the individual selects a new (unvisited) venue located within a radius r, where r is sampled from the empirical distribution of jump lengths observed in D. This radius-based sampling reflects the heavy-tailed nature of human travel patterns, where individuals occasionally make long jumps but typically explore within a constrained geographic range (Gonzalez et al., 2008). Among the venues within this radius, the probability of selecting a specific venue v is proportional to its relevance. The radius-based sampling in the exploration phase reflects the heavy-tailed nature of human mobility (Gonzalez et al., 2008) and aligns with extensions of the EPR model (Barbosa et al., 2018; Pappalardo et al., 2015, 2016). We assume that individuals favour dense or attractive areas, rather than choosing uniformly at random within the radius, as suggested by Pappalardo et al. (2015, 2016). The relevance of a venue $v \in V$ is defined as the number of other venues located within a fixed radius r^* from v, where r^* is set to the median of the jump length distribution computed on D. Intuitively, a venue is more relevant if it lies in a dense area where many other venues fall within this typical range, making it more likely to be part of users' mobility patterns. Formally, the relevance is given by $rel(v) = |\{v' \in V \setminus \{v\} | d(v, v') \leq r^*\}|$.

If a user lacks enough venues of the target category c within the sampled radius r, a fallback mechanism ensures a valid venue is selected. First, the mechanism searches for a venue in a broader category within the same radius. For example, if the target category is *Sushi Restaurant*, the system will attempt to find any venue under the broader *Food* category. If no venue is found within this broader category, the individual is assigned the geographically nearest venue of the original target category, regardless of distance. Similarly, in the return phase, if the individual has not previously visited any venue in the target category, we look for one in the broader category. If that also fails, the individuals explores, visiting a new venue.

3.2.5 Algorithm retraining

The recommender system $\mathcal{A}$ is periodically retrained during the simulation at fixed intervals of Δ time units (line 10, Algorithm 1). At each retraining step, the model is updated using an enriched dataset $D_{\text{train}} \cup \mathcal{S}$, consisting of the original training set D_{train} combined with the set $\mathcal{S}$ of simulated visits generated up to that point. In practice, this step corresponds to updating the parameters θ of the scoring function $\mathcal{R}_\theta$ based on this enriched dataset.

This retraining procedure allows $\mathcal{A}$ to iteratively refine its understanding of individual preferences by incorporating both historical and simulated behavioural data. In doing so, it captures the dynamic nature of the human–AI feedback loop, where algorithmic outputs influence human behaviour, which in turn shapes future recommendations.

3.3 Metrics

We assess the impact of the human–AI feedback loop on urban dynamics from two complementary perspectives: one focused on venues and the other on individuals.

3.3.1 Venue perspective

We evaluate the inequality in the distribution of visits across venues using the Gini coefficient. This measure captures the extent to which visits are concentrated on a subset of venues: a Gini coefficient of 0 indicates perfect equality (all venues receive the same number of visits), while higher values indicate increasing concentration of visits on fewer venues. Our use of the Gini coefficient is motivated by studies showing that recommender systems can amplify the inequality in the distribution of consumed items (Pedreschi et al., 2025; Pappalardo et al., 2024; Boratto et al., 2021; Elahi et al., 2021; Lee & Hosanagar, 2019). In our context, the Gini coefficient quantifies to what extent algorithmic recommendations direct individuals towards a limited number of venues, increasing their popularity while marginalising others.

Given a dataset of visits D, we compute the collective Gini coefficient as:

$$G = \frac{1}{|V|}\left(|V| + 1 - 2 \cdot \frac{\sum_{i=1}^{|V|}(|V| + 1 - i)x_i}{\sum_{v \in V} x_i}\right), \tag{1}$$

where $x_i = |(u, v, t) \in D : v = v_i|$ is the total number of visits to the venue $v_i \in V$, and the values x_i are sorted in ascending order.

To complement this collective view, we also measure inequality from the perspective of individual behaviour. For each individual $u \in U$, we calculate an individual Gini coefficient G_u, applying Eq. 1 to D_u (the visits made by u). That is, we compute the inequality in how frequently each venue is visited by the individual. The average individual Gini coefficient is then given by:

$$\overline{G} = \frac{1}{|U|}\sum_{u \in U} G_u, \tag{2}$$

which captures the typical inequality in venue usage across individuals.

3.3.2 Individual perspective

From the individual perspective, we construct a co-location network in which nodes represent individuals and edges indicate co-presence at the same venue within a specified time window. Intuitively, this network captures how individuals encounter one another through shared spatial activity in the city.

Formally, let $D = (u_i, v_j, t_k)$ be the dataset of individual–venue visits. Given a time window $[t_1, t_2]$, the co-location network is defined as an undirected graph $\mathcal{G} = (U', E)$, where $U' \subseteq U$ is the set of individuals active in that time window (i.e., who move at least

once), and $(u_m, u_n) \in E$ if and only if individuals u_m and u_n visited the same venue v_j within the time window. That is $\exists\,(u_m, v_j, t_k),\ (u_n, v_j, t_l) \in D$ such that $t_k, t_l \in [t_1, t_2]$.

To investigate the presence of hierarchical structure in the social interactions encoded by the co-location network, we analyse the tendency of high-degree nodes to form a densely interconnected core (rich club). There is substantial evidence showing that rich-club structures can influence the dynamics of complex systems (Pedreschi et al., 2022; Bertagnolli & De Domenico, 2022; Berahmand et al., 2018; Colizza et al., 2006), such as fostering socio-economic development (Kang et al., 2022) and shaping the spatial organisation of cities (Jia et al., 2022).

Following the method proposed by Zhou and Mondragón (2004), we identify the set U_{rich} as the h nodes with the highest degree in G, and compute the rich-club density as the ratio of the number of actual links among U_{rich} to the maximum possible number of links in a complete subgraph of h nodes, i.e. $\frac{h(h-1)}{2}$.

We also examine how the degree distribution of social interactions may be affected by the influence of the recommender system. Given that the degree distribution $P(k)$ in real-world networks often exhibits heavy-tailed behaviour, we fit a linear regression to the log-log plot of $P(k)$ derived from the co-location network $\mathcal{G}$, and report the absolute value of the slope of the fitted line, denoted by α. Larger values of α correspond to a steeper decline, reflecting a distribution closer to exponential decay, whereas smaller values reflect a more uniform structure. We estimate $P(k)$ using the empirical degree distribution, defined as $P(k) = \frac{|\{u \in U' : \deg(u) = k\}|}{|U'|}$, where $\deg(\cdot)$ denotes the degree of a node in $\mathcal{G}$.

4 Experimental settings

In this section, we describe the experimental settings of our study. Section 4.1 presents the recommendation systems evaluated in our simulations. Section 4.2 details the dataset on which our experiments are based, and Sect. 4.3 outlines the values and configurations of the parameters used in the simulation framework.

4.1 Benchmark recommenders

Various recommender systems have been designed to assist users in discovering the most relevant and valuable content (Ricci et al., 2021a, b). Among the various approaches, collaborative filtering represents one of the most widely adopted families of algorithms, as it relies solely on binary feedback indicating whether a user has interacted with a particular item (Koren et al., 2021). These techniques have also been successfully applied to the recommendation of venues in urban settings (Rahmani et al., 2022a), consistently achieving high accuracy across benchmark datasets for next-venue recommendation.

We benchmark several recommendation models spanning classical neighbours-based methods, matrix factorization techniques, deep generative models, geographical recommenders, and approaches based on graph neural networks. The list of models included in our study is detailed below:

- *User-KNN* (Aiolli, 2013): Neighbourhood-based method that recommends items based

on the preferences of users with similar interaction histories;
- *Item-KNN* (Aiolli, 2013): Similar to User-KNN, but recommends items that are similar to those the user has previously interacted with;
- *Matrix factorization (MF)* (Sarwar et al., 2000): Model-based approach that decomposes the user–item interaction matrix into latent user and item feature vectors, enabling prediction of missing interactions;
- *Bayesian personalized ranking (BPRMF)* (Rendle et al., 2009): Extension of matrix factorization that optimizes for pairwise ranking by assuming users prefer observed items over unobserved ones, and learning to rank accordingly;
- *Multinomial variational autoencoder (MultiVAE)* (Liang et al., 2018): Generative model based on deep learning that captures user preferences by learning a probabilistic latent space using variational inference;
- *Light graph convolutional network (LightGCN)* (He et al., 2020): Recommendation system based on graph neural networks that propagates user and item embeddings over a user–item bipartite graph;
- *PGN* (Sánchez & Dietz, 2022): Hybrid approach that averages the scores of User-KNN, a popularity-based recommender, and a geographical model that suggests venues near the centroid of the user's previously visited locations.

As different models exhibit varying levels of bias in location-based recommender systems (Rahmani et al., 2022b), and algorithmic bias is directly linked to the long-term effects of feedback loops (Mansoury et al., 2020; Chaney et al., 2018), our comprehensive evaluation is designed to capture the different behavioural shifts and inequality patterns that may emerge from each recommendation strategy over time. Details about the hyperparameters and the training procedure of the recommender systems are in Appendix A. As a sanity check, we report the performance of all the recommenders trained on D_{train} and tested on D_{post} in terms of *HitRate*@20 and *mRR*@20, see Table 2 in Appendix A.

4.2 Dataset

We use a Foursquare dataset containing mobility data from New York City, collected over a period of 304 days—from April 3, 2012, to February 16, 2013 (Yang et al., 2014). The dataset includes 227,428 check-ins, each annotated with a timestamp, GPS coordinates, and detailed venue categories. We also replicate our analysis using a Foursquare dataset from Tokyo. Since the results closely align with those from New York City, we present the NYC results in the main text and report the Tokyo findings in Appendix B.

For consistency, we assign a single category to each venue, selecting the second category level as it provides a balanced granularity between the broader first category and the highly specific third category.[3] For example, a venue with the third category *Italian Restaurant* is classified under the second category *Restaurant* and the first category *Dining and Drinking*.

We apply several preprocessing steps on the raw data. Specifically, we remove check-ins to categories representing familiar locations (e.g. *Office, Home, Meeting Room*), modes of transportation (e.g. *Train, Ferry, Road*), and venues labelled with *Unknown* categories. Appendix C provides the list of excluded categories.

[3] https://observablehq.com/d/94b009d907d7c023

After preprocessing, the New York City dataset contains 166, 306 check-ins from 1, 083 unique users to 23, 459 unique venues that span 159 distinct categories. On average, each category includes approximately 172 venues. In our simulations, we set $T_{\text{train}} = 210$ days and $T_{\max} = 304$ days.

4.3 Parameter settings

Our simulation framework relies on a set of parameters that govern the behavioural rules of the agents involved. Table 1 summarises the symbols used, their corresponding meanings, and the values adopted throughout the simulations. The framework requires only three parameters, aligning with our goal of assessing the effect of recommendation systems on urban dynamics with minimal changes to all other factors, thus isolating the impact of algorithmic intervention.

The adoption rate η is a key parameter in our framework, representing the probability that an individual follows the suggestion provided by the recommender system. To systematically investigate the effects of recommendation systems, we vary η across a predefined grid of values ranging from 0 to 1, with a step size of 0.2.

The parameter Δ reflects how frequently the recommendation system is updated (retrained) to account for the evolving nature of user preferences. We set $\Delta = 7$ days, meaning the recommender is retrained weekly to follow the dynamics introduced by the human-AI feedback loop. Results for $\Delta = 1$ day are similar; we provide them in Appendix D.

When a human agent does not rely on the recommender system, it follows the EPR model (Song et al., 2010). Empirical analyses in Song et al. (2010) confirm the model's ability to reproduce the scaling properties of human mobility, capturing the balance between returning to familiar locations and exploring new ones. This dual behavioural perspective has been validated on real-world data, where the probability of exploring a new venue is given by $p = \rho \times |V|^{-\gamma}$, with $|V|$ denoting the number of unique venues. Following the empirical estimates of ρ and γ reported by Song et al. (2010), we set their values to 0.6 and 0.21, respectively.

5 Results

Figure 2 compares the effects of the feedback loop as a function of the adoption rate η. All values are averaged over five independent simulation runs. Figure 2a and b report the average individual Gini coefficient $\overline{G}$ and the collective Gini coefficient G of venue visits, respectively. We remind that the former captures how evenly each individual distributes their visits across venues, while the latter reflects overall disparities in venue popularity.

For the sake of brevity, we do not report results for PGN in the main text, as they closely mirror the patterns observed for MF and MultiVAE. A detailed comparison that includes PGN is provided in Appendix E.

Table 1 Simulation parameters adopted in our framework

η	Adoption rate	$\{0.0, 0.2, 0.4, 0.6, 0.8, 1.0\}$
Δ	Retraining frequency	7 days
p	Exploration probability	$0.6 \times \lvert V\rvert^{-0.21}$ (Song et al., 2010)

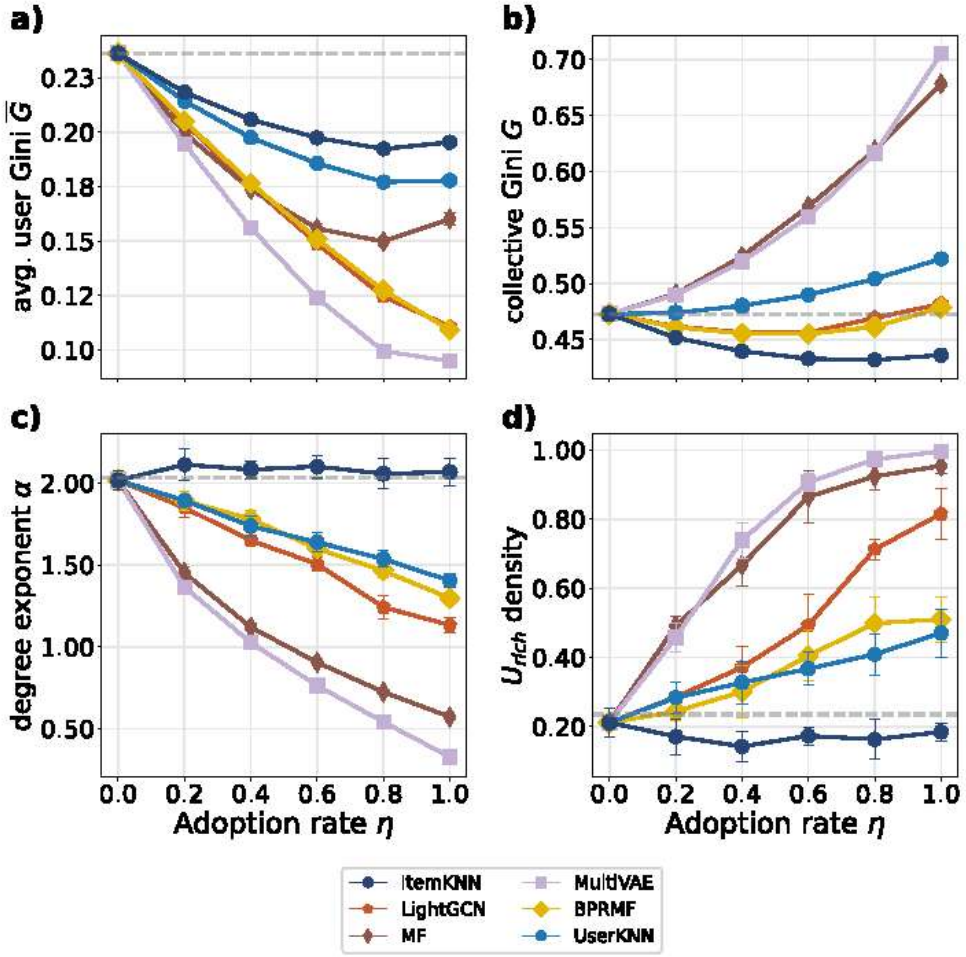


Fig. 2 Effects of recommender system adoption η. **a** Average individual Gini coefficient $\overline{G}$; **b** Gini coefficient of the distribution of visits across venues G; **c** Absolute value of the slope of the co-location network's degree distribution α; **d** Internal density of the rich club (top 15 nodes by degree) in the co-location network. Each curve represents a recommender system; each point represents the average over five simulation runs; the error bars represent the standard deviations

The feedback loop consistently reduces $\overline{G}$ across all simulations, regardless of the specific recommender system embedded within it. Compared to the baseline with no algorithmic influence ($\eta = 0$, dashed line at $\overline{G} \approx 0.2$ in Fig. 2a), the activation of the feedback loop leads to lower values of $\overline{G}$, with the effect intensifying as the adoption rate η increases. While the feedback loop involving ItemKNN and UserKNN results in modest reductions, those involving deep learning-based models—BPRMF, LightGCN, and MultiVAE—produce significantly stronger decreases. Notably, for $\eta = 1$ (full reliance on the recommender system), the MultiVAE-based feedback loop reduces $\overline{G}$ to approximately 0.09, a relative decrease of about 60% compared to the baseline. This result suggests that, when algorithmic influence is strong, individuals tend to diversify their mobility patterns, visiting a broader and more varied set of venues rather than repeatedly returning to the same few.

Key Result 1

The feedback loop involving recommender systems boosts individual diversity, as reflected by a consistent drop in the average individual Gini coefficient.

However, this increase in individual diversity does not necessarily translate into more diverse collective behaviour. The feedback loop leads to an increase in the collective Gini

coefficient G with rising η for some recommender systems (see Fig. 2b). UserKNN, MF, and MultiVAE all result in a sharp increase in G as adoption grows. In particular, for $\eta = 1$, the feedback loop with MultiVAE reaches $G = 0.7$, an increase of approximately 47% compared to $\eta = 0$, indicating a more uneven distribution of visits across venues. Deep learning–based models such as LightGCN and BPRMF contribute to slightly reducing inequality, except in full-adoption scenarios. ItemKNN is the only model that slightly but consistently reduces global inequality compared to the baseline, reaching $G \approx 0.44$, a relative reduction of about 7% compared to $\eta = 0$. These results suggest that a systematic increase in individual diversity may be accompanied by a simultaneous concentration of visits on a smaller subset of venues.

Key Result 2

More individual diversity does not imply more collective diversity: for some recommender systems, the feedback loop concentrates visits on a small set of popular venues.

Analysis of the co-location network reveals that the feedback loop increases the absolute value of the slope of the power-law degree distribution, α, for all recommender systems except ItemKNN, with the effect particularly pronounced for MultiVAE and MF (see Fig. 2c). A lower slope α corresponds to a flatter distribution, indicating that co-location ties are more evenly distributed across individuals rather than concentrated among a few highly connected ones. Thus, in most cases, the feedback loop reduces the skewness of the co-location structure, promoting broader interpersonal mixing. In contrast, for ItemKNN, the slope α remains close to the baseline as the adoption rate η increases, indicating little change in user co-presence patterns (squares in Fig. 2c).

A similar pattern is observed in the rich-club structure U_{rich}, defined as the top 15 individuals with the highest degree in the co-location network. For all recommender systems except ItemKNN, the internal density of connections within this group—the ratio of observed to possible links—increases sharply with higher values of η, approaching nearly 1 for MultiVAE (see Fig. 2d). This suggests that highly connected individuals increasingly tend to co-locate with one another. In contrast, for ItemKNN, the rich-club density remains close to the baseline level of approximately 0.2, indicating little change in elite clustering. Results for Tokyo closely mirror those for New York City, with minor differences in rich-club density likely due to Tokyo's higher initial network density and more compact geography (see Appendix B). In addition, a robustness analysis using MultiVAE—the recommender system showing the strongest effects—confirms that our main findings are stable under alternative retraining configurations. Specifically, we tested both daily retraining and a sliding-window update scheme, observing no substantial differences in the results (see Appendix D).

Key Result 3

The feedback loop promotes broader mixing of co-locations, while rich-club individuals become nearly fully connected within the co-location network.

5.1 Venue perspective

Since the feedback loops involving MultiVAE and ItemKNN represent the two extreme cases across all the metrics discussed, we provide a more detailed analysis of the patterns that emerge from their integration within our framework. The contrasting spatial effects of

MultiVAE and ItemKNN are illustrated in Figs. 3 and 4, which show both individual and collective patterns under the extreme cases of null and full adoption ($\eta = 0$ and $\eta = 1$).

5.1.1 MultiVAE

For MultiVAE, Fig. 3a-b shows the spatial distribution of visits for a representative individual (User 900). Under the scenario of no recommendation ($\eta = 0$), visits are mostly concentrated on just two venues in western Midtown Manhattan (Fig. 3a). Under full adoption ($\eta = 1$), the activity of User 900 becomes more spatially dispersed, with visits more evenly distributed throughout the city (Fig. 3b). This shift is captured by the individual Gini coefficient, which drops from $G_u = 0.41$ to $G_u = 0.19$—a reduction of approximately 54%—with a consistent decline observed as the adoption rate η increases (Fig. 3c).

At the collective level, MultiVAE shows the opposite effect. Figure 3d-e shows the aggregated spatial distribution of visits for a random sample of 250 users. When $\eta = 0$, visits are relatively uniformly spread across the city, including peripheral neighbourhoods and outer boroughs, with $G = 0.32$ (see Fig. 3d). When $\eta = 1$, visits become concentrated in a small number of venues, especially in Midtown and Lower Manhattan. The Gini coefficient increases to $G = 0.50$, marking a 56% rise in venue popularity inequality (see Fig. 3f).

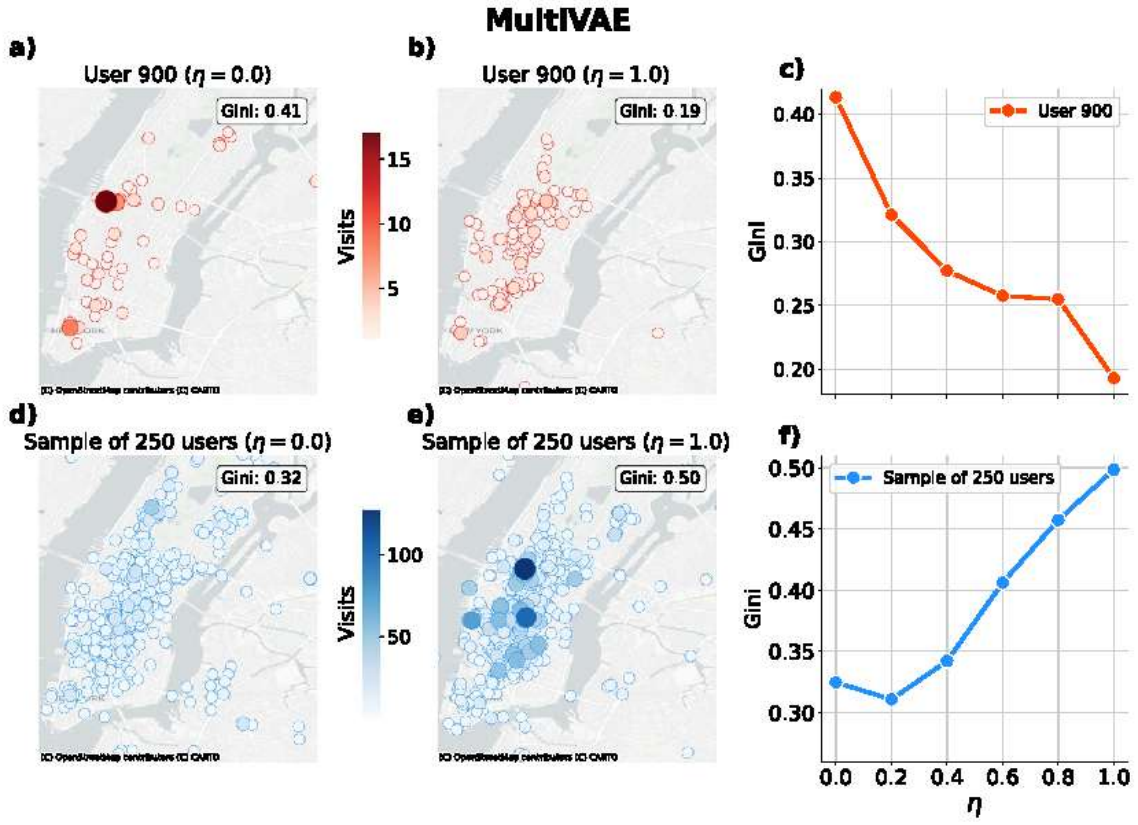


Fig. 3 Venue visitation patterns under MultiVAE for null ($\eta = 0$) and full ($\eta = 1$) adoption. **a**, **b** Geographical distribution of visits by a representative user (User 900). Visit intensity is proportional to point size and color. When $\eta = 1$, visits become more evenly distributed across venues, with the individual Gini coefficient decreasing from $G_u = 0.41$ to $G_u = 0.19$. **c** Individual Gini coefficient for User 900 as a function of η, showing a consistent decline. **d**, **e** Aggregated venue visits for a sample of 250 users. When $\eta = 1$, visits are more concentrated in a few high-traffic venues, particularly in central Manhattan. **f** Gini coefficient of venue visit distribution across the sample, rising from $G = 0.32$ to $G = 0.50$ with increasing η

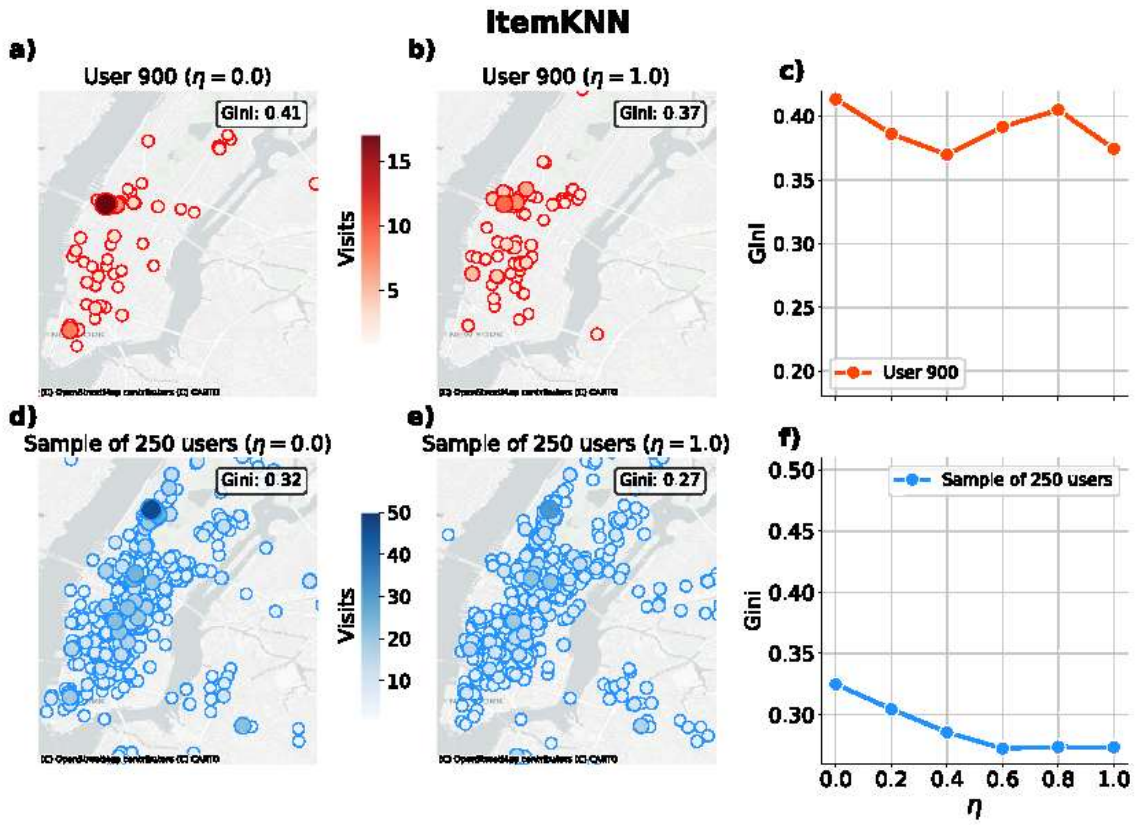


Fig. 4 Venue visitation patterns under ItemKNN for null ($\eta = 0$) and full ($\eta = 1$) adoption. **a**, **b** Geographical distribution of visits by a representative user (User 900). A modest shift toward a more balanced visitation pattern is observed, with the individual Gini decreasing from $G_u = 0.41$ to $G_u = 0.37$. **c** Individual Gini coefficient for User 900 as a function of η, showing minor fluctuations. **d**, **e** Aggregated venue visits for a sample of 250 users. Visit distributions remain fairly dispersed under both conditions. **f** Gini coefficient of venue visit distribution across the sample, decreasing slightly from $G = 0.32$ to $G = 0.27$

These results confirm the results shown in the previous section: while the feedback loop with MultiVAE promotes more balanced behaviour at the individual level, it drives convergence toward a narrow set of highly frequented locations at the collective level, reinforcing popularity inequality. This pattern is typical of the feedback loops involving most other recommender systems we examined.

5.1.2 ItemKNN

ItemKNN produces more stable spatial dynamics than MultiVAE (see Fig. 4). For User 900, the shift from $\eta = 0$ to $\eta = 1$ results in a slightly more balanced, yet still unequal, visitation pattern (see Figs. 4a-b). The individual Gini coefficient decreases from $G_u = 0.41$ to $G_u = 0.37$, a reduction of approximately 9.8%, with only minor fluctuations across η (Fig. 4c). At the collective level, ItemKNN contributes to a small improvement in spatial balance. Figure 4d-e shows that the city-wide distribution of visits remains fairly dispersed across both conditions. The Gini coefficient decreases from $G = 0.32$ to $G = 0.27$ under full adoption (Fig. 4f confirms this downward trend). Unlike MultiVAE, the feedback loop involving ItemKNN does not lead to spatial convergence around a narrow set of venues. Instead, it preserves heterogeneity in mobility patterns while slightly enhancing the overall equity in how urban space is visited.

5.2 Individual perspective

5.2.1 MultiVAE

The feedback loop involving MultiVAE substantially reshapes the structure of urban co-presence (see Fig. 5). We remind that, in the co-location network, two individuals are linked if they visit the same venue within the same epoch of the simulation. Figure 5a,c show the degree distributions of the co-location network under the two extreme cases of null and full adoption ($\eta = 0$ and $\eta = 1$). Under the autonomous choice scenario ($\eta = 0$), the degree distribution follows a steeper power-law decay, with a median degree of 4 and the absolute

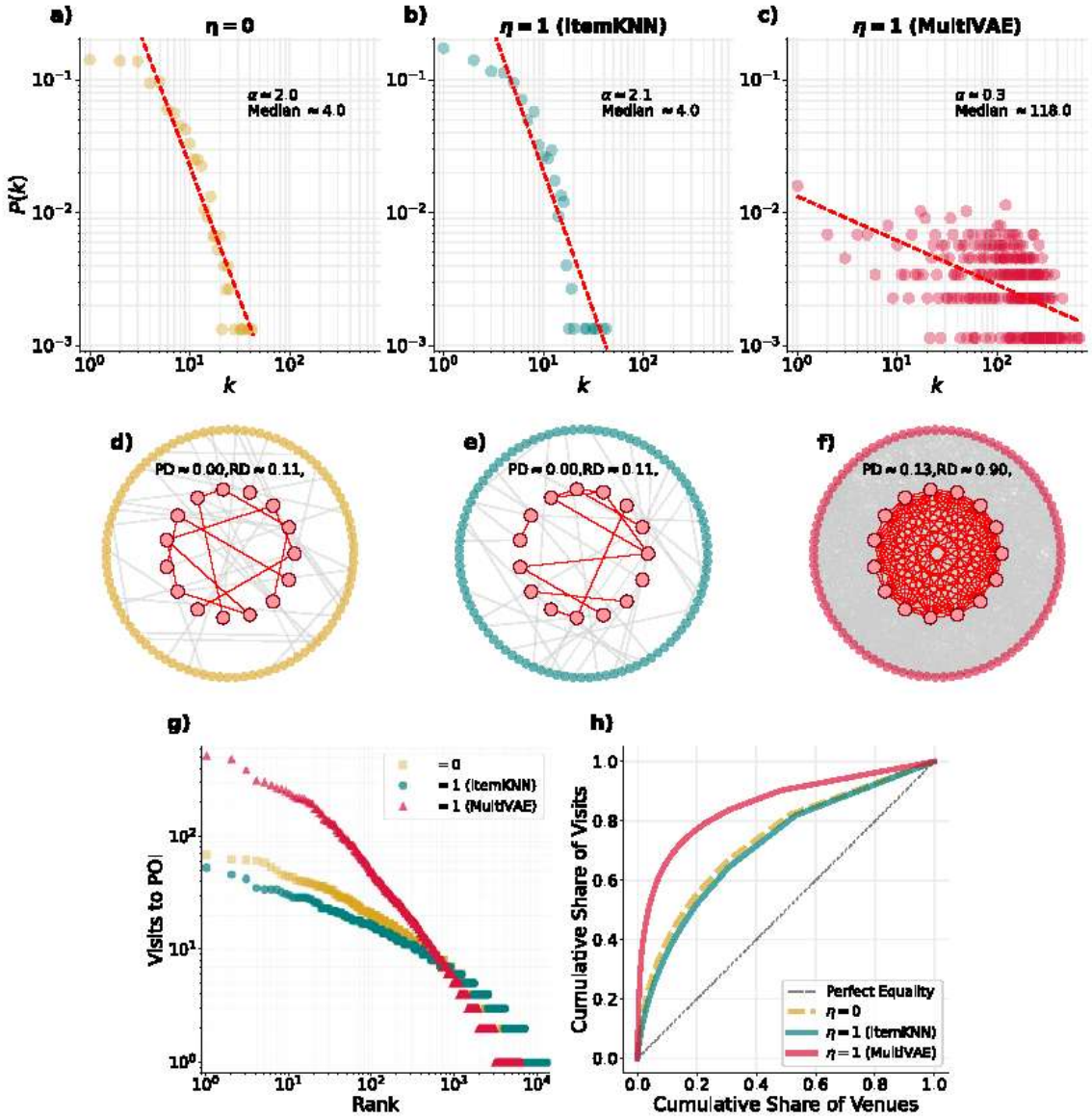


Fig. 5 Effects of the feedback loop under null ($\eta = 0$) and full ($\eta = 1$) adoption. **a–c** Degree distribution of the co-location network, for null adoption, full adoption with ItemKNN and full adoption with MultiVAE. MultiVAE significantly flattens the distribution and increases median degree, while ItemKNN maintains the original network structure. **d–f** Visualization of the co-location network (sampled), highlighting the rich-club structure (red nodes) and peripheral nodes (circular layout) for null adoption, full adoption with ItemKNN and full adoption with MultiVAE. Rich-club density increases dramatically for MultiVAE, while remaining constant for ItemKNN. **g**, **h** Rank-size distribution and Lorenz curves of venue visits: under η=1 for MultiVAE, top-ranked venues receive disproportionately more visits

value of the slope $\alpha \approx 2.0$ (Fig. 5a). When users fully adopt MultiVAE ($\eta = 1$), the degree distribution flattens significantly, with a heavier tail and a substantial increase in median degree (Fig. 5c). Notably, the number of nodes in the network also increases, which implies broader participation in the social layer under algorithmic influence.

Figure 5d,f offers a visual snapshot of the co-location network structure under the null and full adoption scenarios for MultiVAE. The inner red nodes denote the rich club and the peripheral nodes lie along the outer circle. When $\eta = 0$, the peripheral subnetwork is sparse (Peripheral Density, PD is almost 0) and the rich club is weakly connected (Rich-club Density, $RD \approx 0.11$). When $\eta = 1$ with MultiVAE, the structure shifts dramatically: peripheral nodes become more interconnected ($PD \approx 0.13$) and the rich club much more densely connected ($RD \approx 0.90$), forming a dense social core.

These structural shifts in co-location patterns are tightly connected to changes in venue popularity. Figure 5g reports the rank-size distribution of venue visits. With full adoption of MultiVAE, the head of the distribution grows significantly: a few venues receive disproportionately high attention, reflecting an increase in spatial centralisation. This dynamic reinforces not only high-degree users but also high-frequency venues, suggesting that popular places become even more dominant. This pattern is further confirmed in Fig. 5h, which reports the Lorenz curves of visit distributions under different scenarios. The curve for MultiVAE with $\eta = 1$ deviates more strongly from the equality line, indicating higher overall inequality in venue visitation, despite the increased uniformity in individual user behaviour and co-location degree.

5.2.2 ItemKNN

In contrast to MultiVAE, the feedback loop involving ItemKNN has a minimal effect on the structure of the co-location network and the distribution of visits across venues. Figure 5a,b show the degree distributions of the co-location network under $\eta = 0$ and $\eta = 1$ with ItemKNN, respectively. Both distributions follow similar power-law trends, with nearly identical slopes ($\alpha \approx 2.0$ and $\alpha \approx 2.1$) and a stable median degree of 4. This indicates that the feedback loop involving ItemKNN does not substantially alter the topology of social co-presence: individuals maintain similar numbers of co-located interactions regardless of whether they follow the algorithm's suggestions.

A structural view of the co-location network confirms this stability. Figure 5d,e visualises the rich-club structure under null adoption and full adoption with ItemKNN. In both cases, the peripheral density remains at zero ($PD \approx 0.00$), and the rich-club density remains almost constant ($RD = 0.11$ and $RD \approx 0.11$ respectively). This suggests that the feedback loop with ItemKNN neither strengthens the elite core of highly connected individuals nor promotes wider cohesion among peripheral ones.

The implications from a venue perspective are shown in Fig. 5g-h. The rank-size distribution of venue visits (Fig. 5g) shows that ItemKNN largely preserves the pattern of venue visits across adoption levels, with only minor variation in the long tail. The Lorenz curves (Fig. 5h) reveal a slight shift towards greater equality under full adoption of ItemKNN, as its curve moves marginally closer to the diagonal compared to the null adoption scenario. This indicates a modest reduction in the inequality of venue visits, although the general pattern remains similar.

Overall, the feedback loop involving ItemKNN preserves the existing structure of user interactions and venue usage. Unlike more sophisticated recommender systems, it avoids concentrating visits or reinforcing strong social cores, instead maintaining a relatively stable distribution of co-location and visits across the system.

5.3 Understanding the individual-collective diversity trade-off

Within the feedback loop, the recommender systems analysed reduce inequality in how individuals distribute their visits across venues, yet most recommenders simultaneously increase the inequality in how visits are distributed across venues at the collective level. How can these two seemingly contradictory mechanisms coexist?

To address this question, we compare the exploratory visits of individuals during the simulation with those recorded in the historical training data (D_{train}). We model the interaction between individuals and venues as a bipartite network, where one set of nodes represents individuals, the other represents venues, and edges correspond to visits. We construct two bipartite networks: one based on the historical training data, and another based on exploratory visits simulated under full algorithmic adoption ($\eta = 1$). In the training network, an edge connects an individual to a venue if it was visited during the training period. In the simulation network, an edge is added only if the individual visited a venue for the first time, i.e., a venue not in D_{train}.

We focus on MultiVAE, which exhibits the strongest impact on inequality patterns. We group venues into deciles based on their number of unique visitors in the training set. Each decile represents 10% of venues, ranked from least (decile 1) to most popular (decile 10).

The degree of a venue in the bipartite network reflects how many users visited it. Since we model binary interactions, the degree is equivalent to the weighted degree. To normalise across the network, we compute the normalised weighted degree for each venue v:

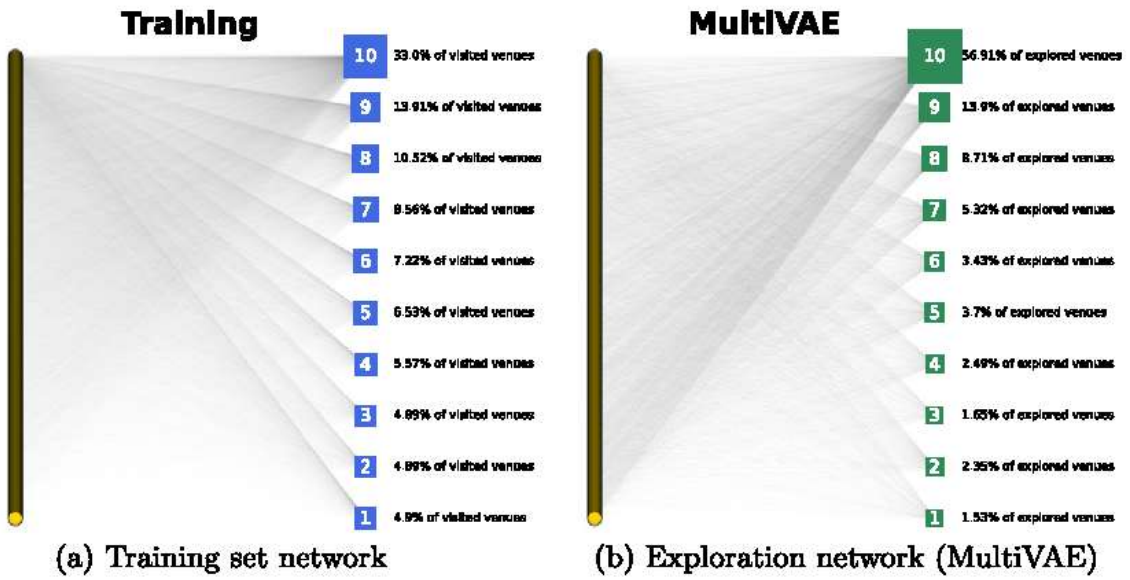
$$\hat{d}_v = \frac{d_v}{\sum_{v' \in V} d_{v'}} \tag{3}$$

We then aggregate these values within each decile to compute the share of attention directed to that group:

$$\hat{D}_g = \sum_{v \in g} \hat{d}_v \tag{4}$$

This measure, $\hat{D}_g$, captures the fraction of all user–venue interactions involving venues in the group g.

Figure 6a shows that while the most popular venues (decile 10) already attract the largest share of attention (33%) in the training data, the other deciles also receive a significant portion of visits. After the simulation, under the influence of MultiVAE, attention becomes far more concentrated (see Fig. 6b): decile 10 captures 56.91% of all interactions in the exploration phase, an increase of nearly 24% compared to the training data. This comes at the expense of nearly every other decile, whose shares decline.



(a) Training set network (b) Exploration network (MultiVAE)

Fig. 6 User-venue bipartite networks. In the bipartite network, individuals are on the left and venue deciles are on the right. Each decile contains 10% of venues, ranked by their popularity in the training data. Percentages indicate each decile's share of total user–venue interactions. In the bipartite network corresponding to the training data (**a**), an edge indicates a visit during the training period. In the bipartite network of the simulation data (**b**), an edge represents a simulated visit to a venue not seen during training (i.e., a novel exploration). Venues with fewer than 3 unique visitors are excluded

To characterise this shift better, we compute the change in attention share for each decile between the training data and simulation data: $\delta_g = \hat{D}_g^{(\text{simulation})} - \hat{D}_g^{(\text{train})}$. As shown in Fig. 7, the shift is stark: decile 10 gains over 20 percentage points in attention, while all other deciles, especially the least popular ones, loose ground.

The analysis shows that the new venues visited by individuals, those driving the increase in individual diversity, are often the same across the population, typically the most popular ones. This reflects a rich-get-richer dynamic: while individual-level diversity improves, it does so by directing many individuals toward a narrow set of highly recommended venues. As a result, the system becomes more centralised, amplifying popularity bias and revealing a fundamental tension: recommendation systems can enhance fairness at the individual level while undermining it collectively. Appendix F offers a complementary perspective on this dichotomy, showing that user similarity increases with the adoption rate η for complex, representation-based models such as MultiVAE, while remaining stable for ItemKNN.

Key Result 4

Individual diversity rises mostly through visits to the same popular venues.

6 Discussion

Our findings contribute to the growing body of research on human–AI interaction by extending the analysis of feedback loops to the domain of human mobility.

Our main result is that the feedback loop leads to a divergence between individual and collective diversity for most recommendation systems. On the one hand, individuals distribute their visits more evenly across venues, resulting in increased individual diversity. On the other hand, the distribution of collective venue popularity becomes more concentrated, with

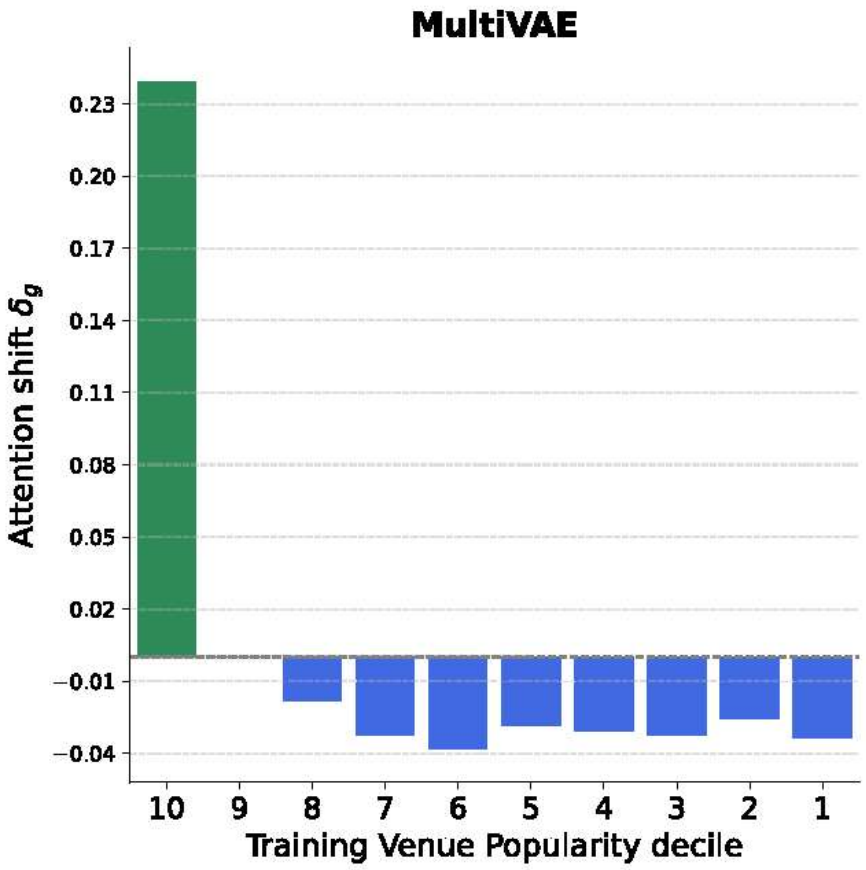


Fig. 7 Change in normalized weighted degree across venue popularity deciles (1–10). Bars show the difference in share of interactions between exploration and training under MultiVAE

a small set of venues attracting a disproportionate share of total visits—a decreased collected diversity. These two apparently contrasting mechanisms coexist because individual diversity rises mostly through visits to the same popular venues, which makes individual visitation patterns increasingly similar.

This individual-collective trade-off mirrors the patterns observed in other human-AI ecosystems (Pappalardo et al., 2024). In online retail, for example, the use of recommender systems often increases individual exposure to niche items while amplifying popularity bias at the aggregate level, leading to market concentration around top-ranked products (Lee & Hosanagar, 2019; Fleder & Hosanagar, 2009). A similar dynamic has been observed in the context of generative AI. Online experiments show that while generative models can boost individual creativity, they often reduce diversity at the collective level (Doshi & Hauser, 2024). Additionally, when data produced by an AI model is fed back into the model for fine-tuning, the resulting feedback loop can lead to mode collapse, producing increasingly homogenized outputs at scale (Shumailov et al., 2024). Our results suggest that the human-AI feedback loop produces this trade-off not only in digital content consumption but also in physical movement patterns within cities. This raises important concerns about the long-term effects of algorithmic mediation on urban accessibility, spatial inequality, and the viability of less popular or peripheral venues. In this regard, our study contributes to the emerging field of human-AI coevolution (Pedreschi et al., 2025), which focuses on measuring and modelling feedback loops between humans and algorithms, and understanding how these interactions shape the dynamics of complex systems.

Another key finding of this study is that the extent of the individual-collective trade-off depend on the algorithmic architecture used to train the recommender systems within the feedback loop. Empirical research shows that different types of recommender systems introduce distinct forms of bias, with no universally accepted solution: for example, neighbourhood-based models tend to exhibit less popularity and exposure bias than deep learning approaches, while graph neural networks are prone to degree bias, disproportionately favouring highly connected nodes (Chen et al., 2023; Klimashevskaia et al., 2024; Mansoury et al., 2019; Lesota et al., 2021; Elahi et al., 2021; Subramonian et al., 2024). Our findings are consistent with these observations: feedback loops with neighbourhood-based recommenders such as UserKNN and ItemKNN produce less pronounced biases than those with deep learning models such as MultiVAE. While our study does not offer a definitive explanation for the origins of these disparities, it adds evidence to the growing literature on how algorithmic design shapes fairness, diversity, and systemic outcomes in human-AI feedback loops.

Our work also contributes to the field of human mobility (Pappalardo et al., 2023; Barbosa et al., 2018), which seeks to analyse, predict, and model the fundamental patterns underlying human movement. In particular, our study challenges the core assumptions of mechanistic mobility models such as the Exploration and Preferential Return (EPR) model (Song et al., 2010) and its extensions (Barbosa et al., 2018), which typically assume that individuals make movement decisions autonomously, driven by internal rules balancing exploration and return. However, in today's algorithmically mediated urban environments, these decisions are often shaped by external influences (most notably, recommender systems) which embed individuals within a human-AI feedback loop that alters the dynamics of mobility itself. As our results show, such feedback loop has the potential to reshape venue visitation patterns, with cascading effects on venue popularity and social co-location. These findings suggest that classical mobility models must be revised to explicitly incorporate algorithmic influence as a primary behavioural driver, enabling more realistic representations of human mobility in algorithmically guided cities.

7 Conclusion

In this article, we modelled the human–AI feedback loop in the context of next-venue recommendation to simulate its impact on the diversity of visits at the individual and collective level, and on the structure of individuals' co-location networks.

Our study has some limitations. First, our analysis of the feedback loop is based on simulation rather than empirical observation, due to the lack of access to real-time platform data such as recommendation outputs and user interactions. As a result, our analysis should be intended as a what-if analysis to explore hypothetical scenarios rather than providing empirical evidence. Regulatory initiatives such as the Digital Services Act will empower researchers to run empirical experiments directly on online digital platforms, enabling rigorous assessments of how recommender systems influence real-world urban dynamics. In the meantime, our work offers a set of testable conjetures—the key results in the grey boxes—that can guide future empirical investigations on these platforms.

Second, we model individual decision-making using simplified probabilistic rules, which do not fully capture the complexity of human behaviour. For example, we assume the same

probability of adopting the recommender system across all individuals, without explicitly accounting for individual differences in receptiveness to recommendations. Our assumption of the same adoption probability across all individuals is supported by the findings of Cinus et al. (2022), which suggest that the qualitative effects of recommender systems remain robust even when individual susceptibility to recommendations is heterogeneous. Nonetheless, incorporating user-level heterogeneity more explicitly, or designing simulations with persistent exposure conditions, would provide a valuable extension to better understand the nuanced interplay between algorithmic influence and venue visitation patterns. We plan to address the limitations discussed above in future works.

Our study opens several directions for future research. One promising avenue is to enrich the autonomous decision-making mechanism by integrating more sophisticated mechanistic models of human mobility, capturing factors such as temporal routines (Jiang et al., 2016; Pappalardo & Simini, 2018), social interactions (Toole et al., 2015; Cornacchia & Pappalardo, 2021), and spatial scales (Alessandretti et al., 2020). In this regard, an interesting direction is to model the decision-making process as a two-step sequence: first selecting an activity category (e.g., dining, shopping, entertainment), and then choosing a specific venue within that category. This would allow us to capture how recommender systems influence not just where individuals go, but also what types of activities they choose to engage in. Another possible improvement would be to account for commercial mechanisms typically implemented in online platforms, such as venue sponsorship or monetized rankings. These mechanisms can alter the visibility of certain venues and may therefore influence the dynamics observed in this study.

Another interesting direction is the mitigation of the observed biases in location-based recommender systems. Model ensembling offers a promising approach in this regard. While our study does not explicitly evaluate ensemble methods, it includes PGN (Sánchez & Dietz, 2022), an ensemble model that integrates collaborative, geographical-proximity, and popularity signals. PGN displays behavioural patterns similar to deep learning models like MultiVAE, suggesting that even architectures integrating diverse signals may still reproduce certain biases. This underscores the need for future research to systematically assess how ensembling and signal integration impact fairness and diversity in human travel patterns.

Concerning the field of human mobility, a direction for future work is to investigate how the human-AI feedback loop may influence the well-established statistical laws that characterise human movement by guiding individuals towards certain venues over others. It would be also important to explore whether the feedback loop reinforces or mitigates the socio-economic and demographic segregation that often emerges in urban visitation patterns (Moro et al., 2021; Barbosa et al., 2021; Gambetta et al., 2023), potentially amplifying inequalities or, conversely, acting as tools for urban inclusion. By shaping venue popularity, location-based recommender systems can also affect property values and the perceived attractiveness of neighbourhoods, potentially driving complex urban processes such as gentrification and overtourism (Mauro et al., 2025).

In conclusion, our study represents the first attempt to study how the human-AI feedback loop reshapes human mobility patterns. Our novel computational framework opens new avenues for studying the coevolution between algorithmic decision support and individual choice in the city, contributing to a deeper understanding of how digital platforms shape, and are shaped by, human mobility patterns.

Table 2 Performance metrics for various recommender models on New York City dataset

Model	HitRate@20	mRR@20
UserKNN	0.1726	0.0576
ItemKNN	0.1703	0.0377
MF	0.1870	0.0474
BPRMF	0.2261	0.0611
MultiVAE	0.1898	0.0656
LightGCN	0.2774	0.0838
PGN	0.2156	0.0704

Table 3 Complete set of simulation parameters adopted in our framework

η	Adoption rate	$\{0.0, 0.2, 0.4, 0.6, 0.8, 1.0\}$
Δ	Retraining frequency	{1 days, 7 days}
p	Exploration probability	$0.6 \times \lvert V \rvert^{-0.21}$ (Song et al., 2010)
	Training approach	{incremental, sliding window}
	Optimizer	Adam
	Learning rate	0.001
	Epochs	500
	Early-stopping patience	5
	Batch size	{16, 32}
	nr of nearest neighbors	10
	KL divergence coeff (MultiVAE)	0.2

Appendix A: Training process and hyperparameters

UserKNN, ItemKNN, and PGN are trained using $k = 10$ nearest neighbors. MF is implemented with 32 latent factors. BPRMF and LightGCN are optimized using the pairwise BPR loss with a batch size of 16 and an L2 regularization term of 0.0001. MultiVAE is trained with cross-entropy loss combined with a KL divergence term (weighted by a coefficient of 0.2) in a full-batch setting.

During the training process, we track the training loss at the end of each epoch and apply early stopping if the loss remains stable for five consecutive epochs. The total number of epochs is capped at 500. We use the Adam optimiser with a learning rate of 0.001. The hyperparameter values we use in our experiments are reported in Table 3.

As a sanity check, we evaluate each trained model on D_{train} and D_{post}. For each user $u \in U$, we generate a top-K list of recommended venues and measure the HitRate and mean Reciprocal Rank (mRR) in predicting u's interactions in D_{post}. We set $K = 20$ and report average HitRate and mRR across all users in Table 2.

The performance of the recommender systems aligns with findings reported in the literature (Sánchez & Dietz, 2022): PGN outperforms BPRMF, which in turn exceeds UserKNN, followed by ItemKNN.

Appendix B: Tokyo dataset

The Foursquare Tokyo dataset contains 573,703 check-ins from 2293 users across 61,858 venues. This is approximately twice the size of the New York City dataset, while the geographic area covered by Tokyo –measured as the bounding box of all venues—is roughly half that of New York City.

To ensure a fair comparison between the two datasets, we randomly sample 1,083 users—the same number as in the raw New York City dataset—and apply the same preprocessing pipeline to both datasets. The resulting Tokyo dataset contains 141,968 check-ins to 34,314 venues by 1,083 users.

Figure 8 shows the impact of the feedback loop in Tokyo as a function of the adoption rate η, averaged over five simulation runs. The results closely mirror those observed for New York City, supporting the robustness of the main findings across different urban contexts.

Figure 8a shows that all models reduce the average individual Gini coefficient $\overline{G}$ with increasing η, indicating that recommenders consistently promote more diverse individual behavior. MultiVAE again achieves the largest reduction, reaching values below $\overline{G} = 0.1$ at full adoption, while the impact of KNN-based models is limited. Figure 8b reports the collective Gini coefficient G, which increases for most models as adoption rises. MultiVAE

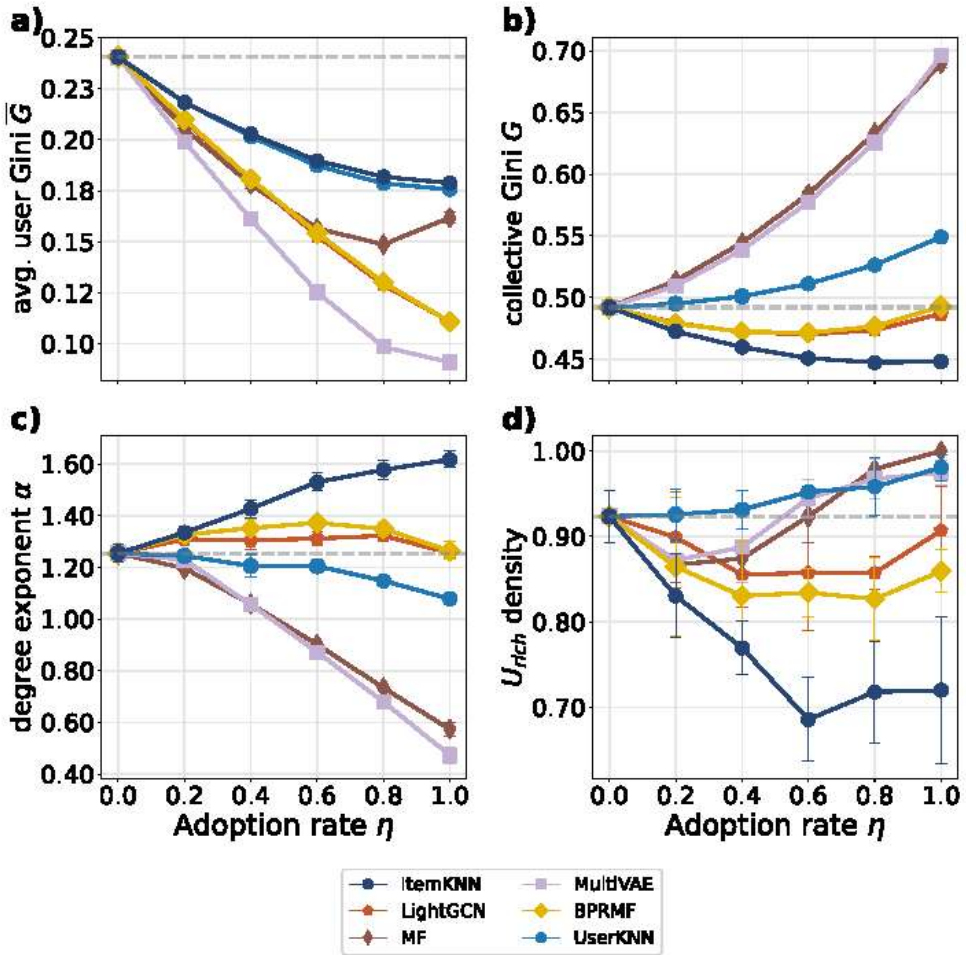


Fig. 8 Effects of different recommender systems on mobility and network-level indicators in Tokyo, as a function of the adoption rate η. Each point represents an average over five simulation runs. **a** average individual Gini coefficient $\overline{G}$; **b** collective Gini coefficient G; **c** degree exponent α of the co-location network; **d** rich-club density U_{rich}

and MF lead to the strongest increases, similarly to what was observed in New York City, reaching values close to $G = 0.7$. ItemKNN remains the only recommender system that consistently reduces collective inequality, similarly to what observed for New York City.

The degree exponent α of the co-location network is shown in Fig. 8c. As observed in New York City, most recommender systems flatten the degree distribution by lowering α, promoting broader co-location across users. The effect is the strongest for MF, MultiVAE, and LightGCN, while ItemKNN and UserKNN exhibit more stable values near the baseline. Figure 8d presents the rich-club density U_{rich}, defined as the internal density among the 15 most connected individuals. At first glance, results differ slightly from those in New York City. In Tokyo, the initial rich-club density is already high (around 0.9) at $\eta = 0$, compared to about 0.2 in New York City. As a result, for most recommender systems, the feedback loop either maintains or slightly reduces U_{rich}, with the exception of MF, MultiVAE, and UserKNN, which preserve or further increase rich-club connectivity. These differences may be explained by the higher population density, more compact spatial distribution of venues, or contextual behavioral patterns specific to Tokyo. Taken together, these results demonstrate that while some quantitative differences emerge—particularly in the rich-club structure—the main trends observed in New York City hold also in Tokyo. This reinforces the generalizability of our findings across different urban environments.

Appendix C: Dataset cleaning procedure

We excluded all ckeck-ins to venues with following categories: Train, Transport Hub, Transportation Service, Travel and Transportation, Boat or Ferry, Platform, Road, Island, River, Housing Development, Meeting Room, Conference Room, Office, Home (private), Apartment or Condo and Unknown.

Appendix D: Impact of retraining frequency

To assess the impact of the retraining frequency of recommender systems, we shorten the retraining interval of the simulation to $\Delta = 1\ day$, i.e., we update the recommender system daily. As shown in Fig. 9, the results for $\Delta = 1$ day are consistent with those for $\Delta = 7$ days, although the effects are more pronounced in the latter case. Our interpretation is that, because the simulation follows the real, longitudinal evolution of user–venue interactions, longer gaps between updates give the algorithm more time to amplify historical biases.

We also replicate our experiments using a sliding-window training scheme instead of an incremental one. In this setting, each new batch of simulated user–venue interactions replaces an equal number of the oldest interactions, keeping the training set size constant over time.

As Fig. 10 shows, the two schemas yield nearly identical outcomes: the curves mostly overlap and their error bars largely coincide, indicating no statistically significant differences. These findings suggest that our results are robust to the choice of training strategy, holding regardless of whether the recommender system is trained incrementally or using a sliding-window approach.

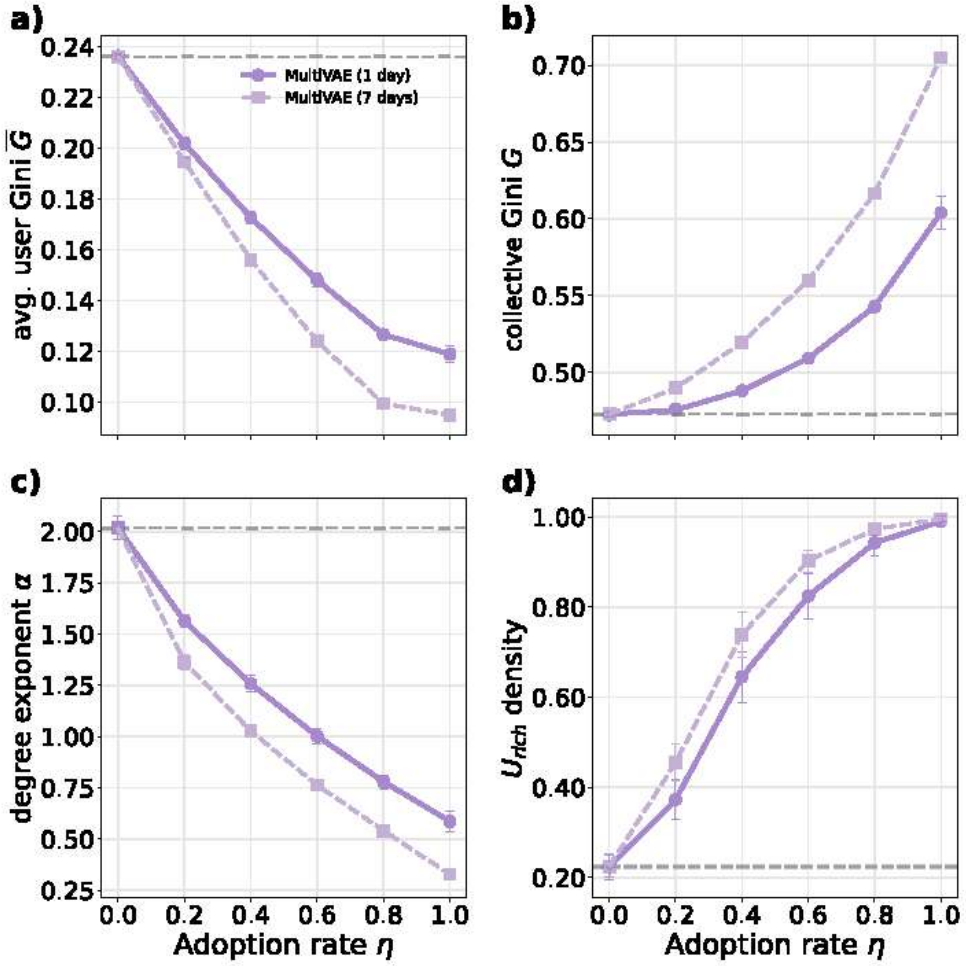


Fig. 9 Results obtained by comparing two different retraining frequencies: $\Delta = 1\ day$ and $\Delta = 7\ days$. No substantial differences are observed. The simulation uses MultiVAE as the recommender algorithm and the New York City dataset

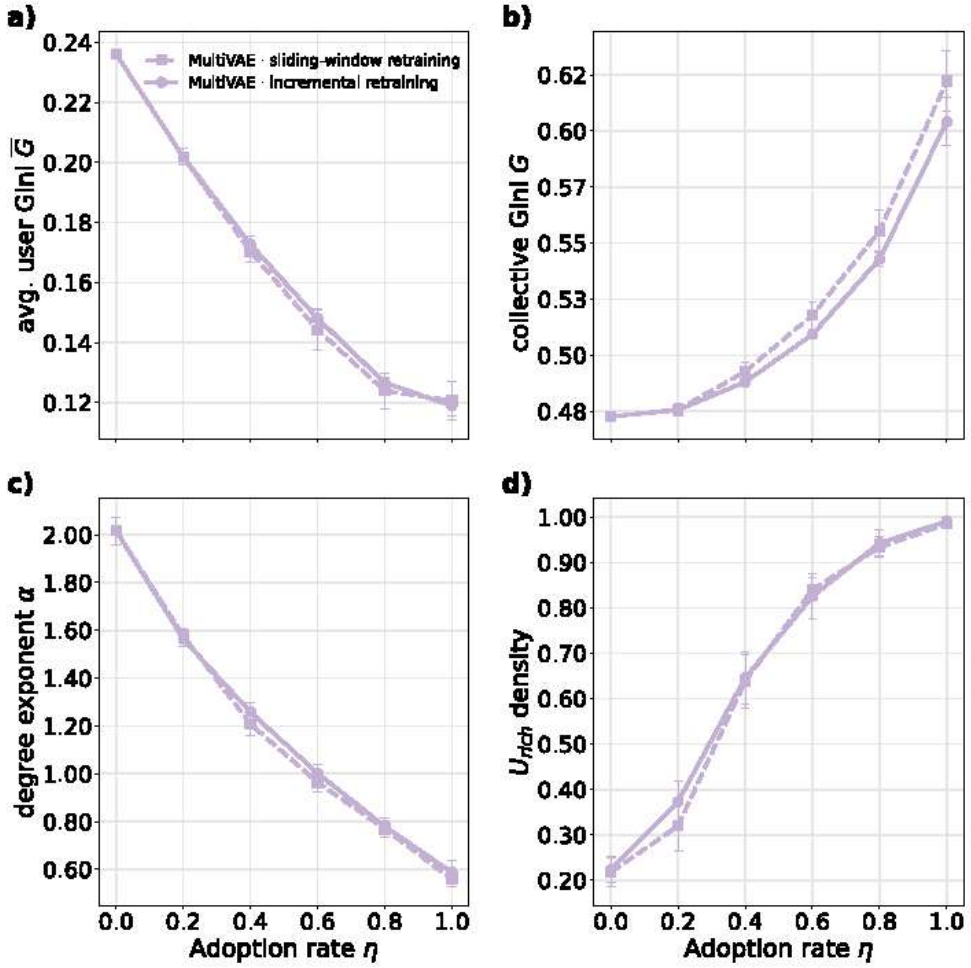


Fig. 10 Results obtained by comparing two training strategies: incremental training (used in the main text) and sliding-window training. No substantial differences are observed. The simulation uses MultiVAE as the recommender algorithm, the New York dataset, and a retraining interval of $\Delta = 1\ day$

Appendix E: Results for PGN

Similarly to MF and MultiVAE, PGN leads to a notable reduction in the individual Gini coefficient, accompanied by a significant increase in the collective Gini coefficient (see Fig. 11). We also observe a substantial decrease in the degree exponent α of the co-location network, suggesting a shift toward more homogenous connectivity, alongside the emergence of a denser rich-club structure.

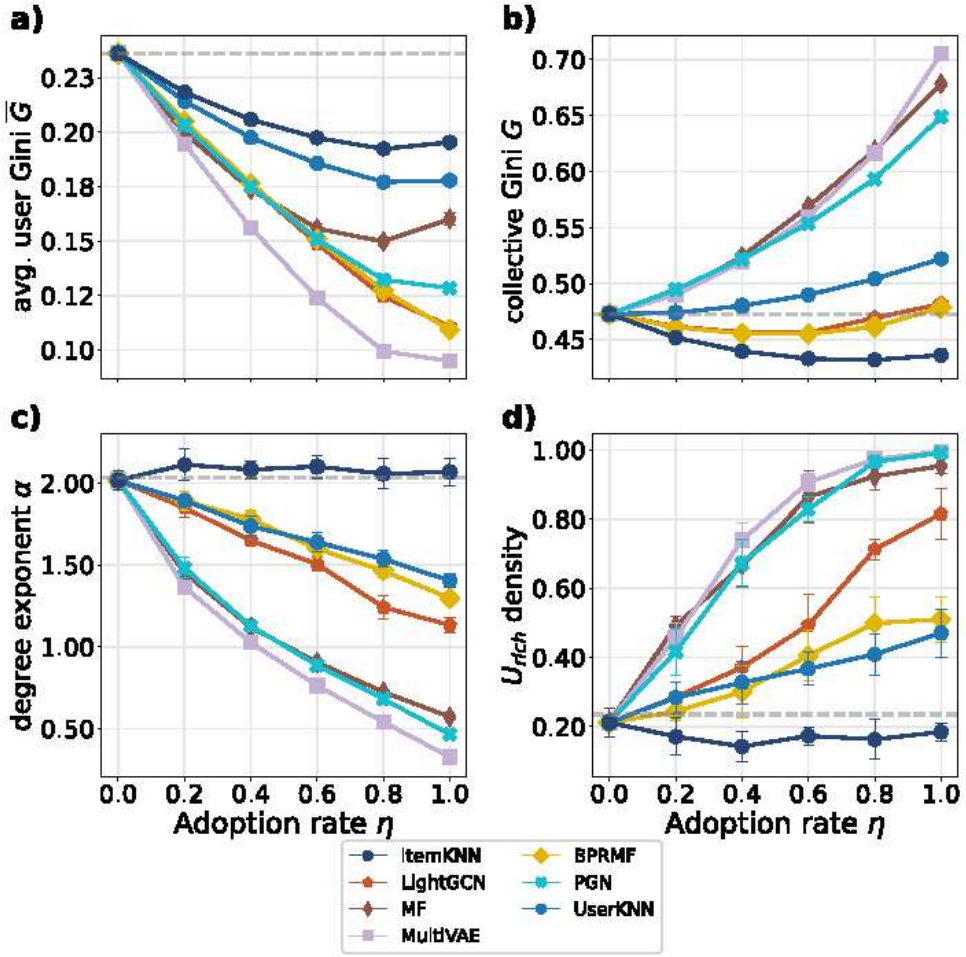


Fig. 11 Effects of recommender system adoption η considering also PGN

Appendix F: User similarity across adoption rates

We study how user similarity evolves as a function of the adoption rate η, offering further insight into the collective impact of different recommendation systems. We define Jensen-Shannon similarity (JS similarity) as:

$$\mathrm{JS}(P, Q) = 1 - \mathrm{JS}(P \| Q)$$

where $\mathrm{JS}(P\|Q)$ is the Jensen-Shannon divergence between two discrete probability distributions P and Q. The JS divergence (Fuglede & Topsoe, 2004) is defined as:

$$\mathrm{JS}(P\|Q) = \frac{1}{2}\mathrm{KL}(P\|M) + \frac{1}{2}\mathrm{KL}(Q\|M) \quad \text{where } M = \frac{1}{2}(P+Q)$$

where KL denotes the Kullback–Leibler divergence (Kullback, 1997; Van Erven & Harremos, 2014) defined as:

$$\mathrm{KL}(P\|Q) = \sum_{x \in X} P(x) \log\left(\frac{P(x)}{Q(x)}\right)$$

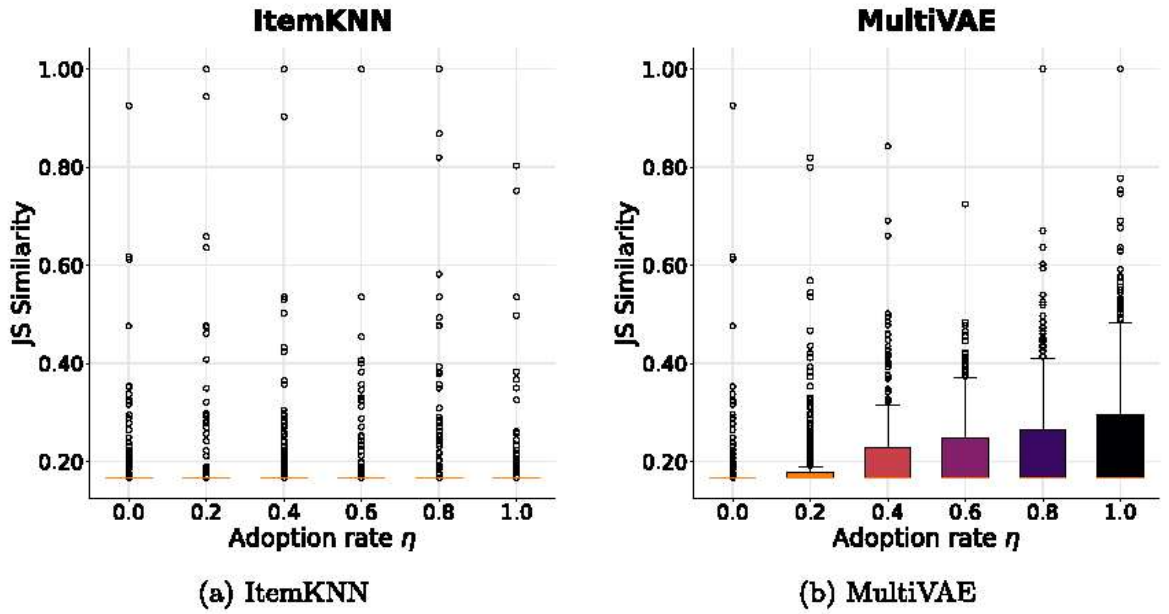


Fig. 12 Distribution of pairwise JS similarity among the top 50 users, computed over the 50 most visited venues, as a function of the adoption rate η. Under ItemKNN, similarity remains relatively stable. In contrast, under MultiVAE, JS similarity increases with adoption, indicating that visitation patterns become more aligned as the recommender is more widely adopted

$\mathrm{JS}(P, Q) \in [0, 1]$, it is symmetric and always well-defined. The higher JS, the more similar the two distributions are.

We select the top 50 most active users and the 50 most visited venues in New York City. For each user, we build a probability distribution over visits to these venues. We then compute the JS similarity between all user pairs at each value of the adoption rate η.

Figure 12 shows the Jensen-Shannon (JS) similarity as the adoption rate η increases for ItemKNN and MultiVAE. For MultiVAE, we observe a clear increase in JS similarity with rising η, indicating that higher reliance on the recommender system leads to more homogeneous visitation patterns. In contrast, ItemKNN does not exhibit this trend. These findings suggest that as users increasingly depend on MultiVAE for venue exploration, they become more similar to one another; they visit the same set of top venues.

Acknowledgements Giovanni Mauro has been supported by SoBigData.it—Strengthening the Italian RI for Social Mining and Big Data Analytics", prot. IR0000013, avviso n. 3264 on 28/12/2021. Marco Minici acknowledges partial support by (*i*) the SERICS project (PE00000014) under the NRRP MUR program funded by the EU—NGEU and (*ii*) "SoBig-Data.it—Strengthening the Italian RI for Social Mining and Big Data Analytics"—Prot. IR0000013, avviso n. 3264 on 28/12/2021. Luca Pappalardo has been supported by PNRR (Piano Nazionale di Ripresa e Resilienza) in the context of the research program 20224CZ5X4_PE6_PRIN 2022 "URBAI—Urban Artificial Intelligence" (CUP B53D23012770006), funded by European Union—Next Generation EU. We thank Daniele Fadda for the invaluable help with the data visualisations. We thank Andrea Frasson for his master's thesis work, which inspired this study. We thank Dino Pedreschi, Giuliano Cornacchia, Gabriele Barlacchi, Daniele Gambetta and Margherita Lalli for the useful discussions.

Author contributions GM and MM implemented the code for simulations and the data analysis. GM, MM, and LP conceptualised the work and designed the figures. GM made the plots. GM, MM and LP wrote the paper. LP directed and supervised the research. All authors contributed to the scientific discussion, read and approved the paper.

Funding Open access funding provided by Scuola Normale Superiore within the CRUI-CARE Agreement.

Data availability Data is provided within the manuscript or supplementary information files.

Declarations

Conflict of interest The authors declare no Conflict of interest.

Reproducibility and Code Our framework and analysis are fully reproducible, with all code available at https://github.com/mauruscz/UrbanFeedbackLoop. The simulations were executed on high-performance computing platforms with 64 CPU cores (AMD EPYC 7313 16-Core Processor) and ∼1.18 TB RAM. While the simulations are computationally intensive, we implemented a highly parallelized and modular architecture to optimize performance and resource utilization. Experiments with neural recommenders, i.e., MultiVAE, LightGCN, BPRMF, have been conducted on a DGX Server equipped with 4 NVIDIA Tesla V100 GPU (32GB) and CUDA Version 12.2.

References

Afèche, P., Liu, Z., & Maglaras, C. (2023). Ride-hailing networks with strategic drivers: The impact of platform control capabilities on performance. *Manufacturing and Service Operations Management, 25*(5), 1890–1908.

Aiolli, F. (2013). Efficient top-n recommendation for very large scale binary rated datasets. In *Proceedings of the 7th ACM conference on Recommender systems* (pp. 273–280).

Alessandretti, L., Aslak, U., & Lehmann, S. (2020). The scales of human mobility. *Nature, 587*(7834), 402–407.

Arora, N., Cabannes, T., Ganapathy, S., Li, Y., McAfee, P., Nunkesser, M., Osorio, C., Tomkins, A., & Tsogsuren, I. (2021). Quantifying the sustainability impact of google maps: A case study of salt lake city. CoRR arXiv:2111.03426

Barbosa, H., Barthelemy, M., Ghoshal, G., James, C. R., Lenormand, M., Louail, T., Menezes, R., Ramasco, J. J., Simini, F., & Tomasini, M. (2018). Human mobility: Models and applications. *Physics Reports, 734*, 1–74.

Barbosa, H., Hazarie, S., Dickinson, B., Bassolas, A., Frank, A., Kautz, H., Sadilek, A., Ramasco, J. J., & Ghoshal, G. (2021). Uncovering the socioeconomic facets of human mobility. *Scientific Reports, 11*(1), 8616.

Berahmand, K., Samadi, N., & Sheikholeslami, S. M. (2018). Effect of rich-club on diffusion in complex networks. *International Journal of Modern Physics B, 32*(12), 1850142.

Bertagnolli, G., & De Domenico, M. (2022). Functional rich clubs emerging from the diffusion geometry of complex networks. *Physical Review Research, 4*(3), Article 033185.

Bokányi, E., & Hannák, A. (2020). Understanding inequalities in ride-hailing services through simulations. *Scientific Reports, 10*(1), 6500.

Boratto, L., Fenu, G., & Marras, M. (2021). Connecting user and item perspectives in popularity debiasing for collaborative recommendation. *Information Processing & Management, 58*(1), Article 102387.

Chaney, A. J., Stewart, B. M., & Engelhardt, B. E. (2018). How algorithmic confounding in recommendation systems increases homogeneity and decreases utility. In *Proceedings of the 12th ACM conference on recommender systems*, (pp 224–232).

Chen, J., Dong, H., Wang, X., Feng, F., Wang, M., & He, X. (2023). Bias and debias in recommender system: A survey and future directions. *ACM Transactions on Information Systems, 41*(3), 1–39.

Cinus, F., Minici, M., Monti, C., & Bonchi, F. (2022). The effect of people recommenders on echo chambers and polarization. In *Proceedings of the international AAAI conference on web and social media* (pp. 90–101)

Colizza, V., Flammini, A., Serrano, M. A., & Vespignani, A. (2006). Detecting rich-club ordering in complex networks. *Nature Physics, 2*(2), 110–115.

Coppolillo, E., Mungari, S., Ritacco, E., Fabbri, F., Minici, M., Bonchi, F., & Manco, G. (2025). Algorithmic drift: A simulation framework to study the effects of recommender systems on user preferences. *Information Processing & Management, 62*(4), Article 104125.

Cornacchia, G., & Pappalardo, L. (2021). A mechanistic data-driven approach to synthesize human mobility considering the spatial, temporal, and social dimensions together. *ISPRS International Journal of Geo-Information, 10*(9), 599.

Cornacchia, G., Böhm, M., Mauro, G., Nanni, M., Pedreschi, D., & Pappalardo, L. (2022). How routing strategies impact urban emissions. In *Proceedings of the 30th international conference on advances in geographic information systems* (pp. 1–4).

Cornacchia, G., Nanni, M., & Pappalardo, L. (2023). One-shot traffic assignment with forward-looking penalization. In *SIGSPATIAL/GIS*. ACM (pp. 87:1–87:10).

Cornacchia, G., Lemma, L., & Pappalardo, L. (2024). Alternative routing based on road popularity. In *Proceedings of the 2nd ACM SIGSPATIAL workshop on sustainable urban mobility*. Association for Computing Machinery, New York, NY, USA, SUMob '24 (pp. 14–17). https://doi.org/10.1145/3681779.3696836

Doshi, A. R., & Hauser, O. P. (2024). Generative AI enhances individual creativity but reduces the collective diversity of novel content. *Science Advances, 10*(28), eadn5290.

Edelman, B. G., & Luca, M. (2014). Digital discrimination: The case of airbnb. com. *Harvard Business School NOM Unit Working Paper (14-054)*.

Elahi, M., Kholgh, D. K., Kiarostami, M. S., Saghari, S., Rad, S. P., & Tkalčič, M. (2021). Investigating the impact of recommender systems on user-based and item-based popularity bias. *Information Processing & Management, 58*(5), Article 102655.

Ensign, D., Friedler, S. A., Neville, S., Scheidegger, C., & Venkatasubramanian, S. (2018). Runaway feedback loops in predictive policing. In: Friedler, S.A., Wilson, C. (eds.) Proceedings of the 1st conference on fairness, accountability and transparency, proceedings of machine learning research, vol. 81. PMLR (pp. 160–171). https://proceedings.mlr.press/v81/ensign18a.html

Fleder, D., & Hosanagar, K. (2009). Blockbuster culture's next rise or fall: The impact of recommender systems on sales diversity. *Management Science, 55*(5), 697–712.

Fuglede, B., & Topsoe, F. (2004). Jensen-Shannon divergence and Hilbert space embedding. In *International symposium on information theory*, 2004. ISIT 2004. Proceedings., IEEE (pp. 31)

Gambetta, D., Mauro, G., & Pappalardo, L. (2023). Mobility constraints in segregation models. *Scientific Reports, 13*(1), 12087.

Gonzalez, M. C., Hidalgo, C. A., & Barabasi, A. L. (2008). Understanding individual human mobility patterns. *Nature, 453*(7196), 779–782.

He, X., Deng, K., Wang, X., Li, Y., Zhang, Y., & Wang, M. (2020). Lightgcn: Simplifying and powering graph convolution network for recommendation. In *Proceedings of the 43rd international ACM SIGIR conference on research and development in Information Retrieval* (pp. 639–648).

Jia, T., Cai, C., Li, X., Luo, X., Zhang, Y., & Yu, X. (2022). Dynamical community detection and spatiotemporal analysis in multilayer spatial interaction networks using trajectory data. *International Journal of Geographical Information Science, 36*(9), 1719–1740.

Jiang, R., Chiappa, S., Lattimore, T., György, A., & Kohli, P. (2019). Degenerate feedback loops in recommender systems. In *Proceedings of the 2019 AAAI/ACM conference on AI, ethics, and society*. Association for Computing Machinery, New York, NY, USA, AIES '19 (pp. 383–390). https://doi.org/10.1145/3306618.3314288

Jiang, S., Yang, Y., Gupta, S., Veneziano, D., Athavale, S., & González, M. C. (2016). The timegeo modeling framework for urban mobility without travel surveys. *Proceedings of the National Academy of Sciences, 113*(37), E5370–E5378.

Kang, C., Jiang, Z., & Liu, Y. (2022). Measuring hub locations in time-evolving spatial interaction networks based on explicit spatiotemporal coupling and group centrality. *International Journal of Geographical Information Science, 36*(2), 360–381.

Klimashevskaia, A., Jannach, D., Elahi, M., & Trattner, C. (2024). A survey on popularity bias in recommender systems. *User Modeling and User-Adapted Interaction, 34*(5), 1777–1834.

Koren, Y., Rendle, S., & Bell, R. (2021). Advances in collaborative filtering. *Recommender systems handbook* (pp. 91–142).

Kullback, S. (1997). Information theory and statistics.

Lee, D., & Hosanagar, K. (2019). How do recommender systems affect sales diversity? A cross-category investigation via randomized field experiment. *Information Systems Research, 30*(1), 239–259.

Lesota, O., Melchiorre, A., Rekabsaz, N., Brandl, S., Kowald, D., Lex, E., & Schedl, M. (2021). Analyzing item popularity bias of music recommender systems: Are different genders equally affected? In *Proceedings of the 15th ACM conference on recommender systems* (pp. 601–606).

Lesota, O., Geiger, J., Walder, M., Kowald, D., & Schedl, M. (2024). Oh, behave! country representation dynamics created by feedback loops in music recommender systems. In *Proceedings of the 18th ACM conference on recommender systems*. Association for Computing Machinery, New York, NY, USA, RecSys '24 (pp. 1022–1027).

Liang, D., Krishnan, R. G., Hoffman, M. D., & Jebara, T. (2018). Variational autoencoders for collaborative filtering. In *Proceedings of the 2018 world wide web conference* (pp. 689–698).

Luca, M., Barlacchi, G., Lepri, B., & Pappalardo, L. (2021). A survey on deep learning for human mobility. *ACM Computing Surveys (CSUR), 55*(1), 1–44.

Luca, M., Pappalardo, L., Lepri, B., & Barlacchi, G. (2023). Trajectory test-train overlap in next-location prediction datasets. *Machine Learning, 112*(11), 4597–4634.

Mansoury, M., Mobasher, B., Burke, R., & Pechenizkiy, M. (2019). Bias disparity in collaborative recommendation: Algorithmic evaluation and comparison.

Mansoury, M., Abdollahpouri, H., Pechenizkiy, M., Mobasher, B., & Burke, R. (2020). Feedback loop and bias amplification in recommender systems. In *Proceedings of the 29th ACM international conference on information & knowledge management*. Association for Computing Machinery, New York, NY, USA, CIKM '20, (pp 2145–2148), https://doi.org/10.1145/3340531.3412152,

Mauro, G., Pedreschi, N., Lambiotte, R., & Pappalardo, L. (2025). Dynamic models of gentrification. *Advances in Complex Systems*. https://doi.org/10.1142/S0219525925400065

Moro, E., Calacci, D., Dong, X., & Pentland, A. (2021). Mobility patterns are associated with experienced income segregation in large us cities. *Nature Communications, 12*(1), 4633.

Nguyen, T. T., Hui, P. M., Harper, F. M., Terveen, L., & Konstan, J. A. (2014). Exploring the filter bubble: The effect of using recommender systems on content diversity. In *Proceedings of the 23rd international conference on world wide web*. Association for Computing Machinery, New York, NY, USA, WWW '14 (pp. 677–686). https://doi.org/10.1145/2566486.2568012,

Pappalardo, L., & Simini, F. (2018). Data-driven generation of spatio-temporal routines in human mobility. *Data Mining and Knowledge Discovery, 32*(3), 787–829.

Pappalardo, L., Simini, F., Rinzivillo, S., Pedreschi, D., Giannotti, F., & Barabási, A.-L. (2015). Returners and explorers dichotomy in human mobility. *Nature Communications, 6*(1), 8166.

Pappalardo, L., Rinzivillo, S., & Simini, F. (2016). Human mobility modelling: Exploration and preferential return meet the gravity model. *Procedia Computer Science, 83*, 934–939.

Pappalardo, L., Manley, E., Sekara, V., & Alessandretti, L. (2023). Future directions in human mobility science. *Nature Computational Science, 3*(7), 588–600.

Pappalardo, L., Ferragina, E., Citraro, S., Cornacchia, G., Nanni, M., Rossetti, G., Gezici, G., Giannotti, F., Lalli, M., Gambetta, D. R., Mauro, G., Morini, V., Pansanella, V., & Pedreschi, D. (2024). A survey on the impact of AI-based recommenders on human behaviours: Methodologies, outcomes and future directions. arXiv preprint https://arxiv.org/abs/2407.01630

Pedreschi, D., Pappalardo, L., Ferragina, E., Baeza-Yates, R., Barabási, A.-L., Dignum, F., Dignum, V., Eliassi-Rad, T., Giannotti, F., Kertész, J., Knott, A., Ioannidis, Y. E., Lukowicz, P., Passarella, A., Pentland, A., Shawe-Taylor, J., & Vespignani, A. (2025). Human-AI coevolution. *Artificial Intelligence, 339*, Article 104244.

Pedreschi, N., Battaglia, D., & Barrat, A. (2022). The temporal rich club phenomenon. *Nature Physics, 18*(8), 931–938.

Perez-Prada, F., Monzón, A., & Valdés, C. (2017). Managing traffic flows for cleaner cities: The role of green navigation systems. *Energies, 10*, 791. https://doi.org/10.3390/en10060791

Piao, J., Liu, J., Zhang, F., Su, J., & Li, Y. (2023). Human-AI adaptive dynamics drives the emergence of information cocoons. *Nature Machine Intelligence, 5*(11), 1214–1224.

Quijano-Sánchez, L., Cantador, I., Cortés-Cediel, M. E., & Gil, O. (2020). Recommender systems for smart cities. *Information Systems, 92*, Article 101545.

Rahmani, H. A., Deldjoo, Y., & Di Noia, T. (2022). The role of context fusion on accuracy, beyond-accuracy, and fairness of point-of-interest recommendation systems. *Expert Systems with Applications, 205*, Article 117700.

Rahmani, H. A., Deldjoo, Y., Tourani, A., & Naghiaei, M. (2022b). The unfairness of active users and popularity bias in point-of-interest recommendation. In *International workshop on algorithmic bias in search and recommendation*, Springer (pp. 56–68)

Rendle, S., Freudenthaler, C., Gantner, Z., & Schmidt-Thieme, L. (2009). Bpr: Bayesian personalized ranking from implicit feedback. In *Proceedings of the twenty-fifth conference on uncertainty in artificial intelligence* (pp. 452–461).

Ricci, F., Rokach, L., & Shapira, B. (2021a) Recommender systems: Techniques, applications, and challenges. *Recommender systems handbook* pp 1–35

Ricci, F., Rokach, L., & Shapira, B. (2021b). Recommender systems: Techniques, applications, and challenges. *Recommender systems handbook* pp 1–35

Sánchez, P., & Dietz, L. W. (2022). Travelers vs. locals: The effect of cluster analysis in point-of-interest recommendation. In *Proceedings of the 30th ACM conference on user modeling, adaptation and personalization* (pp. 132–142).

Sarwar, B., Karypis, G., Konstan, J., & Riedl, J. (2000). Application of dimensionality reduction in recommender system-a case study. In *ACM WebKDD workshop*, Citeseer (pp. 285–295).

Shumailov, I., Shumaylov, Z., Zhao, Y., Papernot, N., Anderson, R., & Gal, Y. (2024). Ai models collapse when trained on recursively generated data. *Nature, 631*(8022), 755–759.

Song, C., Koren, T., Wang, P., & Barabási, A.-L. (2010). Modelling the scaling properties of human mobility. *Nature Physics, 6*(10), 818–823.

Subramonian, A., Kang, J., & Sun, Y. (2024). Theoretical and empirical insights into the origins of degree bias in graph neural networks. *Advances in Neural Information Processing Systems, 37*, 8193–8239.

Sun, W., Khenissi, S., Nasraoui, O., & Shafto, P. (2019). Debiasing the human-recommender system feedback loop in collaborative filtering. In *Companion Proceedings of The 2019 World Wide Web Conference. Association for Computing Machinery*, New York, NY, USA, WWW '19, p 645–651

Toole, J. L., Herrera-Yaqüe, C., Schneider, C. M., & González, M. C. (2015). Coupling human mobility and social ties. *Journal of The Royal Society Interface, 12*(105), 20141128.

Erven, T., & Harremos, P. (2014). Rényi divergence and Kullback-Leibler divergence. *IEEE Transactions on Information Theory, 60*(7), 3797–3820.

Wang, Y., Tao, L., & Zhang, X. X. (2025). Recommending for a multi-sided marketplace: A multi-objective hierarchical approach. *Marketing Science, 44*(1), 1–29.

Yang, D., Zhang, D., Zheng, V. W., & Yu, Z. (2014). Modeling user activity preference by leveraging user spatial temporal characteristics in LBSNs. *IEEE Transactions on Systems, Man, and Cybernetics: Systems, 45*(1), 129–142.

Zhang, S., Mehta, N., Singh, P. V., & Srinivasan, K. (2021). Frontiers: Can an artificial intelligence algorithm mitigate racial economic inequality? An analysis in the context of airbnb. *Marketing Science, 40*(5), 813–820.

Zhou, S., & Mondragón, R. J. (2004). The rich-club phenomenon in the internet topology. *IEEE Communications Letters, 8*(3), 180–182.

Zhou, Y., Dai, S., Pang, L., Wang, G., Dong, Z., Xu, J., & Wen, J.-R. (2024). Source echo chamber: Exploring the escalation of source bias in user, data, and recommender system feedback loop. arXiv:2405.17998