Mover: Generalizability Verification of Human Mobility Models via Heterogeneous Use Cases

WENJUN LYU, Rutgers University, United States
GUANG WANG, Rutgers University, United States
YU YANG, Lehigh University, United States
DESHENG ZHANG, Rutgers University, United States

Human mobility models typically produce mobility data to capture human mobility patterns individually or collectively based on real-world observations or assumptions, which are essential for many use cases in research and practice, e.g., mobile networking, autonomous driving, urban planning, and epidemic control. However, most existing mobility models suffer from practical issues like unknown accuracy and uncertain parameters in new use cases because they are normally designed and verified based on a particular use case (e.g., mobile phones, taxis, or mobile payments). This causes significant challenges for researchers when they try to select a representative human mobility model with appropriate parameters for new use cases. In this paper, we introduce a MObility VERification framework called MOVER to systematically measure the performance of a set of representative mobility models including both theoretical and empirical models based on a diverse set of use cases with various measures. Based on a taxonomy built upon spatial granularity and temporal continuity, we selected four representative mobility use cases (e.g., the vehicle tracking system, the camera-based system, the mobile payment system, and the cellular network system) to verify the generalizability of the state-of-the-art human mobility models. MOVER methodically characterizes the accuracy of five different mobility models in these four use cases based on a comprehensive set of mobility measures and provide two key lessons learned: (i) For the collective level measures, the finer spatial granularity of the user cases, the better generalization of the theoretical models; (ii) For the individual-level measures, the lower periodic temporal continuity of the user cases, the theoretical models typically generalize better than the empirical models. The verification results can help the research community to select appropriate mobility models and parameters in different use cases.

CCS Concepts: • **Human-centered computing** → **Ubiquitous and mobile computing**;

Additional Key Words and Phrases: Mobility modeling, Generalizability, Heterogeneous use cases

ACM Reference Format:
Wenjun Lyu, Guang Wang, Yu Yang, and Desheng Zhang. 2021. Mover: Generalizability Verification of Human Mobility Models via Heterogeneous Use Cases. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 5, 4, Article 171 (December 2021), 21 pages. https://doi.org/10.1145/3494997

1 INTRODUCTION

Human mobility models generate synthetic mobility data to capture collective human flows or individual spatiotemporal human locations, which are of great significance for researchers to conduct their research without real mobility data. Examples include mobile networking [5][70][13], location-based services [57][58][63][65], epidemic

Authors' addresses: Wenjun Lyu, Rutgers University, NJ, United States, wenjun.lyu@rutgers.edu; Guang Wang, Rutgers University, NJ, United States, guang.wang@rutgers.edu; Yu Yang, Lehigh University, Bethlehem, Pennsylvania, United States; Desheng Zhang, Rutgers University, New Jersey, United States, desheng.zhang@cs.rutgers.edu.

2474-9567/2021/12-ART171 $15.00
https://doi.org/10.1145/3494997

control [2][3][9][51], transportation management[23][24][57][60][66], and emergency management [19][29][44]. In particular, some concrete use cases include (1) modeling how large-scale vehicles move in real-time can help manage autonomous vehicles and provide large-scale simulations for Vehicular Edge Computing [30]; (2) human mobility models can provide guidance for base station deployment of ongoing 5G cellular networks; (3) human mobility models can help know how infectious virus spread in crowds, which is extremely important to control epidemics, such as the recent COVID-19 [25]. To enable the practical effectiveness of human mobility models in these cases, it is essential to understand the generalizability of different mobility models in different use cases, which can provide the best guidance for model selection.

In this work, our goal is to verify the generalizability of existing human mobility models in heterogeneous use cases. Existing human mobility models can be generally divided into two categories: *theoretical* mobility models and *empirical* mobility models. (i) Theoretical models such as Gravity Model [21][31], Random Way-point model [26], and Levy Walk [48] are generally based on scientific hypotheses or fitted by small-scale data from interviews and surveys [40]. (ii) *Empirical* mobility models such as [14][16][64][66] are driven by large-scale real-world data to improve the accuracy of mobility modeling. However, these models are generally only validated in some use cases while it is unclear how it performs when generalized to other use cases. The reason that we care about generalizability is, in practical usage, it is not always promising that we can find the exactly same data source as used in the existing works for human mobility modeling because of many practical constraints such as data ownership and data privacy. Instead, we may need to use different data sources as replacement to build the mobility models, which leads to unverified performance. For example, the performance of mobility models resulted from smartphone data is unclear when applied to vehicle use cases due to their different spatiotemporal mobility features. To the best of our knowledge, little work, if any, has comprehensively verified the performance of various mobility models with real-world data in multiple use cases with different spatial and temporal features to exam their generalizability because of the lack of data in multiple mobility use cases.

Benefited from the ubiquity of sensing devices and the decrease of data transmission and storage cost in recent years, various mobility data have been collected from different use cases at urban scale in real time recently, which provide an unprecedented opportunity to verify the generalizability of different mobility models with heterogeneous use cases at a city scale. For example, cellular network data [14][10], vehicle trace data [59][53][27][55][34][28], mobile payment data[64][66], and camera-based system data (e.g., traffic management [45], public safety [47]) have been gathered for various use cases in many big and developed cities, e.g., New York City, Beijing, Shanghai and Shenzhen.

In this work, we design a MObility VERification framework called `MOVER` to verify existing human mobility models in a systematic way. `MOVER` is based on city-scale mobility datasets covering different spatial and temporal granularity from four real-world systems including a vehicle tracking system, a mobile payment system, a camera-based system, and a cellular system deployed in Chinese city Shenzhen. The verification is conducted in 4 use cases with 5 representative mobility models based on 9 mobility measures involving more than 3.3 million users. The goal of our work is to provide a practical guidance for the research community in mobility model selection and parameter adjustment when conducting large-scale mobility-driven experiments and emulations, e.g., how autonomous vehicles move and share sensor data with others, deploying new base stations for 5G cellular networks in cities, and understanding the virus (e.g., COVID-19) spreading mechanism based on human mobility. We summarize the contributions of our work as follows.

- To the best of our knowledge, `MOVER` is the first work to verify the generalizability of the multiple mobility models for heterogeneous mobility use cases to produce practical guidance for model and parameter selection. Specifically, the verification involves two theoretical mobility models (Random Way-point and Levy walk) and three empirical mobility models (WHERE, Universal model, and Buscope) in four typical use cases with one-month city-scale data and different spatial-temporal granularity. It covers four categories

of use cases including mobile payment, telecommunication, transportation, and public safety. Therefore, our work has fairly complete data and model coverage and gives a comprehensive understanding of how mobility models generalize to different real-world use cases.
- MOVER systematically verifies mobility models' generalizability based on mobility measures at both collective mobility and individual mobility levels. At the collective level, we adopt four essential measures including trip counts, original Destination matrix, users counts, and location entropy to capture location features, which can be used to estimate mobility flows between location pairs for various use cases, e.g., 5G base station deployment and traffic management. At the individual level, we adopt five essential measures including average jump length, maximum jump length, distinct visited locations, entropy, and gyration to capture individual mobility features, which can be used for more fine-grained use cases such as large-scale emulation of autonomous cars or epidemic control.

Based on the verification results, we discuss two lessons learned about human mobility model selection and design implications for other use cases.

- **Lesson 1: Impact of Spatial Granularity in Use Cases.** For the collective level measures such as visit counts, origin-destination matrix, and location entropy, the finer spatial granularity of the use cases, the better generalizability of the theoretical models. This insight is counter-intuitive because normally the performance of a theoretical mobility model should be better with data of coarser spatial granularity compared to finer spatial granularity. In contrast, the performance of empirical models is not sensitive to spatial granularity of the use cases. This insight provides important guidance for model selection given use cases with spatial granularity when focusing on collective level measures, e.g., cellular user flows from one area to another area.
- **Lesson 2: Impact of Temporal Granularity in Use Cases.** For the individual level measures such as jump length, gyration, and distinct visited locations, the lower periodic temporal continuity of the use cases (i.e., the mobility data are logged with a lower frequency), the better generalizability of the theoretical models compared to empirical models. This insight provides guidance for model selection in use cases with lower periodic temporal continuity (such as mobile payment or phone call based mobility logs) when focusing on individual-level measures, e.g., maximum daily travel distance.

2 METHODOLOGY

In this section, we introduce the data from four use cases with different spatial granularity and temporal continuity for the generalizability verification first, followed by the description of five mobility models, e.g., two theoretical models and three empirical models. Then we illustrate the collective and individual level measures involved. Finally, we describe the metric we used for verify the performance of the models.

2.1 Four Mobility Use Cases

We built a taxonomy for mobility use cases from two dimensions, e.g., spatial granularity and temporal continuity as shown in Figure 1. For instance, the vehicle GPS data in the bottom left corner has the smallest spatial granularity and temporal continuity because the onboard GPS devices update the fine-grained longitude and latitude for vehicles about every 10s. In contrast, the camera-based system only collected data in the stationary locations, e.g., the highway toll stations (e.g., 70 in Shenzhen), which are sparsely distributed. The highway toll stations only collect the vehicle transaction data when the vehicles enter and left the highway, so the temporal continuity is coarse in this use case.

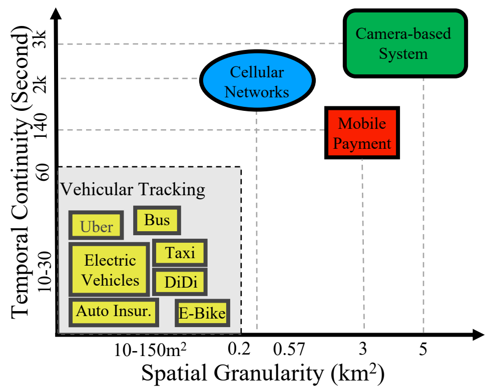


Fig. 1. Taxonomy for Different Mobility Use Cases

Table 1. Use Case Data Description

Description	Data Used for Generalizability Verification in Four Use Cases			
	Vehicle Tracking Systems	Camera-based Systems	Payment Systems	Cellular Network Systems
# of users	10K	0.8M	1.3M	1.2M
Daily data size	84.6M	201M	550M	1.44G
# of daily records	1.2M	2.4M	4M	14M
Format	Device ID, Date&time GPS, Speed	Plate ID, Date&time O/D Station, O/D Road	Station ID, Date&time Machine ID, In/Out	Card ID, Date&time GPS, Celltower ID

For this work, We choose four typical real-world use cases with different spatial-temporal features covering 4 urban domains (e.g., payment, telecommunication, transportation, and public safety). We have obtained one-month data in Shenzhen, collected by service providers and the Shenzhen Committee of Transportation (SCT). The details about data sets are shown in Table 1.

- **Use Case 1: the vehicle tracking system** collects the personal vehicle GPS trace data by an onboard GPS device. We divide the Shenzhen map into 500m × 500m grids, 1km × 1km grids, and 2km × 2km grids, and map the vehicle locations into different grids according to the longitude and latitude, which a state of practice approach for vehicular data processing[35]. As a result, if a vehicle stopped the GPS trace uploading, then we assume it stops and the driver visits the corresponding locations. The average daily individual hop (i.e., the number of distinct consecutive locations visited by an individual a day) is 2.48 after data preprocessing. The use cases for vehicle tracking systems include travel time estimation [12] and vehicular sensing task scheduling [59].
- **Use Case 2: the camera-based system** collects the coarse-grained information when a vehicle enters or leaves the highway. There are 70 highway toll stations in Shenzhen, and the individual daily hop is 2.18 on average. The use cases for the camera-based system include the real-time vehicle locations for anomaly detection and risk assessment in highways [64].
- **Use Case 3: the mobile payment system** collects passengers' trip origins and destinations when they tap in and tap out of the subway stations. The data sets contain 118 subway stations in total, and the average daily individual hop is 1.56. The use cases for the mobile payment systems include the subway scheduling [56] and travel time estimation [18].

- **Use Case 4: the cellular network system** collects the spatial-temporal information with the cell tower spatial granularity. Cellular data is collected when cellphone users are connected to nearby towers for phone calls or messages. There are more than 3600 cell towers involved in our data sets, and the daily number of hops is 6.9 on average for each individual. The use cases for the cellular network system include the cellular data usage prediction [41], urban population modeling [14], and transportation modes inference [69].

2.2 Five Mobility Models

In this paper, we choose two typically and widely used theoretical models, Random Way-point [26] and Levy walk [48] in mobile computing [52] [43]. For empirical models, we choose the representative models WHERE [16] and Universal Model [61], which considering the collective and individual level mobility features, respectively. We also select the advanced model Buscope [32] that considering both the collective and individual level mobility behavior. We list the notations and the meaning in this paper in Table 2 and the comparisons of the models are shown in Table 3.

Table 2. Comparison of Mobility Models

Category	Model	Factors considered				Math expression
		Population	Distance	Individual History	Collective History	
Theoretical models	Random Way-point [26]	×	×	×	×	–
	Levy Walk [48]	×	√	×	×	$P_{i,j} = \frac{1}{Z\|\|d_{i,j}\|\|^{\alpha}}$
Empirical models	WHERE [16]	×	×	×	√	–
	Universal Model [61]	√	×	√	×	$P_{i,j} \propto \frac{m_j}{s_{j,i}}(1 + \frac{\lambda}{rank_j})$
	Buscope [32]	×	×	√	√	–

Table 3. Notations used in this paper

Notation	Meaning
$s_{j,i}$	number of visits in a circle centered at location j and the radius is the distance between i and j
$P_{i,j}$	the probability a resident moves to j when she is now at location i
r_j	the attractiveness of location j, the smaller of r_j, the more attractive of j
$P_j(u)$	the probability that location j is visited by u
$P^u(j)$	the probability of j in u's probability vector
r_0^u	the center of mass of u's trip
r_i^u	the location vector of i-th record for u
E_j	the uncorrelated location entropy for j
E^u	the uncorrelated entropy for individual u
r_g^u	the radius of gyration for u
N_u	number of locations visited by u

- **Theoretical Random Way-point model:** In Random Way-point [26] model, a synthetic resident starts from an initial location and then randomly chooses the next destination from all nodes. When the user arrived at the destination and pause for a period of time, she would reselect the destination randomly and repeat the process. Random Way-point is a non-parameter model and without history information for generating traces for individuals.
- **Theoretical Levy walk model:** Existing studies have revealed that distance of human movement follows heavy-tailed distribution [15], so compared with Random Way-point, the probability for next location selection for a synthetic user is related to the distance of two location pairs with Levy walk [48]. The larger the distance to j, the lower probability to choose j as the destination. As shown in Table 3, α is a parameter that can affect the performance of the model.
- **Empirical WHERE model:** Random Way-point and Levy walk are theoretical models without taking history information into consideration. WHERE [16] is a statistical mobility model based on Call Detail Records(CDRs). The key insight is the majority of people move between some important locations, e.g., home and working places, so we can get five different probability distributions with the history collective information, e.g., home, distance, work, call time and per user calls per day, and generate the individual trace according to the collective information. As a complementary to WHERE, WHERE3 is proposed to add another important location for generating individual trace to improve the effectiveness and scalability of the model.
- **Empirical Universal model :** Compared with WHERE, which utilizes the collective history information, universal model [61] captures both individual and collective level patterns of human mobility, taking the number of people and the individual history trace into consideration. λ is a parameter for UM, and normally, the value of λ is smaller in a better-developed country or area, in this work, we choose 3 different values for λ [61].
- **Empirical BuScope:** BuScope also [32] combines individual and aggregated information to predict the next location for a user. But different from the universal model, BuScope only gives the individual history data higher priority. If a user starts from s with high support according to historical data, then the destination is the highest-confidence location from s. If the support is low, then the destination prediction is based on the probability of the aggregated flows from s.

For models that need the collective history information (e.g., WHERE) and individual history information (e.g., Universal model and BuScope), we use about three weeks of data as the historic records and the remaining data as the verification data to show the accuracy of the models.

2.3 Mobility Model Measures

We measure the accuracy of different mobility models from two perspectives, e.g., collective and individual level. The verification measures are also divided into two categories, e.g., collective and individual measures. We select four collective measures, which can capture the aggregated feature and has been widely used [17][37][8] for transportation network optimization and investments. We also select five typical individual measures used previous works [6][36][39] to capture the residents' move patterns, such as variability (e.g.,Radius of gyration), and predictability (e.g., Uncorrelated entropy).

2.3.1 Four Collective Measures. The collective measures measure the flows migrating between two locations, e.g., the cell towers in cellular networks and subway stations in AFC systems. The selected collective measures are given as follows.

- `Visit counts` [37] computes the number of daily visits $\sum_i T_{i,j}$ for each location j.

- `Origin Destination (OD) matrix` [7] computes the daily number of trips $T_{i,j}$ between arbitrary two locations i and j during the measurement period.
- `User counts` computes the daily number of distinct individuals for each location j.
- `Uncorrelated location entropy` [8] computes the temporal-unrelated location entropy of the locations and indicates the historic probability that the location is visited by each individual. The uncorrelated location entropy can be calculated as

$$E_j = -\sum_u P_j(u) \log P_j(u). \tag{1}$$

2.3.2 Five Individual Measures. Different from the collective measures focusing on aggregated features, individual measures can capture the mobility patterns from a personal perspective. We exploit five individual-level measures as follows.

- `Average jump length` [6] computes the average daily jump length (the distance between two consecutive locations that a resident visits) for each individual during the measurement period.
- `Maximum jump length` [29] computes the largest daily jump length for each individual.
- `Number of distinct visited locations` [15] computes the average daily number of distinct locations that an individual has been visited.
- `Uncorrelated entropy` [39] computes the temporal uncorrelated entropy of a set of individuals, and indicates the historic probability that a location was visited by an individual. The uncorrelated entropy is calculated as

$$E^u = -\sum_j P^u(j) \log P^u(j). \tag{2}$$

- `Radius of gyration` [36][15] computes the radii of gyration of a set of individuals and represents the characteristic distance travelled by an individual in mobility analysis. The radius if gyration of u is calculated as:

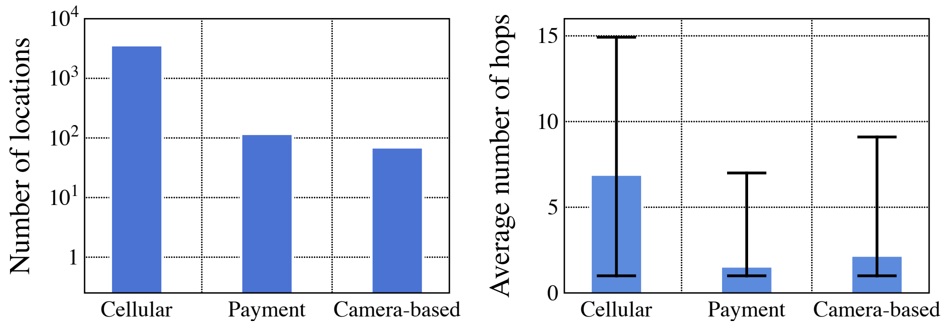
$$r_g^u = \sqrt{\frac{1}{N_u} \sum_{i=1}^{N_u} (r_i^u - r_0^u)^2}. \tag{3}$$

2.4 Verification Metric: MAPE

To verify the model generalizability accuracy, we use the real-world data as the ground truth. We first utilize the models to generate collective flows or individual traces and then compute the mobility measures in different level. We use the Mean Absolute Percent Error (MAPE) [11] as the accuracy evaluation for each measure because compared with it can measure the deviation degree of the predicted value to the true values. MAPE is defined as

$$MAPE = \frac{1}{n} \sum_{i=1}^{n} \frac{|\hat{y}_i - y_i|}{y_i}. \tag{4}$$

where $\hat{y}_i$ is the value of the measure calculated according to the model, and y_i is derived from the real data.



(a) Number of Locations

(b) Average Number of Hops

Fig. 2. Number of locations and average number of hops for each individual in three use cases. (a) Locations represents stations in camera-based and payment systems, and cell towers in cellular network system; (b) The average daily number of hops means the quantity of continuously distinct location pairs in each individual trace, e.g. A⟶ B ⟶ A with the number of 2, while A⟶ A ⟶ B with the number of 1.

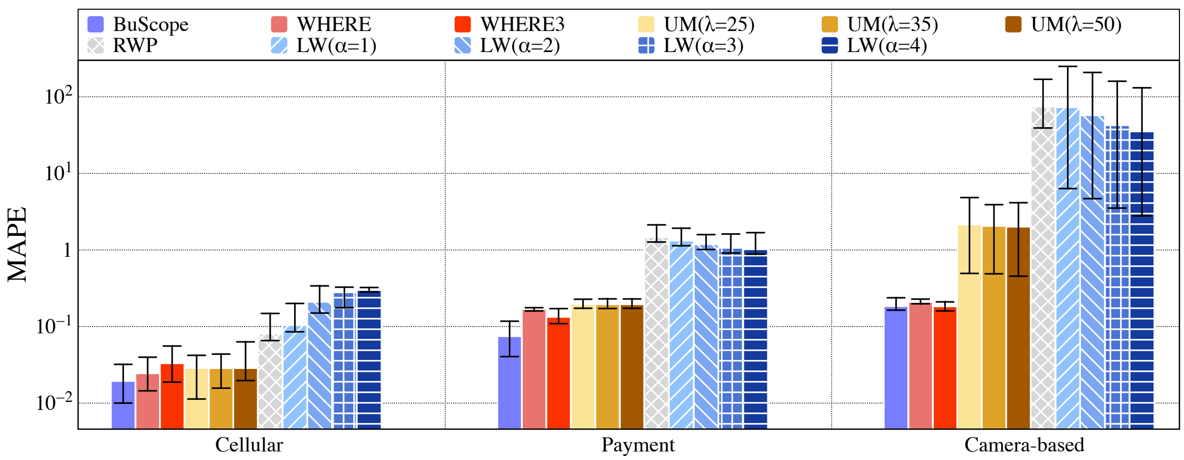


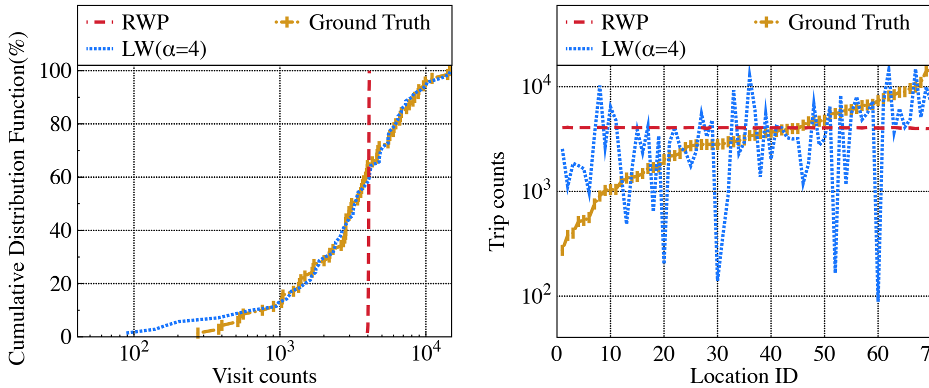
Fig. 3. Visit counts in collective level: Units from left to right in x-axis are with finer to coarser spatial granularity

3 VERIFICATION RESULTS

In this section, we present the verification results in collective and individual measures, the detailed analysis the some key lessons learned from the results. We show mobility spatial and temporal features in different use cases in Figure 2, which can be used for explaining most of the following results. We use UM, RWP, LW to represent the Universal model, Random Way-Point model, and Levy walk model in the figures shown in this section due to the space limitation.

3.1 Collective Level Mobility Measures

3.1.1 Visit Counts. **Visit Counts for Theoretical Models:** As shown in Figure 3, two theoretical models, i.e., Random Way-point (RWP) and Levy walk (LW), achieve the worst performance in the camera-based use case compared with the other two use cases because of the fewest number of locations in this use case (e.g., 70 stations



(a) CDF of visit counts

(b) Visit counts in different locations

Fig. 4. Visit counts in camera-based use case

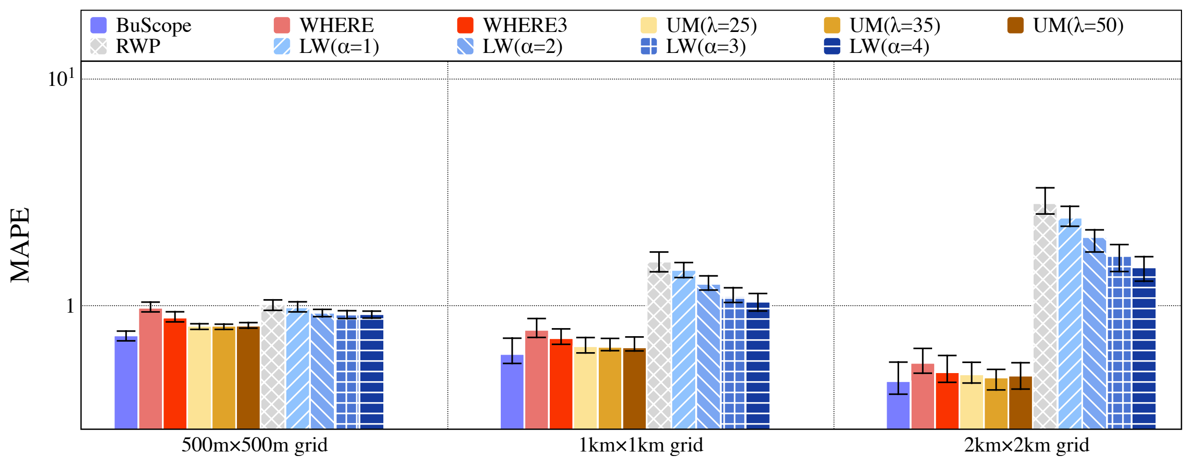


Fig. 5. Visit counts with different grid partition in vehicle tracking system

involved as shown in Figure 2 (a)). The RWP model randomly selects the next destination among all the locations with the same probability, which generates similar visit counts for each location. The LW model with different α adjusts the selection probability of the next location according to the distance between origin and destination pairs, and the distribution of the visit counts of Levy walk with α =4 is similar to the ground truth as shown in Figure 4 (a). But Levy walk is not a context-aware model, which produces extremely large gap for `visit counts` in some locations with ground truth and generated trace as shown in Figure 4 (b). Due to the small number of locations in a camera-based use case, the large MAPE for `visit counts` in some locations cannot be averaged by that in other locations.

Visit Counts for Empirical Models: Three empirical models (BuScope, Universal Model, and WHERE) perform better than two theoretical models at the collective level as shown in Figure 3. For example, in the mobile payment use case, the MAPE of theoretical models are about 2 to 13× than that of the empirical models. BuScope performs better than the universal model because it utilizes both the more detailed individual history trace and the effective combination of individual and collective information. The universal models with different parameters achieve a similar performance in all use cases, which shows their stability.

Visit Count with Different Spatial Granularity: The vehicle tracking use case generates individual traces with the finest spatial granularity and temporal continuity (e.g., GPS locations every 10 seconds), but due to its lowest penetration, the models' performance can be affected by insufficient and less representative data. We divide the Shenzhen map into grids with different sizes, e.g., 500m × 500m, 1km × 1km, and 2km × 2km to evaluate the impact of spatial granularity on measures. As in Figure 5, with the increase of the grid size, the empirical models perform better; whereas the theoretical models perform worse. We explore the reason for this trend. The average daily visit counts for most locations are less than 2 with 500m × 500m grid partition as in Figure 6. As a result, the wrong destination selection due to the inaccurate prediction in the empirical models would cause nearly 1 for MAPE (e.g., zero in the ground truth and non-zero in the generated trace). With a coarser area partition, the average `visit counts` is increased, and the negative effect of the inaccurate prediction for the empirical models can be effectively mitigated.

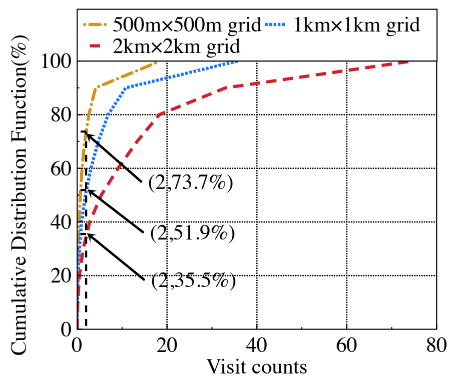


Fig. 6. Visit counts in the vehicle tracking use case

3.1.2 OD Matrix. **OD Matrix vs. Visit Counts:** Generally, the MAPE of the `OD matrix` is larger compared with `visit counts` because `OD matrix` is a more fine-grained measure. Specifically, `visit counts` only cares about the total mobility flows to a destination, which neglects each part of the flows; whereas `OD matrix` takes all the origins into consideration. For example, as in Figure 7, compared with `visit counts` in Figure 3, the cellular system has about 7 to 50× larger MAPE. But it is not the case in the camera-based use case with theoretical models as in Figure 7. The large value gap between some location pairs with the ground truth and the generated trace can be decreased more effectively compared with `visit counts` calculation because there are more elements in `OD matrix`.

OD Matrix for Empirical Models: WHERE and WHERE3 can capture the collective features effectively, e.g., flows between two locations, especially for use cases with a few hops (e.g., 1.56 in the mobile payment system and 2.18 in the camera-based system). Because WHERE and WHERE3 only care about two and three important locations for each individual, respectively.

3.1.3 Lesson Learned on Collective Level Measures. **Impact of Spatial Granularity of Use Cases**: In terms of the collective level measures, the finer spatial granularity of the user cases, the better generalization of the theoretical models. As shown in Figure 3 and Figure 7, the `visit counts` and `OD matrix` have the smaller MAPE in cellular use cases compared with that in mobile payment and camera-based use cases. This insight is

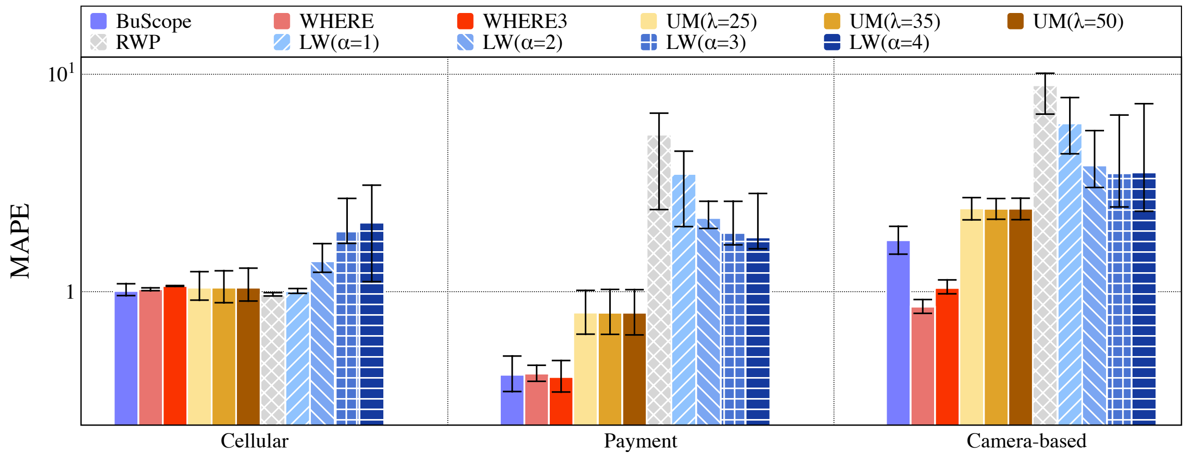


Fig. 7. OD matrix in collective level: Units from left to right in x-axis are with finer to coarser spatial granularity

counter-intuitive because normally we think the performance of a theoretical mobility model should be better with coarser spatial granularity, instead of finer spatial granularity. In contrast, the performance of the empirical models in collective measures is not sensitive to the spatial granularity of the user cases. This insight provides guidance for model selection given use cases with spatial granularity when focusing on collective level measures, e.g., cellular user flows from one area to another area.

3.2 Individual Level Mobility Measures

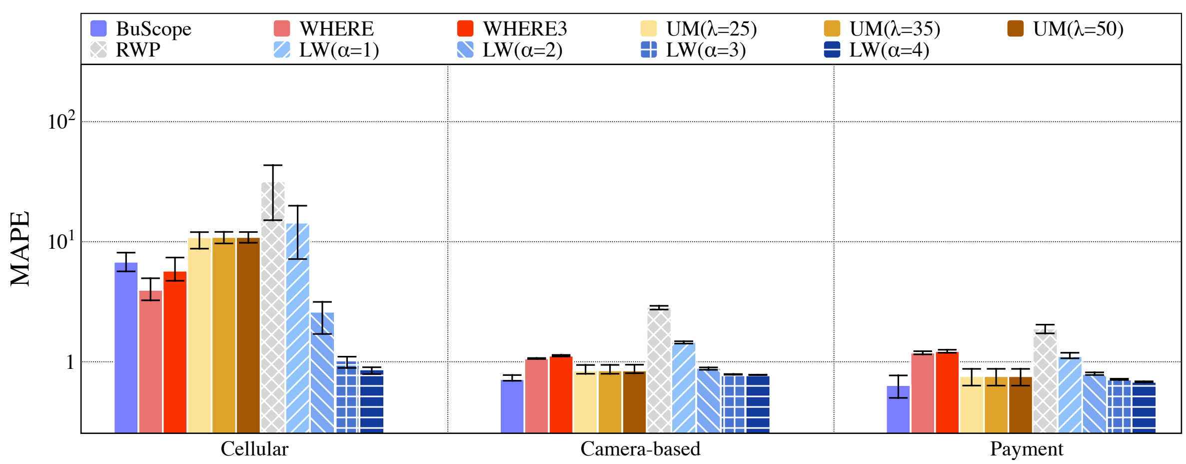


Fig. 8. Average jump length in individual level: Units from left to right are with lower to higher periodic temporal continuity

3.2.1 Average Jump Length. **Average Jump Length for Theoretical Models:** Compared with collective level measures, e.g., `OD matrix`, individual measures care about individual trace, which can capture individual features

better. Levy walk with $\alpha = 3$ or 4 can achieve good performance for this measure in all use cases because of the heavy-tailed travel distance distribution and their location-insensitive measurement.
Average Jump Length for Empirical Models: WHERE and WHERE3 perform much worse in the cellular network use case as shown in Figure 8 compared with in other two use cases. However, the individual average number of hops is 6.94 in the cellular system, which is obviously larger, because WHERE and WHERE3 generates individual traces focusing on two or three important locations for an individual, As a result, WHERE cannot generate reasonable traces for most individuals. The Universal model and BuScope also show a similar trend because of the less periodicity of the temporal continuity. Specifically, the cellular system use case generated a data point when cellphone users have phone calls or messages, which is determined by users' behavior and less periodic compared to other use cases such as vehicular tracking.

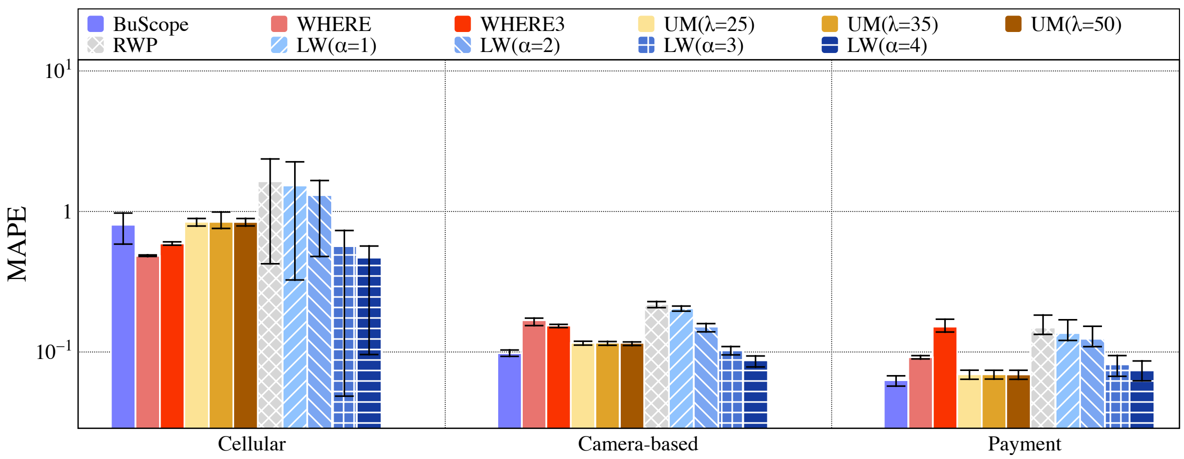


Fig. 9. Number of distinct locations in individual level: Units from left to right in x-axis are with lower to higher periodic temporal continuity

3.2.2 Number of Distinct Visited Locations. **Distinct Visited Locations vs. Average Jump Length:** Compared with the `average jump length`, the MAPE of the distinct visited location is lower (31.7 vs. 2 with RWP in-vehicle system) as in Figure 9. Because the former measure cares more detailed information for an individual, which generates more accurate measurement with the same MAPE.
Distinct Visited Locations for Empirical Models: WHERE utilizes aggregated statistical features to generate the individual traces. As a result, as shown in Figure 8 and Figure 9, WHERE achieves worse performance compared with other empirical models in payment and camera-based use cases. BuScope achieves the best performance in both camera-based and payment use cases because the two mobility modalities are similar to the bus system, which is BuScope designed for. Universal models with different parameters achieve similar performance in individual measures.

3.2.3 Lesson Learned on Individual-Level Measures. **Impact of Temporal Granularity of Use Cases**: In terms of the individual-level measures, the lower periodic temporal continuity of the user cases, the theoretical models typically generalize better than the empirical models. For example, as shown in Figure 8 and Figure 9, in the cellular use case, due to the least periodic temporal continuity, i.e., the cellular users' mobility are logged with lower periodicity, e.g., based on the events of a phone call or messages, an empirical model cannot capture the mobility features effectively. This insight is important for the human mobility model and parameter selection given use cases with lower periodic temporal continuity when we focus on the individual measures.

3.3 Overall Verification Results

We summarize the verification results as follows: (1) the trend of individual measure `user counts` is similar to `visit counts`; (2) the trends of `uncorrelated location entropy`, `maximum jump length` and `radius of gyration` are similar to `average jump length`; (3) the trend of `uncorrelated entropy` is similar to `number of distinct visited locations`. Note that `uncorrelated location entropy` is more sensitive to individual mobility compared with other collective metrics, so its trend is similar to individual measures. The verification results are put into four categories based on a two-dimensional taxonomy, i.e., theoretical and empirical models for both collective and individual measures, as shown in Table 4. The results provide important guidance for model selection and parameter adjustment.

Table 4. Summary of the verification results

Categories	Summary
Theoretical models with collective measures	1. The coarser spatial granularity of the use cases, the lower generalizability. (Figure 3, 7) 2. The MAPE of OD matrix is much larger than visit counts in camera-based system, which is different from that in cellular and payment system. (Figure 3, 7)
Theoretical models with individual measures	Levy walk with larger α shows better generalizability. (Figure 8, 9)
Empirical models with collective measures	1. The empirical models utilizing individual and collective history information can achieve similar generalizability. (Figure 3, 7) 2. The empirical models achieve better generalizability compared to the theoretical models. (Figure 3, 7) 3. Without enough data for analysis, it is better to have coarser spatial partition for empirical models. (Figure 5)
Empirical models with individual measures	1. In the payment and camera-based use case, the empirical models utilizing individual information perform better compared with the model utilizing the collective information. (Figure 8, 9) 2. The generalizability of the empirical models is affected by the periodic temporal continuity significantly. (Figure 8, 9) 3. The universal models with different λ achieve similar generalizability, which shows the stability of these models. (Figure 3, 5, 7, 8, 9)

4 DISCUSSION

4.1 Limitations

In this paper, even large scale data has been used, the main limitation is that we only choose the data collected in one of the most biggest and developed cities in China, e.g., Shenzhen, to verify the generalizability of different models, which may have a bias in other countries or small-scale cities due to different mobility models, culture diversity, and use cases. The second limitation is that we only collect four use cases with different spatial granularity and temporal continuity as shown in Figure 1 to verify the generalizability of the models, our results may not be generalized to other use cases with different spatiotemporal features. However, the four use cases are representative with obviously different spatial granularity and temporal continuity. Another limitation is that we only focus on the performance of five typical mobility models, e.g., two theoretical and three empirical models. There are many generative mobility models have been proposed [16] [33] [46] [68] [62] [66] [64] [32] in recent years. It is infeasible to analyze all the models in different categories. We believe that the verification results in the paper can provide important guidance for the model selection and the parameter adjustment in other use cases as well.

4.2 Potential Implication of the Results

Understanding the accuracy of heterogeneous mobility models with different modality and measures is significant to help to recommend the models and guide the parameter selection for many use cases such as transportation management, epidemic control, urban planning, and autonomous driving. We show a practical guideline for model selection and parameter selection with two use cases: (1) In mobile computing areas, estimating `OD matrix` is important for optimizing autonomous car scheduling and conducting large-scale simulations. If we focus on the cars' mobility on highways, we can refer to the performance of models in camera-based system because of the similar spatial-temporal features. As a result, WHERE and WHERE3 can be good choices. Estimating the `visit counts` of a place can help to build the new urban-scale infrastructures, e.g., 5G base station, to meet the increasing demand of the residents in the city. In this scenario, we can select empirical models due to the obviously superior performance compared with theoretical models in the cellular network system. Overall, the model selection is based on the similar spatial-temporal features; (2) In human behavior analysis, it is important to get human mobility regularity for knowing the spreading mechanism of diseases, e.g., COVID-19. Specifically, we can evaluate the risk of the individual based on the `radius of gyration` and the `number of distinct locations` visited by an individual by choosing the Levy walk model with α=3 or 4 and BuScope model in the cellular network system and mobile payment system, respectively, because of the heavy-tailed distribution of the individual mobility.

4.3 Privacy and Ethics

All data used were legally collected by service providers and Shenzhen Committee of Transportation under user consent. In this work, we focus on mobility model accuracy instead of caring about individual resident in the city. All data has been anonymized by our collaborators, so we cannot use these data to trace back the individual information, e.g., the driver. Besides, we only store and utilize the data involved in our work and the other individual information has been dropped for minimal information exposure, e.g., the user connection details in cellular network system.

4.4 Public Data Access

Having access to the datasets is important for people to study the mobility modeling. However, many researchers usually cannot obtain the data. As an initial step, we are working with Shenzhen smart city research group to release some sample data to benefit the research community and promote more effective and robust mobility models. Due to the privacy issue, we need to deal with the raw data first, e.g., doing the hash operation for individual ID, discarding sensitive information and adding some noise to the data with the differential privacy techniques.

5 RELATED WORK

Human mobility modeling has received considerable attention in recent years due to its importance to various sensing applications [54] [2] [22] [19] [5]. We divide the existing work into four categories based on a two-dimension taxonomy, i.e., collective and individual measures for both modeling single source and multiple source data as shown in Table 5.

Table 5. Categories of human mobility modeling

Categories		Sensing system: Single	Sensing system: Multiple
Measures	Collective	[16][33] [46][14]	[42][68] [62][67]
	Individual	[18][64][66] [20][32]	[61][38]

5.1 Models Focusing on Collective Measures

Gravity model is proposed based on the Newton's law of gravitation [71], which assumes that the people flow between two locations is proportional to the population of these two locations and inversely proportional to the distance between the two locations [1] [4]. Simini *et al.* propose Radiation model [49] to address the drawback of Gravity model which can not capture the difference of flow between two locations in reverse directions. However, it is difficult to obtain the population information for Gravity model and Radiation model because of the dynamic human mobility characteristic. Isaacman *et al.* design *WHERE* model based on Call Detail Records (CDRs) to model general human mobility, which focuses on a few important locations, e.g., home and working places, and evaluate it on the population movement between different locations [16]. Mir *et al.* propose *DP-WHERE* model which adds noise to the probability distributions in *WHERE* to prevent the privacy leak [33]. Shafiq *et al.* characterize the scheme of flow migration during the real-life crowded events based on the cellular network data [46]. Fang *et al.* model real-time urban population based on Call Detailed Records (CDRs) from multiple cellular networks deployed in Shenzhen to avoid over-fitting of models with only data collected from single cellular network [14]. Rezaei *et al* provide an improved upper bound for information flood time with Manhattan random way-point Model and verify the accuracy based on the bike renting data and taxi GPS data [42]. Zhang *et al* model real-time human mobility with transportation and cellphone data to solve the over-fitting of single-view modeling [68]. Yan *et al* propose a non-parameter human mobility model based on population-weighted opportunities with taxi GPS and cellphone data [62]. Zhang *et al* explore the human mobility based on multi-source data, e.g., vehicle data, smart card data and cellphone data, to avoid data bias with single-source data and improve the accuracy of the modeling [67].

5.2 Models Focusing on Individual Measures

Random Way-point is a typical and simple theoretical model to generate individual trajectory [26], which selects the next location randomly from all the locations for each individual. Levy walk [15] adjusts the probability of generating the next location for the individual based on the distance between origin and destination. Random Way-point and Levy walk models are theoretical models, which can not utilize the individual or collective history information. [66] [64] exploits the ETC data for mobility modeling on highways at an individual level, e.g., the travel time and real-time location. Jang *et al* estimate travel times among stops based on smart cart data from automatic fare collection systems [18]. Jiang *et al* provide a data mining framework to extract individual mobility patterns from CDRs collected in Singapore [20]. Lakmal *et al* apply real-time data collected by Singapore public transit system to model individual mobility [32]. Yan *et al* develop an individual model combing history information and the population ranks to capture the mobility patterns [61]. Pappalardo *et al* extract two distinct profiles in human mobility modeling, e.g., returners and explorers, and combines Gravity model with EPR model [50] to improve the modeling accuracy [38] with CDR and GPS datasets. Meegahapola *et al* utilize individual and aggregated history information to predict the destination of each individual with public transportation data in Singapore [32].

5.3 Summary

The most of existing research works mainly focus on collective mobility modeling, e.g., the flows between two locations and the flows to/from a location, or based on the single sensing systems. Even though some works propose individual models to capture detailed mobility patterns, they only exploit data collected from one or two use cases, e.g., cellular networks and GPS systems. However, to our knowledge, MOVER is the first work to verify the generalizability of both theoretical and empirical models at both individual level and collective level based on a taxonomy of the mobility use case (as shown in Figure 1) from two dimensions, which are spatial granularity and temporal continuity.

6 CONCLUSION

In this paper, we design and implement a mobility model verification framework called MOVER, aiming to verify the model generalizability from both the collective level and individual level with four diverse use cases. MOVER takes these four use cases, five models as input, and outputs the model accuracy and generalizability through four collective and five individual metrics.

The verification results in this work show: (i) For the collective level measures, the finer spatial granularity of the user cases, the better generalization of the theoretical models; (ii) For the individual-level measures, the lower periodic temporal continuity of the user cases, the theoretical models typically generalize better than the empirical models. We believe our verification results will provide better model selection and parameter adjustment for improving many related use cases in mobile computing, cellular communication, epidemic control, and urban planning.

ACKNOWLEDGMENTS

The authors would like to thank anonymous reviewers for their valuable comments and suggestions. This work is partially supported by NSF 1849238, NSF 1932223, NSF 1952096, NSF 1951890, NSF 2003874, and NSF 2047822.

REFERENCES

[1] James E Anderson. 2011. The gravity model. *Annu. Rev. Econ.* 3, 1 (2011), 133–160.
[2] Duygu Balcan, Vittoria Colizza, Bruno Gonçalves, Hao Hu, José J Ramasco, and Alessandro Vespignani. 2009. Multiscale mobility networks and the spatial spreading of infectious diseases. *Proceedings of the National Academy of Sciences* 106, 51 (2009), 21484–21489.
[3] Duygu Balcan and Alessandro Vespignani. 2011. Phase transitions in contagion processes mediated by recurrent mobility patterns. *Nature physics* 7, 7 (2011), 581–586.
[4] Marc Barthélemy. 2011. Spatial networks. *Physics Reports* 499, 1-3 (2011), 1–101.
[5] Federica Bogo and Enoch Peserico. 2013. Optimal throughput and delay in delay-tolerant networks with ballistic mobility. In *Proceedings of the 19th annual international conference on Mobile computing & networking*. 303–314.
[6] Dirk Brockmann, Lars Hufnagel, and Theo Geisel. 2006. The scaling laws of human travel. *Nature* 439, 7075 (2006), 462–465.
[7] F. Calabrese, G. Di Lorenzo, Liu Liang, and C. Ratti. 2011. Estimating Origin-Destination Flows Using Mobile Phone Location Data. *IEEE Pervasive Computing* 10, 4 (2011), 36–44.
[8] Eunjoon Cho, Seth A Myers, and Jure Leskovec. 2011. Friendship and mobility: user movement in location-based social networks. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*. 1082–1090.
[9] Stephen Eubank, Hasan Guclu, VS Anil Kumar, Madhav V Marathe, Aravind Srinivasan, Zoltan Toroczkai, and Nan Wang. 2004. Modelling disease outbreaks in realistic urban social networks. *Nature* 429, 6988 (2004), 180–184.
[10] Zhihan Fang, Guang Wang, Shuai Wang, Chaoji Zuo, Fan Zhang, and Desheng Zhang. 2020. CellRep: Usage Representativeness Modeling and Correction Based on Multiple City-Scale Cellular Networks. In *Proceedings of The Web Conference 2020*. 584–595.
[11] Zhihan Fang, Guang Wang, and Desheng Zhang. 2020. Modeling Fine-Grained Human Mobility on Cellular Networks. In *Companion Proceedings of the Web Conference 2020*. 35–37.
[12] Zhihan Fang, Yu Yang, Shuai Wang, Boyang Fu, Zixing Song, Fan Zhang, and Desheng Zhang. 2019. MAC: Measuring the impacts of anomalies on travel time of multiple transportation systems. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 2 (2019), 1–24.

[13] Zhihan Fang, Yu Yang, Guang Yang, Yikuan Xian, Fan Zhang, and Desheng Zhang. 2021. CellSense: Human Mobility Recovery via Cellular Network Data Enhancement. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 5, 3 (2021), 1–22.

[14] Zhihan Fang, Fan Zhang, Ling Yin, and Desheng Zhang. 2018. MultiCell: Urban population modeling based on multiple cellphone networks. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 3 (2018), 1–25.

[15] Marta C Gonzalez, Cesar A Hidalgo, and Albert-Laszlo Barabasi. 2008. Understanding individual human mobility patterns. *nature* 453, 7196 (2008), 779–782.

[16] Sibren Isaacman, Richard Becker, Ramón Cáceres, Margaret Martonosi, James Rowland, Alexander Varshavsky, and Walter Willinger. 2012. Human mobility modeling at metropolitan scales. In *Proceedings of the 10th international conference on Mobile systems, applications, and services*. 239–252.

[17] AK Jain, MN Murty, and PJ Flynn. 1999. Estimating origin-destination flows using mobile phone location data. *Comput. Surveys* 31, 3 (1999), 264–323.

[18] Wonjae Jang. 2010. Travel time and transfer analysis using transit smart card data. *Transportation research record* 2144, 1 (2010), 142–149.

[19] Jayson S Jia, Jianmin Jia, Christopher K Hsee, and Baba Shiv. 2017. The role of hedonic behavior in reducing perceived risk: evidence from postearthquake mobile-app data. *Psychological science* 28, 1 (2017), 23–35.

[20] Shan Jiang, Joseph Ferreira, and Marta C Gonzalez. 2017. Activity-based human mobility patterns inferred from mobile phone data: A case study of Singapore. *IEEE Transactions on Big Data* 3, 2 (2017), 208–219.

[21] Woo-Sung Jung, Fengzhong Wang, and H Eugene Stanley. 2008. Gravity model in the Korean highway. *EPL (Europhysics Letters)* 81, 4 (2008), 48005.

[22] Xiangjie Kong, Ximeng Song, Feng Xia, Haochen Guo, Jinzhong Wang, and Amr Tolba. 2018. LoTAD: Long-term traffic anomaly detection based on crowdsourced bus trajectory data. *World Wide Web* 21, 3 (2018), 825–847.

[23] Xiangjie Kong, Feng Xia, Jinzhong Wang, Azizur Rahim, and Sajal K Das. 2017. Time-location-relationship combined service recommendation based on taxi trajectory data. *IEEE Transactions on Industrial Informatics* 13, 3 (2017), 1202–1212.

[24] Xiangjie Kong, Zhenzhen Xu, Guojiang Shen, Jinzhong Wang, Qiuyuan Yang, and Benshi Zhang. 2016. Urban traffic congestion estimation and prediction based on floating car trajectory data. *Future Generation Computer Systems* 61 (2016), 97–107.

[25] Moritz UG Kraemer, Chia-Hung Yang, Bernardo Gutierrez, Chieh-Hsi Wu, Brennan Klein, David M Pigott, Louis Du Plessis, Nuno R Faria, Ruoran Li, William P Hanage, et al. 2020. The effect of human mobility and control measures on the COVID-19 epidemic in China. *Science* 368, 6490 (2020), 493–497.

[26] Jean-Yves Le Boudec and Milan Vojnovic. 2006. The random trip model: stability, stationary regime, and perfect simulation. *Ieee/Acm Transactions On Networking* 14, 6 (2006), 1153–1166.

[27] Yexin Li and Yu Zheng. 2019. Citywide bike usage prediction in a bike-sharing system. *IEEE Transactions on Knowledge and Data Engineering* 32, 6 (2019), 1079–1091.

[28] Yuxuan Liang, Kun Ouyang, Lin Jing, Sijie Ruan, Ye Liu, Junbo Zhang, David S Rosenblum, and Yu Zheng. 2019. Urbanfm: Inferring fine-grained urban flows. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 3132–3142.

[29] Xin Lu, Linus Bengtsson, and Petter Holme. 2012. Predictability of population displacement after the 2010 Haiti earthquake. *Proceedings of the National Academy of Sciences* 109, 29 (2012), 11576–11581.

[30] Yuyi Mao, Changsheng You, Jun Zhang, Kaibin Huang, and Khaled B Letaief. 2017. A survey on mobile edge computing: The communication perspective. *IEEE Communications Surveys & Tutorials* 19, 4 (2017), 2322–2358.

[31] László Mátyás. 1997. Proper econometric specification of the gravity model. *World Economy* 20, 3 (1997), 363–368.

[32] Lakmal Meegahapola, Thivya Kandappu, Kasthuri Jayarajah, Leman Akoglu, Shili Xiang, and Archan Misra. 2019. Buscope: Fusing individual & aggregated mobility behavior for" live" smart city services. In *Proceedings of the 17th Annual International Conference on Mobile Systems, Applications, and Services*. 41–53.

[33] Darakhshan J Mir, Sibren Isaacman, Ramón Cáceres, Margaret Martonosi, and Rebecca N Wright. 2013. Dp-where: Differentially private modeling of human mobility. In *2013 IEEE international conference on big data*. IEEE, 580–588.

[34] Zheyi Pan, Yuxuan Liang, Weifeng Wang, Yong Yu, Yu Zheng, and Junbo Zhang. 2019. Urban traffic prediction from spatio-temporal data using deep meta learning. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 1720–1730.

[35] Luca Pappalardo, Gianni Barlacchi, Roberto Pellungrini, and Filippo Simini. 2019. Human Mobility from theory to practice: Data, Models and Applications. In *Companion Proceedings of The 2019 World Wide Web Conference*. 1311–1312.

[36] Luca Pappalardo, Salvatore Rinzivillo, Zehui Qu, Dino Pedreschi, and Fosca Giannotti. 2013. Understanding the patterns of car travel. *The European Physical Journal Special Topics* 215, 1 (2013), 61–73.

[37] Luca Pappalardo and Filippo Simini. 2018. Data-driven generation of spatio-temporal routines in human mobility. *Data Mining and Knowledge Discovery* 32, 3 (2018), 787–829.

[38] Luca Pappalardo, Filippo Simini, Salvatore Rinzivillo, Dino Pedreschi, Fosca Giannotti, and Albert-László Barabási. 2015. Returners and explorers dichotomy in human mobility. *Nature communications* 6, 1 (2015), 1–8.

[39] Luca Pappalardo, Maarten Vanhoof, Lorenzo Gabrielli, Zbigniew Smoreda, Dino Pedreschi, and Fosca Giannotti. 2016. An analytical framework to nowcast well-being using mobile phone data. *International Journal of Data Science and Analytics* 2, 1-2 (2016), 75–92.

[40] Constance Potter. 2010. NEW QUESTIONS IN THE 1940 CENSUS. *Prologue (Washington, D.C.)* 42, 4 (2010), 46–52.

[41] Zhou Qin, Fang Cao, Yu Yang, Shuai Wang, Yunhuai Liu, Chang Tan, and Desheng Zhang. 2020. CellPred: A Behavior-aware Scheme for Cellular Data Usage Prediction. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 1 (2020), 1–24.

[42] Aria Rezaei, Jie Gao, Jeff M Phillips, and Csaba D Tóth. 2018. Improved bounds on information dissemination by manhattan random waypoint model. In *Proceedings of the 26th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*. 139–148.

[43] Injong Rhee, Minsu Shin, Seongik Hong, Kyunghan Lee, Seong Joon Kim, and Song Chong. 2011. On the levy-walk nature of human mobility. *IEEE/ACM transactions on networking* 19, 3 (2011), 630–643.

[44] Alex Rutherford, Manuel Cebrian, Sohan Dsouza, Esteban Moro, Alex Pentland, and Iyad Rahwan. 2013. Limits of social mobilization. *Proceedings of the National Academy of Sciences* 110, 16 (2013), 6281–6286.

[45] Todd N Schoepflin and Daniel J Dailey. 2003. Dynamic camera calibration of roadside traffic management cameras for vehicle speed estimation. *IEEE Transactions on Intelligent Transportation Systems* 4, 2 (2003), 90–98.

[46] M Zubair Shafiq, Lusheng Ji, Alex X Liu, Jeffrey Pang, Shobha Venkataraman, and Jia Wang. 2016. Characterizing and optimizing cellular network performance during crowded events. *IEEE/ACM Transactions on Networking* 24, 3 (2016), 1308–1321.

[47] Barrie Sheldon. 2011. Camera surveillance within the UK: Enhancing public safety or a social threat? *International Review of Law, Computers & Technology* 25, 3 (2011), 193–203.

[48] Michael F Shlesinger, Joseph Klafter, and Gert Zumofen. 1999. Above, below and beyond Brownian motion. *American Journal of Physics* 67, 12 (1999), 1253–1259.

[49] Filippo Simini, Marta C González, Amos Maritan, and Albert-László Barabási. 2012. A universal model for mobility and migration patterns. *Nature* 484, 7392 (2012), 96–100.

[50] Chaoming Song, Tal Koren, Pu Wang, and Albert-László Barabási. 2010. Modelling the scaling properties of human mobility. *Nature Physics* 6, 10 (2010), 818–823.

[51] Michele Tizzoni, Paolo Bajardi, Adeline Decuyper, Guillaume Kon Kam King, Christian M Schneider, Vincent Blondel, Zbigniew Smoreda, Marta C González, and Vittoria Colizza. 2014. On the use of human mobility proxies for modeling epidemics. *PLoS Comput Biol* 10, 7 (2014), e1003716.

[52] Cheng-Lin Tsao, Yueh-Ting Wu, Wanjiun Liao, and Jia-Chun Kuo. 2006. Link duration of the random way point model in mobile ad hoc networks. In *IEEE Wireless Communications and Networking Conference, 2006. WCNC 2006.*, Vol. 1. IEEE, 367–371.

[53] Guang Wang, Yongfeng Zhang, Zhihan Fang, Shuai Wang, Fan Zhang, and Desheng Zhang. 2020. FairCharge: A data-driven fairness-aware charging recommendation system for large-scale electric taxi fleets. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 1 (2020), 1–25.

[54] Jinzhong Wang, Xiangjie Kong, Azizur Rahim, Feng Xia, Amr Tolba, and Zafer Al-Makhadmeh. 2017. IS2Fun: Identification of Subway Station Functions Using Massive Urban Data. *IEEE Access* 5 (2017), 27103–27113.

[55] Leye Wang, Xu Geng, Xiaojuan Ma, Feng Liu, and Qiang Yang. 2018. Cross-city transfer learning for deep spatio-temporal prediction. *arXiv preprint arXiv:1802.00386* (2018).

[56] Yuqiang Wang, Yuguang Wei, Qi Zhang, Hua Shi, and Pan Shang. 2019. Scheduling overnight trains for improving both last and first train services in an urban subway network. *Advances in Mechanical Engineering* 11, 5 (2019), 1687814019848920.

[57] Feng Xia, Azizur Rahim, Xiangjie Kong, Meng Wang, Yinqiong Cai, and Jinzhong Wang. 2017. Modeling and Analysis of Large-scale Urban Mobility for Green Transportation. *IEEE Transactions on Industrial Informatics* (2017), 1–1.

[58] Feng Xia, Jinzhong Wang, Xiangjie Kong, Zhibo Wang, Jianxin Li, and Chengfei Liu. 2018. Exploring Human Mobility Patterns in Urban Scenarios: A Trajectory Data Perspective. *IEEE Communications Magazine* 56, 3 (2018), 142–149.

[59] Xiaoyang Xie, Zhihan Fang, Yang Wang, Fan Zhang, and Desheng Zhang. 2020. RISC: Resource-Constrained Urban Sensing Task Scheduling Based on Commercial Fleets. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 2 (2020), 1–20.

[60] Xiaoyang Xie, Yu Yang, Zhihan Fang, Guang Wang, Fan Zhang, Fan Zhang, Yunhuai Liu, and Desheng Zhang. 2018. coSense: Collaborative urban-scale vehicle sensing based on heterogeneous fleets. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 4 (2018), 1–25.

[61] Xiao-Yong Yan, Wen-Xu Wang, Zi-You Gao, and Ying-Cheng Lai. 2017. Universal model of individual and population mobility on diverse spatial scales. *Nature communications* 8, 1 (2017), 1–9.

[62] Xiao-Yong Yan, Chen Zhao, Ying Fan, Zengru Di, and Wen-Xu Wang. 2014. Universal predictability of mobility patterns in cities. *Journal of The Royal Society Interface* 11, 100 (2014), 20140834.

[63] Yu Yang, Zhihan Fang, Xiaoyang Xie, Fan Zhang, Yunhuai Liu, and Desheng Zhang. 2020. Extending Coverage of Stationary Sensing Systems with Mobile Sensing Systems for Human Mobility Modeling. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 4, 3 (2020), 1–21.

[64] Yu Yang, Xiaoyang Xie, Zhihan Fang, Fan Zhang, Yang Wang, and Desheng Zhang. 2019. VeMo: Enabling Transparent Vehicular Mobility Modeling at Individual Levels with Full Penetration. In *The 25th Annual International Conference on Mobile Computing and Networking*. 1–16.

[65] Yu Yang, Xiaoyang Xie, Zhihan Fang, Fan Zhang, Yang Wang, and Desheng Zhang. 2020. Vemo: Enabling transparent vehicular mobility modeling at individual levels with full penetration. *IEEE Transactions on Mobile Computing* (2020).

[66] Yu Yang, Fan Zhang, and Desheng Zhang. 2018. SharedEdge: GPS-free fine-grained travel time estimation in state-level highway systems. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 2, 1 (2018), 1–26.

[67] Desheng Zhang, Jun Huang, Ye Li, Fan Zhang, Chengzhong Xu, and Tian He. 2014. Exploring human mobility with multi-source data at extremely large metropolitan scales. In *Proceedings of the 20th annual international conference on Mobile computing and networking*. 201–212.

[68] Desheng Zhang, Juanjuan Zhao, Fan Zhang, and Tian He. 2015. coMobile: real-time human mobility modeling at urban scale using multi-view learning. In *Proceedings of the 23rd SIGSPATIAL international conference on advances in geographic information systems*. 1–10.

[69] Yi Zhao, Xu Wang, Jianbo Li, Desheng Zhang, and Zheng Yang. 2019. CellTrans: Private Car or Public Transportation? Infer Users' Main Transportation Modes at Urban Scale with Cellular Data. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* 3, 3 (2019), 1–26.

[70] Yu Zheng, Furui Liu, and Hsun-Ping Hsieh. 2013. U-air: When urban air quality inference meets big data. In *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*. 1436–1444.

[71] George Kingsley Zipf. 1946. The P 1 P 2/D hypothesis: on the intercity movement of persons. *American sociological review* 11, 6 (1946), 677–686.

A APPENDIX

We complete the verification results of different models with the remaining measures mention in Section 2.3, including two collective and three individual level measures.

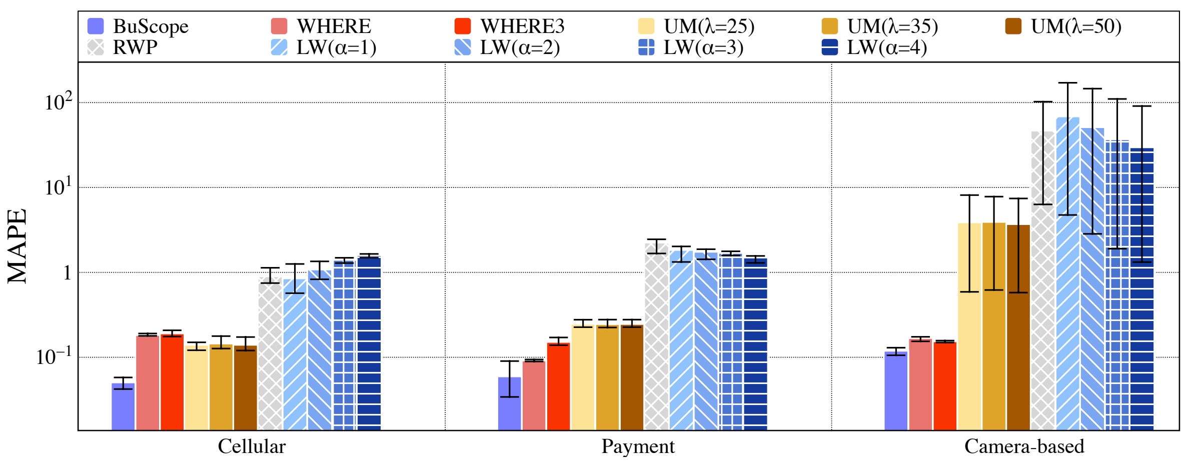


Fig. 10. User counts in collective level: Units from left to right in x-axis are with finer to coarser spatial granularity

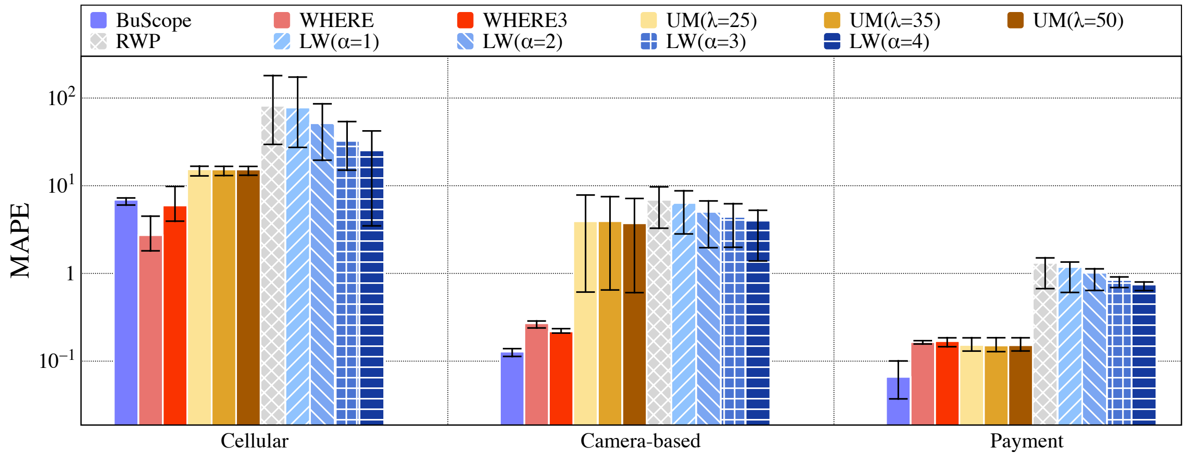


Fig. 11. Uncorrelated location entropy in collective level: Units from left to right in x-axis are with lower to higher periodic temporal continuity

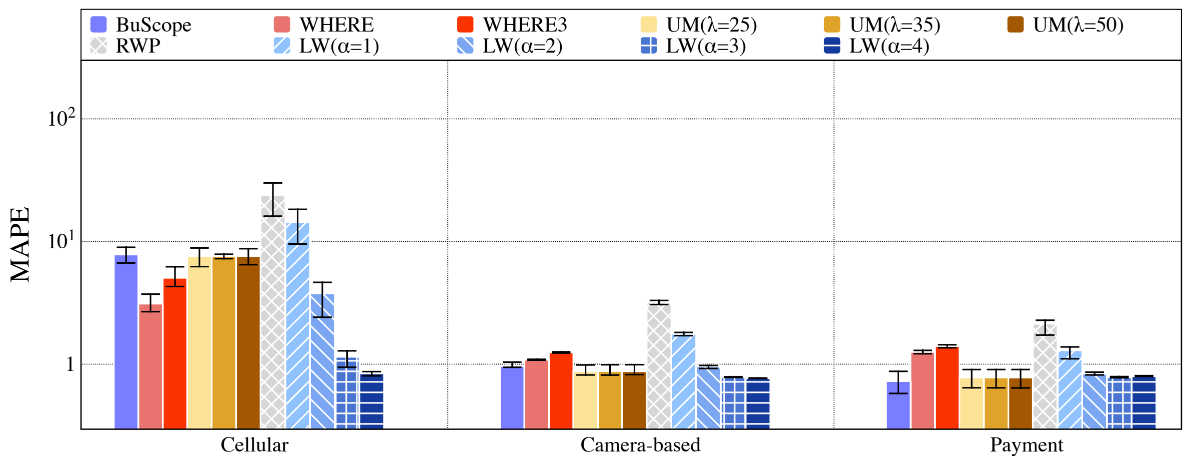


Fig. 12. Maximum jump length in individual level: Units from left to right in x-axis are with lower to higher periodic temporal continuity

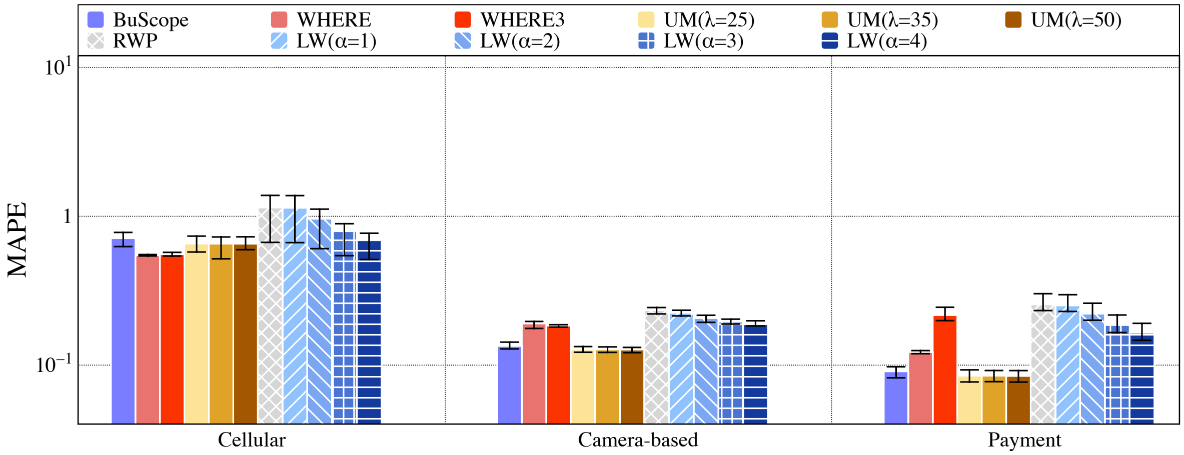


Fig. 13. Uncorrelated entropy in individual level: Units from left to right in x-axis are with lower to higher periodic temporal continuity

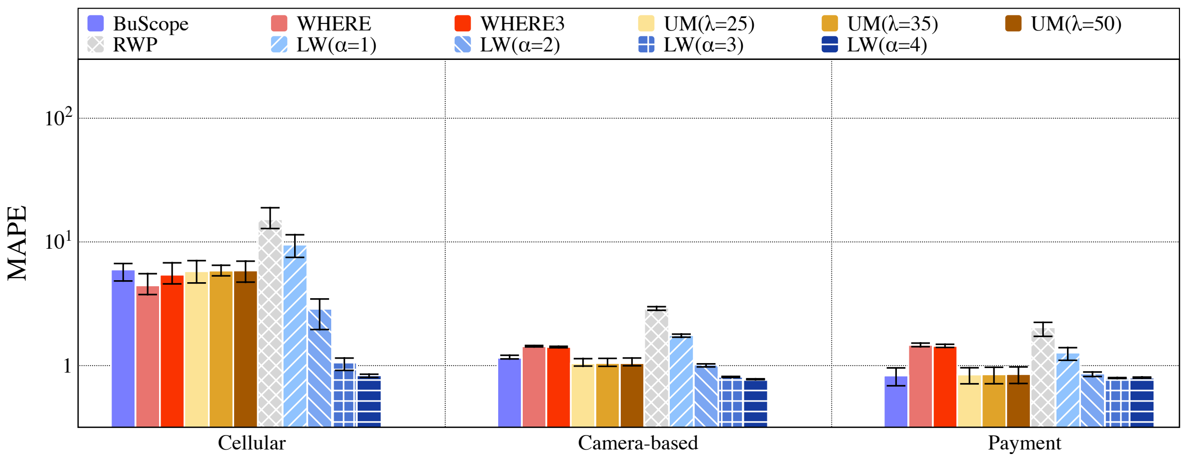


Fig. 14. Radius of gyration in individual level: Units from left to right in x-axis are with lower to higher periodic temporal continuity