Universal model of individual and population mobility on diverse spatial scales

Xiao-Yong Yan[1], Wen-Xu Wang [2], Zi-You Gao[1] & Ying-Cheng Lai[3]

Studies of human mobility in the past decade revealed a number of general scaling laws. However, to reproduce the scaling behaviors quantitatively at both the individual and population levels simultaneously remains to be an outstanding problem. Moreover, recent evidence suggests that spatial scales have a significant effect on human mobility, raising the need for formulating a universal model suited for human mobility at different levels and spatial scales. Here we develop a general model by combining memory effect and population-induced competition to enable accurate prediction of human mobility based on population distribution only. A variety of individual and collective mobility patterns such as scaling behaviors and trajectory motifs are accurately predicted for different countries and cities of diverse spatial scales. Our model establishes a universal underlying mechanism capable of explaining a variety of human mobility behaviors, and has significant applications for understanding many dynamical processes associated with human mobility.

[1] Institute of Transportation System Science and Engineering, Beijing Jiaotong University, Beijing 100044, China. [2] School of Systems Science and Center for Complexity Research, Beijing Normal University, Beijing 100875, China. [3] School of Electrical, Computer and Energy Engineering, Arizona State University, Tempe, AZ 85287, USA. Correspondence and requests for materials should be addressed to W.-X.W. (email: wenxuwang@bnu.edu.cn) or to Z.-Y.G. (email: zygao@bjtu.edu.cn) or to Y.-C.L. (email: Ying-Cheng.Lai@asu.edu)

Human movements typically occur in spatial regions/domains, such as countries or cities, which can have vastly different scales. For example, in China or the United States, the size of the region can be on the order of millions of square kilometers, while in small countries such as Belgium, the domain size is only about 10 km. (Here we regard international travel as atypical and exclude it from our consideration.) Comparing Belgium with China or the United States, the difference in the spatial scale in terms of areas is at least two orders of magnitude. Typical examples of human movements at both the individual and population levels in countries of diverse size are shown in Fig. 1. Overall, for human mobility, there are large scales exemplified by countries such as China and the United States, and small scales as represented by small countries or big cities in a large country.

A remarkable discovery in complexity science in the past decade is that human mobility obeys certain universal scaling laws[1–20]. It was recently revealed[16], however, that human mobility at small spatial scales tends to exhibit different scaling behaviors. Existing models of human mobility are tailored at understanding and characterizing scaling laws either at large (e.g., big country) or small (e.g., city) scales—we lack a universal model capable of quantifying human movements across all spatial scales. Another deficiency of existing models is that they can explain human mobility patterns either at the individual or at the population level, but not both. The purpose of this paper is to develop a model to fill this gap in our knowledge about human mobility.

To understand the dynamics of human movements and to uncover scaling laws underlying human mobility are of fundamental importance as they are relevant to problems such as disease control, social stability, congestion alleviation, information propagation, and e-commerce[21–28]. Data-based modeling research on human mobility started about a decade ago[1], where the trajectories of bank notes were traced over a reasonably long time period. On the basis of empirical data, the study unveiled two scaling laws on geographical scales: (1) the distribution of the traveling distance exhibits a power-law decay, which can be described by Lévy flights, and (2) the probability of remaining in a small region exhibits an algebraically long tail with an exponential cutoff, which is characteristic of a superdiffusive behavior. Existence of universal patterns in the statistical description of human movements was hinted at through an analysis of the mobile phone data[6], and the issue of predictability of human movement patterns was also addressed[16,29]. The correlation between human movements in the cyberspace and in the physical space has also been studied through big data analysis[15,20]. Quite recently, a scaling law connecting human mobility and social interaction (communication) patterns was uncovered[30].

From a modeling perspective, the classic gravity model[31] represented perhaps the earliest attempt to mathematically understand the mobility flow between two locations. For human mobility on large spatial scales, e.g., as revealed by the trajectories of bank notes, a two-parameter continuous-time random walk model was derived to explain, at a detailed and quantitative level, the empirically observed scaling laws[1]. A statistical, self-consistent microscopic mobility model[2] and a macroscopic model, the so-called radiation model[3–5] that takes into account local mobility decisions, were developed. Inspired by these models, a variety of alternative mechanisms aiming at understanding and characterizing the empirical scaling laws obtained from data have been conceived[6–20]. Most classic gravity-based
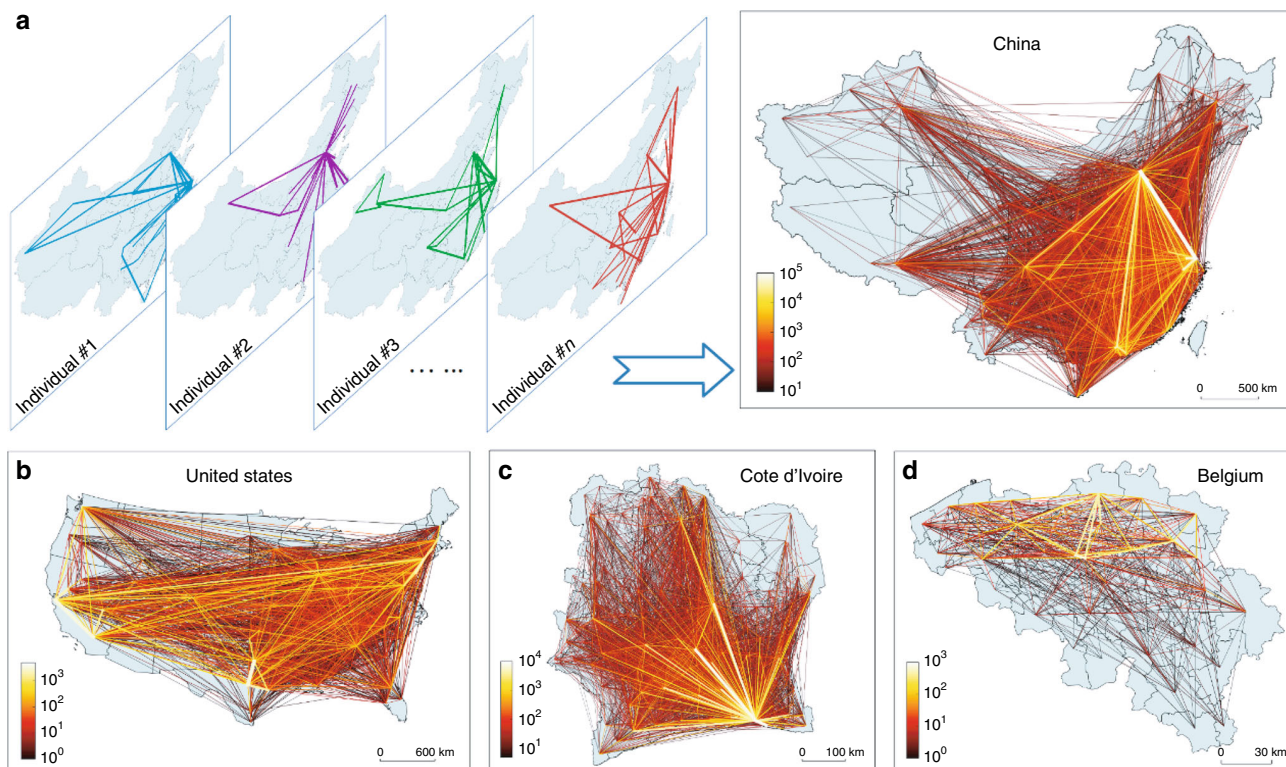

Fig. 1 Real-world examples of individual trajectories and collective movements. **a** Four examples of an individual trajectory from an empirical data set from mainland China and the corresponding collective movements. **b–d** Collective movements embedded in the data sets from the continental United States, Cote d'Ivoire, and Belgium. Here the color bar represents the amount of mobility flux among locations per unit time, where a brighter (darker) line indicates a stronger (weaker) flux. Note that the spatial scales associated with these data sets are drastically different
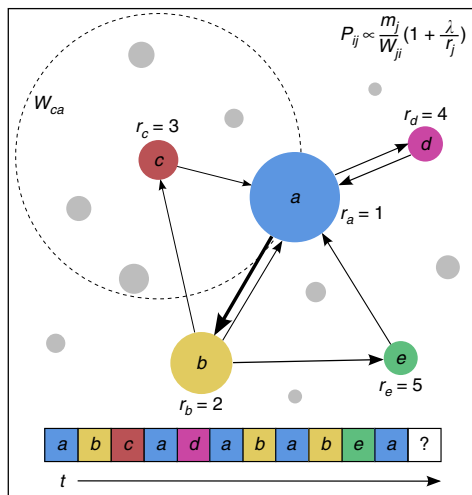
Fig. 2 Model illustration. A typical trajectory visiting the five locations denoted by letters $a − e$ with different colors is indicated at the bottom: $a \rightarrow b \rightarrow c \rightarrow a \rightarrow$ .... The size of the circles that contain a letter indicates the relative attractiveness of the corresponding location as characterized by the index r_j, where r_1 is the most attractive location, r_2 is the second most attractive one, and so on. The dashed circle centered at location c indicates that the travelers moves from c to a, whose radius is the distance between c and a, and the total population within the dashed circle is W_{ca}. The model contains a single parameter, λ, which can be determined from each set of empirical data

macroscopic models are static in the sense that they assume certain mobility decisions so that the transition matrix between spatial locations can be constructed through the corresponding distances and population distribution. While the models can explain the scaling laws to a certain degree, the detailed dynamical features associated with human movements at the individual level are often lost, which are important to understand issues such as the spreading speed and range of diseases. The existing microscopic models[2,8,15,20,32–34] can capture the individual movements and the associated scaling laws, but the mutual interactions among the individuals at the population level were largely ignored.

In this paper, we aim to articulate a human mobility model capable of predicting statistical and scaling behaviors on all spatial scales at both the individual and population levels. To accomplish this goal, it is necessary to identify the general mechanisms underlying human movements independent of the spatial scale. An essential element upon which any human mobility modeling is built is the transition probability for an individual to move from one location to another at any time, from which all kinds of scaling laws can be derived. There are two basic elements that we exploit to construct the transition probability. First, locations are differentiated according to their relative attractiveness, so human movements tend to be biased toward the more attractive locations. This can be modeled by assigning each location a fixed amount of attractiveness. The second element is the memory effect, by which individuals tend to move preferentially to previously visited locations. At a quantitative level, the memory effect can be taken into account by assigning previously visited locations with a relatively high amount of attractiveness. As a result, in our model a location contains two kinds of attractiveness: one simply determined by the population at the location, which is analogous to that in the radiation model[3] or the population-weighted opportunities model[16], and another determined by the memory effect. For any location, the transition probability for an individual to move into

it is proportional to its total attractiveness. More specifically, before making a movement, an individual evaluates the attractiveness of all the available destinations and then moves to a specific destination according to the transition probability. With the transition probability so determined, our model contains a single parameter, which can be determined from each set of empirical data. After the parameter is fixed, our model can simultaneously generate a number of key scaling laws at both the individual and population level, as well as the trajectory motifs, which are in good agreement with empirical results from data associated with arbitrarily spatial scales.

Results

Model. Typical examples of trajectories of human movements at the individual and population levels are shown in Fig. 1. Our aim is to develop a model that can capture the statistical features and predict the scaling laws associated with trajectories at both the individual and population levels, regardless of the spatial scale. A key quantity is the transition probability. In the recent population-weighted opportunities model[16], the transition probability to a destination is proportional to its attractiveness. This probabilistic rule determines the movement of any individual but for only one time step. In order to quantify the statistical behaviors and the scaling laws, sufficiently long trajectories of a large number of individuals are needed.

An important characteristic of human movement, which distinguishes its dynamics from the diffusion dynamics of physical particles, is the memory effect[2,20,33]. In particular, individuals tend to frequently return to previously visited locations. There are different approaches to taking memory effect into account. For example, in the exploration and preferential return (EPR) model[2], it was assumed that the probability for an individual to visit a new location is $p \propto S^{-\gamma}$, where S is the total number of locations that the individual has already visited and $\gamma > 0$ is a model parameter. The probability for the individual to visit a previous location is thus $1 − p$. The algebraic dependence of p on S indicates that the more locations that an individual has visited, the smaller the probability would be for him/her to explore any new location. That is, there is a strong preference for an individual to move among locations that have been visited previously. The model also assumes[2] that the probability for an individual to move to a previously visited location is proportional to the frequency at which it has been visited. This model can successfully reproduce the visiting frequency distribution of the locations obtained from empirical data, as well as the rate of increase in the number of locations. A subsequent model[33] emphasizing the importance of the memory effect assumes that the probability distribution of the return time interval, $P(\tau)$, is known. An individual chooses a value of τ from the distribution to determine the location that he/she wishes to return to. While this model can reproduce the empirically obtained rate of increase of new locations, the choice of $P(\tau)$ is mostly heuristic.

The basic idea underlying the development of our model is that the attractiveness of a location for an individual is determined by both the memory of the individual and the population at the location. Let A be a quantity measuring the effect of memory on the attractiveness of a location to an individual. It is reasonable to assume that a more attractive location can in general impose greater impression on the visitors, resulting in stronger memory and, consequently, enhancing the probability for the individual to visit the location in the future. That is, the attractiveness of a location will be reinforced by good memory and vice versa.

To characterize A in a quantitative manner, we rely on empirical evidence of human travel, in which the frequencies of

3

Table 1 Description of empirical data sets

Country	Data type	Number of cities	Population	Total traveling steps	λ	GDP per capita ($)
China (mainland)	Sina Weibo check-ins	340	1,571,056	4,976,255	35	8141
US (lower 48 states)	Foursquare check-ins	125	32,040	194,730	32	56,084
Cote d'Ivoire	Mobile phone CDR	237	229,335	8,747,801	50	1325
Belgium	Gowalla check-ins	43	1352	21,156	25	40,529

Information of data sets DS1–DS4 and the values of their memory parameter λ are shown
The four data sets correspond to four countries with different spatial scales, and the features of the four countries from the data sets include the number of cities, population, total traveling steps, and GDP per capita. Here, the GDP data is obtained from http://www.imf.org. An individual traveling from one city to another represents one travel step. The number of total traveling steps is the sum of all recorded individual steps

individual visit to different locations are distributed according to the Zipf's law[6]. It is thus reasonable to assume that A is distributed in a similar manner. That is, the Zipf's law stipulates that visitors rank the locations visited such that the probability to visit a location is inversely proportional to its rank. For example, the probability of visiting the most frequently visited location is $A_1 = \lambda/1$ and the probability of visiting the second most frequently visited location is $A_2 = \lambda/2$, and so on, where λ is a constant. Due to an aging effect, the most frequently visited location is often the "oldest" one. We can thus replace the rank of a location by the order with which it is visited. These considerations lead to the following formula to quantify the memory effect:

$$A_j = 1 + \frac{\lambda}{r_j}, \qquad (1)$$

where λ is a parameter characterizing the strength of the memory effect and the index r_j denotes that location j is the rth newly visited location associated with the movement trajectory. The unity in the formula represents the initial attractiveness of location j that has not been visited, i.e., if a location has not been visited, its A value is always unity.

Following the classic gravity model, we assume that the attractiveness of a location is proportional to its population. Let B be a quantity characterizing the population-induced effect on the attractiveness of a location, and let m_i be the population of location i and N be the total number of locations that can possibly be toured by all the individuals. As illustrated in Fig. 2, the attractiveness B_{ij} for a visitor to travel from location i to destination j is

$$B_{ij} = \frac{m_j}{W_{ji}}, \qquad (2)$$

where W_{ji} is the total population in the circular region centered at j, the radius of which is the distance between locations j and i. Note that B_{ij} reflects the competition for opportunities among different locations. For instance, if a traveler at location i wishes to visit a potential destination j, more populations between i and j imply more fierce competitions for limited opportunities at those locations, leading to a lower probability of being offered some opportunity. It is thus reasonable to assume that the attractiveness B_{ij} of destination j for a visitor from location i is inversely proportional to the population between i and j, as quantified by formula (2).

The transition probability p_{ij} of traveling from location i to j is then proportional to both A_j and B_{ij}, which can be written as

$$p_{ij} \propto \frac{m_j}{W_{ji}} \left(1 + \frac{\lambda}{r_j}\right). \qquad (3)$$

We see that the model contains a single adjustable parameter, the memory strength λ, that can be determined from empirical data. For any location i, we place a number of travelers. Each traveler is assigned a number L, the total number of movement steps, which can be obtained from an actual distribution from empirical data. A traveler thus executes a trajectory of length L and, at each step, he/she moves to a destination according to the transition probability p_{ij}.

Model prediction and validation. Our model, as illustrated in Fig. 2, is capable of predicting the statistical behaviors of human mobility at both the individual and population levels, regardless of the spatial scale. At the individual level, we focus on the following quantities: (I1) the total number of locations visited by time t, (I2) return time distribution to any location, (I3) distribution of frequency of visits to a location, and (I4) emergence of traveling motifs and their probability of occurrence in a long trajectory. At the population level, we seek to predict: (P1) distribution of the travel distance of collective movement and (P2) distribution of the number of traveling steps between two locations.

To validate the model predictions, we employ four empirical data sets, as illustrated in Fig. 1. They are: (DS1) record of user check-ins at Sina Weibo in mainland China (Fig. 1a), (DS2) check-in record of the site Foursquare[35] for users in the continental United States (Fig. 1b), (DS3) communication record of mobile phone users in the whole country of Cote d'Ivoire[36] (Fig. 1c), and (DS4) check-in record of the site Gowalla[37] in Belgium (Fig. 1d). Each data set contains spatial and temporal information about continuous user mobility, from which data of movements among various locations (e.g., cities) can be extracted (Methods). The single free parameter λ can be determined from data (Methods). We obtain $\lambda = 35$, 32, 50, and 25 for data sets DS1–DS4, respectively. A heuristic observation is that λ assumes a relatively smaller value for a better developed country (Methods). An explanation is that, in general, in a country with a higher gross domestic product (GDP), individuals can afford more travel, leading to more visited locations and a higher probability of exploring new places. In contrast, in a country with a lower GDP, it is more difficult for people to travel frequently and they tend to stay in their home cities and familiar places. That is, a higher GDP induces a weaker memory effect and a higher probability of visiting new locations, as reflected by the smaller values of the memory strength λ in well developed countries.

Since the data sets contain no information about the individuals' cities of residence, for each individual, we assign the city that he/she signs in with the highest frequency as his/her home city and use it as the initial location in the model. From the data, we calculate the distribution $P(L)$ of the total number of times of movement and choose L accordingly, which is effectively the trajectory length for each individual. It is worth noting that an effective way to test our mobility model is to use the same distribution of the trajectory length as that from the empirical data. We also study analytically the impact of trajectory length on the statistical properties of mobility at both the population and individual levels, with the finding that, for a sufficient number of moving steps, simulation results are in good agreement with the
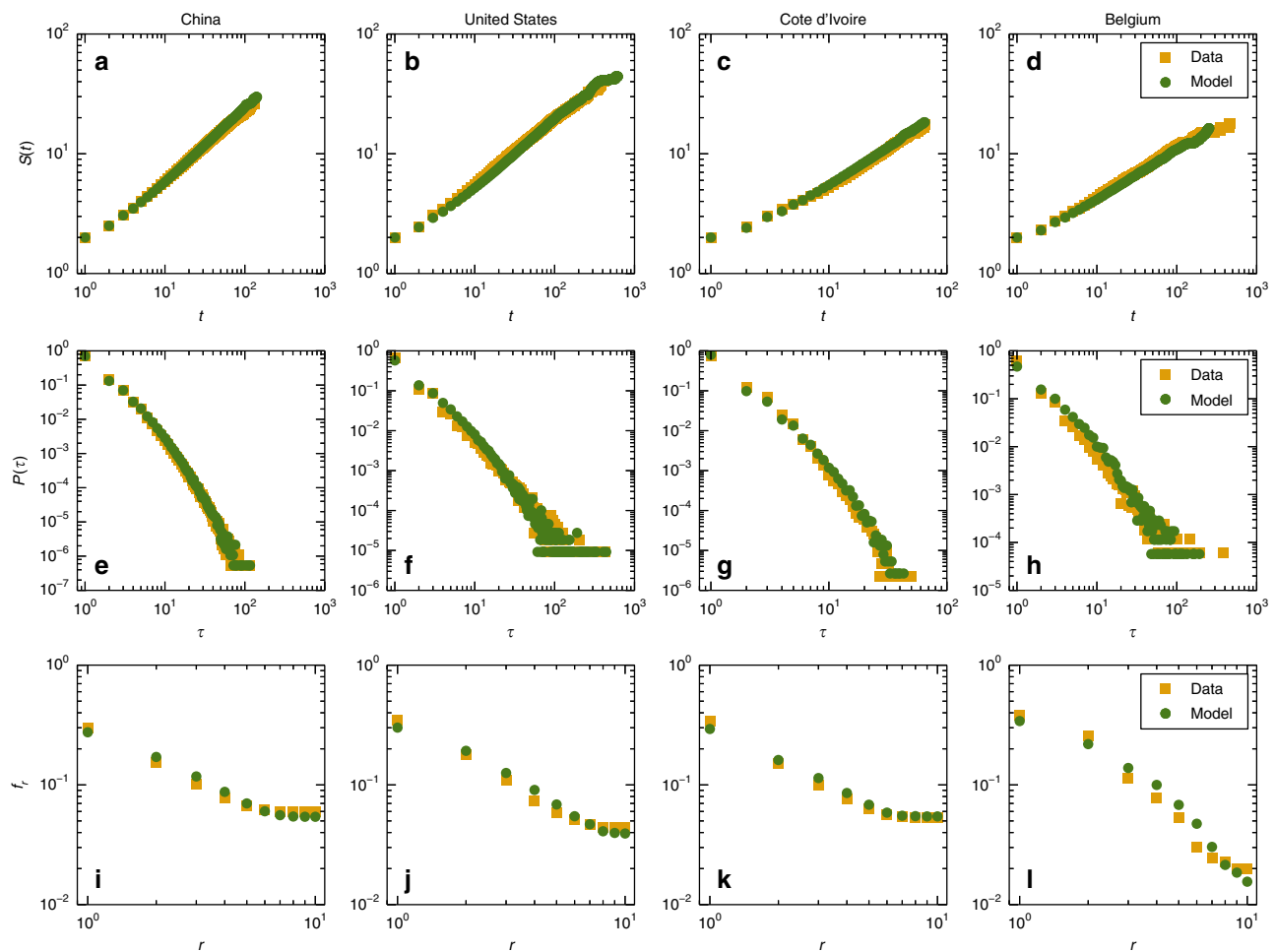
Fig. 3 Model-predicted and empirical statistical behaviors of human mobility at the level of individual trajectories. **a–d** Algebraic increase with time t in the total number of locations visited within t, **e–h** algebraic decaying behavior in the return time distribution, and **i–l** Zipf-like frequency distribution of visits. In all panels, the green color specifies results predicted by our model, and orange denotes the empirical results obtained directly from data. There is a generally strong agreement between the two types of results. The algebraically increasing and decreasing behaviors in **a–d** and **e–h**, respectively, are manifestations of the memory effect

analytical prediction. This indicates that the trajectory length has little effect on the mobility patterns in the long time regime. For the empirical data (Table 1), the total numbers of steps are much larger than those on the population records, so the mobility patterns produced by our model are stable and robust.

Figure 3a–d show, for the data sets DS1–DS4, respectively, the model-predicted algebraic increase with t in the total number of locations visited by time t (green), together with the corresponding results calculated directly from the data (orange). We obtain an excellent agreement between model prediction and the empirical result. The algebraically increasing behavior, as opposed to an exponential growth in the number of cities visited in certain time, is a natural consequence of the memory effect, which is a key ingredient in our model. Figure 3e–h show the predicted and actual return time distributions for the data sets DS1–DS4, respectively, which are algebraic. There is again an excellent agreement between the model prediction and the empirical result. The algebraic decay in the return time distribution can also be attributed to the memory effect. Thus, both the algebraically increasing behavior in Fig. 3a–d and the algebraically decaying behavior in Fig. 3e–h are manifestations of the same memory effect. Figure 3i–l show, for the data sets DS1–DS4, respectively, the model-predicted and empirical frequency distributions of

visits to all locations, which agree with each other reasonably well and follow approximately the Zipf's law. The emergence of the Zipf-like scaling behavior is indicative of the heterogeneity in the location attractiveness, an assumption of our model. The results in Fig. 3a–l validate our model with respect to the statistical behaviors of individual trajectories.

A characteristic of human mobility is the emergence of motifs associated with movement trajectories[38], which are referred to as certain simple and fixed patterns of visit that occur repeatedly in a long trajectory. For an individual initially at his/her home location (the one visited with the highest frequency), a motif is defined as a successive sequence of locations visited with the last location being the initial one. From the empirical data, we identify nine distinct motifs (shown at the top in Fig. 4) and calculate the frequencies of their occurrences from the entire data set. With parameter λ extracted from the data, our model can generate long trajectories from which the possible motifs and their frequencies of occurrence can be determined. Remarkably, our model yields exactly the same set of motifs with frequencies that agree with the empirical results reasonably well, as shown in Fig. 4. Due to the significance of travel motifs in determining the microscopic mobile patterns of travelers, the agreement provides further validation of our model at the individual level.
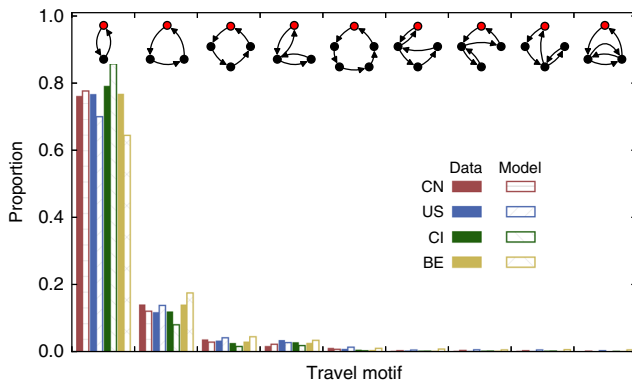
Fig. 4 Frequency of occurrence of motifs associated with movement trajectories. From the four empirical data sets, nine distinct motifs contained in movement trajectories are identified. Exactly the same motifs are obtained from model generated trajectories. The frequencies of occurrence of the motifs from the model and from actual trajectories agree with each other reasonably well. There exist many more types of possible motifs. However, the nine types included here have the highest frequencies of occurrence, the sum of which exceeds 0.97. The frequency of the 10th highest motif, for example, is <0.001. Considering the nine motifs thus suffices

Our model also has strong predictive power for human movements at the population level on all spatial scales. As shown in Fig. 5, the predicted behaviors of $P(d)$ and $P(T)$, the distributions of the travel distance and of the number of traveling steps between two locations, agree well with the statistical results from the empirical data. In particular, Fig. 5a–d show, for the data sets DS1–DS4, respectively, that $P(d)$ decays exponentially. Figure 5e–h reveal that $P(T)$ exhibits a robust algebraic scaling for all four data sets. Figure 5i–l demonstrate that the model-predicted and real values of T are nearly statistically indistinguishable (albeit with fluctuations). Our model is then universally applicable to characterizing human movements across vast different spatial scales at the population level.

Theoretical analysis. In our model, the fixed amount of attractiveness of a location is calculated based on its population. Since the population distribution is typically highly heterogeneous without a closed mathematical form, it is not feasible to treat our model exactly. To gain analytic insights, we consider a simplified model obtained by imposing the approximation that the population is uniformly distributed among the locations, and focus on analytically predicting the individual trajectories and the collective mobility pattern with a special emphasis on the role of the memory effect. Although the simplified model deviates from real scenarios, the analytical predictions enable a good understanding of the real mobility patterns at both the individual and population levels.

In the simplified model, an individual travels among N locations. At each time step, the probability to move to a destination is proportional to its attractiveness. At $t = 0$, the initial attractiveness is identical (unity) for all locations. During the travel, the attractiveness of the rth first visited location is updated to $1 + \lambda N/r$, where $\lambda > 0$ is a parameter. The model describes essentially a random-walk process with time varying location attractiveness, with parameter λ characterizing the memory strength of (or preference to) locations previously visited. For $\lambda = 0$, the model is reduced to an unbiased random walk. For $\lambda = \infty$, the walker can travel between only the first two locations. The total number $S(t)$ of locations visited by time t can be used to characterize how fast the underlying mobile process takes place.

For a uniform random walk, $S(t)$ increases with t linearly: $S(t) \propto t$. For the EPR model[2], $S(t)$ increases with t but in a sublinear fashion: $S(t) \propto t^\beta$ with $0 < \beta < 1$. For our random walk model with memory, $S(t)$ can be obtained analytically (Supplementary Note 1), as shown in Fig. 6a. We see that, as the memory strength parameter λ is increased, the overall rate of increase in $S(t)$ becomes smaller. In addition, for a fixed value of λ, the time derivative of $S(t)$ tends to increase with time, which is consistent with the result from real data (c.f., Figs. 1a and 3 in ref. [2]).

Another characteristic quantity is f_r, the distribution of the frequency of visit to location r. For an unbiased random walk, f_r is uniformly distributed. For the EPR model[2], f_r decays algebraically: $f_r \propto r^{-\alpha}$, where $\alpha > 0$ is a constant. For our model, we analytically obtain (Supplementary Note 1)

$$f_r \propto \frac{\lambda S}{r} + 1 - \lambda. \qquad (4)$$

For $\lambda = 0$, Eq. (4) reduces to a uniform distribution. For $\lambda = 1$, we recover the Zipf's law for f_r. Figure 6b shows the analytic and simulation results of f_r for a number of λ values, where the curves represent the theoretical prediction. We see that, for $\lambda > 1$, there is an apparent deviation from the Zipf's law, as signified by the emergence of an exponential cutoff toward the tail end of f_r. The physical meaning is that, as the memory effect is intensified, a walker tends to travel among only a few locations.

The return time distribution $P(\tau)$ is defined as the probability for a walker to return to one of the previously visited locations after τ steps, which is also a reflection of the memory effect. In our model, $P(\tau)$ contains two different algebraic terms but with the same exponent -1 (Supplementary Note 1). As λ is increased, $P(\tau)$ tends to a single algebraic distribution with essentially zero values near the tail, indicating an extremely low probability for the walker to return to a previous location after many time steps. Figure 6c shows the analytic and simulation results of $P(\tau)$ from our model. The agreement is reasonable, and the deviation of the analytic from the simulation result in the large τ region is due to the finite time used in the simulation.

Finally, we remark on an appealing feature of our model. Consider the probability for the walker to choose a new location at the next time step. Analysis of our model leads to (Supplementary Note 1)

$$P_{\mathrm{new}} = \frac{1}{1 + \lambda(\ln S + C)}. \qquad (5)$$

Thus, in our model, P_{new} decreases with S, which occurs naturally as a consequence of a basic and intuitive assumption, namely the memory effect.

A further simplification of the model by assuming that each individual can move one step only renders analytically predictable collective mobility patterns at the population level. In particular, we place m individuals at each location and exploit the previously discovered[39], common fractal feature in the spatial distribution of locations in the real world: $W_{ji} \propto d_{ij}^D$, where D is the fractal dimension. Equation (3) can be formulated as (Supplementary Note 2)

$$T_{ij} = m_i p_{ij} \propto \frac{m_i m_j}{d_{ij}^D}, \qquad (6)$$

where T_{ij} is the total number of traveling steps from i to j for the whole population. Equation (6) is a standard gravity model with a power-law distance function. Since Eq. (6) indicates that the number of traveling steps T between two locations of distance d is $T(d) \propto d^{-D}$, the travel distance distribution is given by the same form: $P(d) \propto d^{-D}$, as validated by Fig. 6d for four typical fractal domains. The number of location pairs with distance $\leq d$ in a
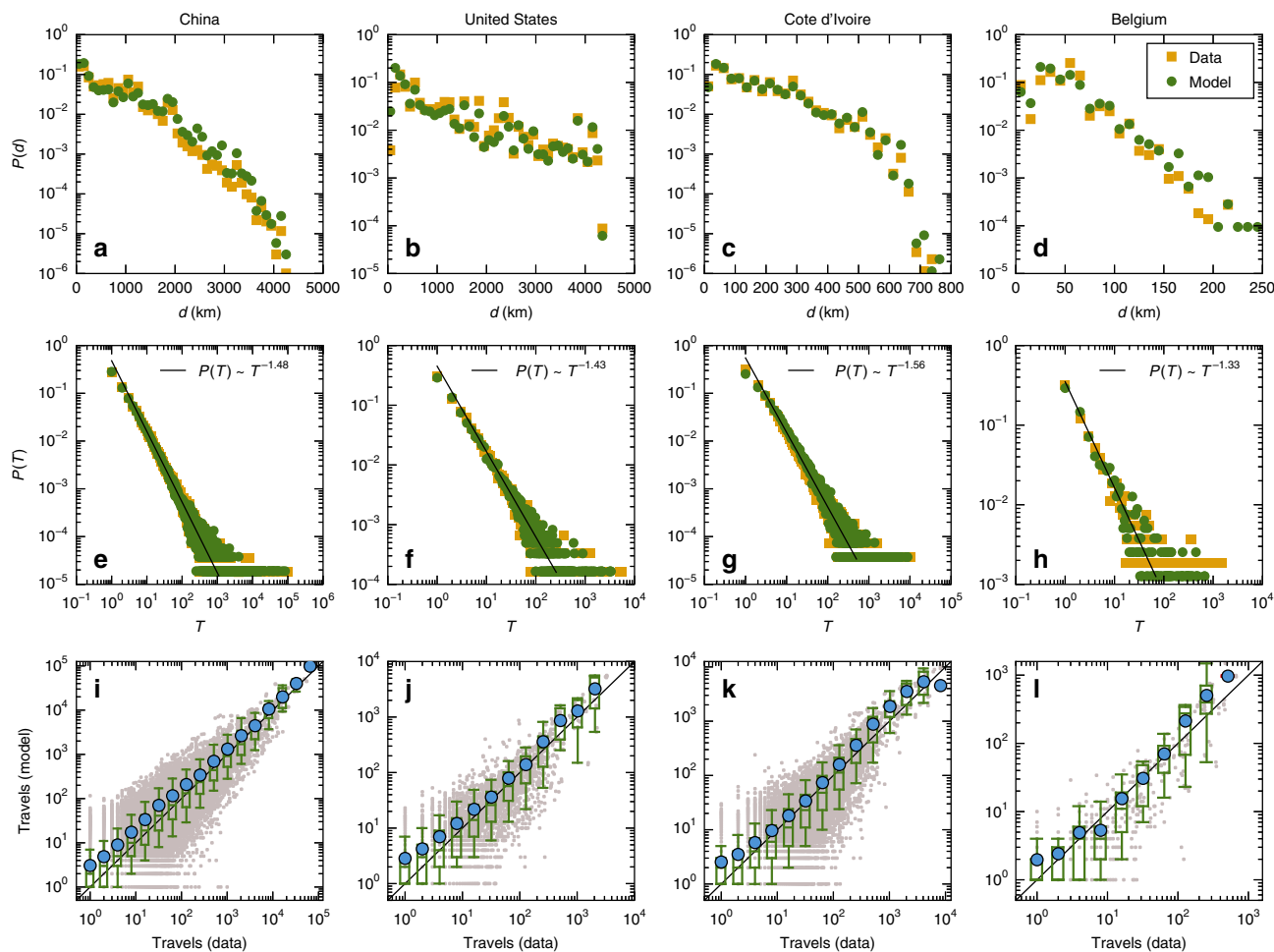
Fig. 5 Model prediction at the population level and validation. For the four empirical data sets with vast difference in the spatial scale, model-predicted and real distributions of travel distance d (**a–d**) and the number of traveling steps, T, between two locations **e–h**. The lines in **e–h** represent a power-law fit to the real data. **i–l** Statistical display of the model-predicted and real values of T for the four data sets, respectively, which are nearly indistinguishable. The gray points are scatter plot for each pair of locations. The blue points represent the average number of predicted travels in different bins. The standard boxplots represent the distribution of the number of predicted travels in different bins of the number of observed travels. A box is marked in green if the line $y = x$ lies between 10% and 91% in that bin and in red otherwise

fractal domain is $N(d) \propto d^D$. Thus, the number of traveling steps T obeys a power-law distribution (Supplementary Note 2):

$$P(T) \propto T^{-2}. \qquad (7)$$

It is worth noting that the algebraic exponent −2 is universal, regardless of the fractal dimension D of the location distribution in the simplified model, as shown in the insert of Fig. 6d. However, in the real world, the heterogeneous nature of the population distribution at different locations can cause a deviation of the exponent from −2. As shown in Fig. 5e–h, the fit of the empirical data demonstrates that their algebraic exponents range from −1.33 to −1.56 (see Supplementary Note 2 for a detailed explanation of the effect of heterogeneous population distribution on the algebraic exponent). Nonetheless, the power-law distribution predicted by our simplified model is robust, which captures the essential features of the collective mobility patterns in the real world.

In the development of our human mobility model, Zipf's law is naturally included as an essential component. However, the Zipf's law is closely related to diminishing exploration. To elucidate the interplay between the two, we articulate an extended individual mobility model based on the generalized Zipf's law. This is guided by the previous evidence that there are situations where individuals tend to choose locations to travel into by following the generalized Zipf's law[40]. Specifically, we assume $f \propto r^{-\zeta}$, where the exponent $\zeta > 1$ is an adjustable parameter. For the extended model, we analytically obtain $P_{\text{new}} = \rho S^{-\gamma}$, where $\gamma = \zeta - 1$. We see that the formula of P_{new} is a direct manifestation of the basic assumptions in the EPR model[2]. This suggests that the generalized Zipf's law and the power-law relation between P_{new} and S have a mutually causal relationship, and the individual mobility models based on the former and latter are equivalent to each other. For the extended model, we also derive the return time distribution $P(\tau)$ for sufficiently large values of S. A detailed description of the extended model, the analysis, and results are presented in Supplementary Note 1.

Discussion

The past decade has witnessed a great deal of efforts into uncovering and understanding the general dynamical behaviors of human mobility. A variety of real data sets have been analyzed, leading to a spectrum of mathematical models being devised to explain the phenomena revealed by data. While universal scaling laws have been unveiled, it turns out that spatial scale has a significant effect on the dynamics. In particular, human
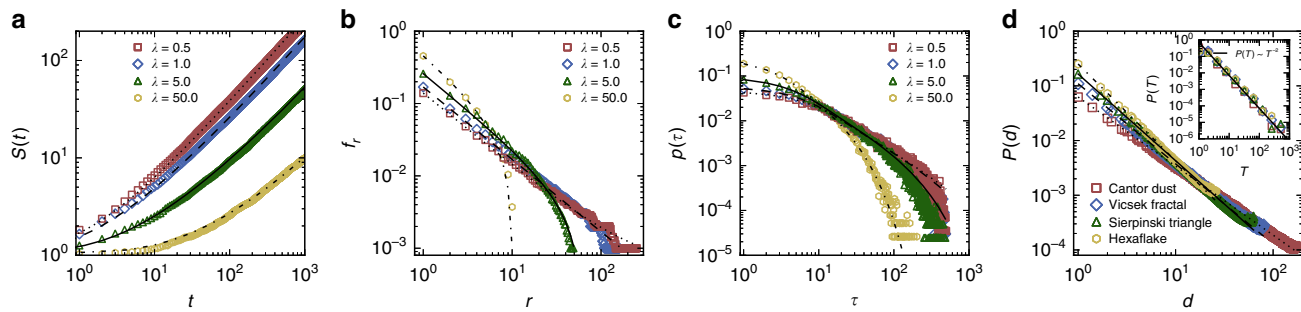
Fig. 6 Analytic predictions of a simplified model in comparison with simulation results. **a–c** Analytically predicted results of $S(t)$—the total number of locations visited by time t, f_r—frequency distribution of visiting distinct locations, and $P(\tau)$—return time distribution. The curves represent the theoretical results for different values of the memory parameter λ, and the data points are simulation results with $N = 1000$ and $L = 1000$. **d** Travel distance distributions $P(d)$ for the simplified model implemented in four classic fractal domains including a 2D Cantor dust, Vicsek fractal, Sierpinski triangle, and hexaflake with fractal dimension $D = 1.262$, 1.465, 1.585, and 1.771, respectively (Supplementary Note 3), where the lines represent the theoretical results for different values of D. The inset indicates the distributions $P(T)$ of the number of traveling steps at the population level for the four fractal domains, with the solid line being the analytic prediction of $P(T)$

mobility at large (e.g., big countries) and small (e.g., small countries or big cities, see Supplementary Note 4 for a city example) scales tends to exhibit distinct scaling behaviors. The representative existing models are suitable to describe human mobility on either large or small scales at either the individual or population level, motivating us to articulate a model that can describe the statistical and scaling behaviors of human behaviors at all spatial scales as well as at both the individual and population levels.

There are two essential ingredients in our model construction: memory and population-induced competition effects. Both effects jointly determine the attractiveness of a location (see Supplementary Note 5 for results and a detailed discussion). On the basis of the attractiveness of locations, we obtain the key quantity in microscopic model of human mobility: the transition probability for an individual to move from one location to another. Our unifying model contains a single adjustable parameter: the strength of the memory effect, and enables us to make predictions about the scaling laws associated with the key statistical behaviors of human mobility at both the individual and population levels, regardless of the spatial scales. The relevant quantities include the total number of locations visited within certain time, the frequency distribution of visits to different locations, and the distribution of time interval of successive visits to any location. Our model also allows us to identify a few kinds of distinct motifs embedded in typical trajectories. All these results have been verified using empirical data from countries having drastically different spatial scales.

Modeling and predicting mobility patterns and scaling laws at both the individual and population levels are a fundamental problem for exploring many dynamical processes associated with human mobility. A typical example is disease propagation in the society. As discussed in ref. [41], in the metapopulation model, both the population level mobility, e.g., travel flux among subpopulations, and the individual level mobility, e.g., transition probability of individuals, are necessary to model the contagion dynamics and predict disease spreading in the society. The empirical mobilities may be obtained by directly measuring the travel flux among locations and the travel trajectories of all individuals in a certain time interval. However, to accomplish this task, vast amounts of private data, such as the data of cell phones with GPS function in specific locations, are required, making the task impractical. Our universal model, because it is based solely on the population distribution, provides an alternative approach to unraveling the important mobility patterns with reasonable accuracy. Likewise, our model may find potential use in

alleviating congestion in urban areas, which is closely related to human mobility behaviors.

Methods

Empirical data sets and processing method. The four data sets DS1–DS4 are from mainland China (Supplementary Data 1 and 2), the contiguous United States, Cote d'Ivoire, and Belgium, respectively. Sets DS1, DS2, and DS4 are the check-in records of social networks[35,37] in their respective countries, which contain the time and locations of user check-ins. Set DS3 is a mobile phone call detail record[36] that collects the time and positions of users making phone calls or sending text messages in a 5-month period, where the spatial locations are determined within counties. In this case, the central city of each county is taken as the location of the individual. Since we focus on movements among cities, all the positions within a city are regarded as the same with an identical city label. Table 1 lists the detailed information about each data set, from which a complete trajectory of each user moving among different cities can be obtained for the entire time duration of the data record. The results of statistical analysis of individual trajectories are shown in Fig. 3, while those at the population level are presented in Fig. 5.

Parameter estimation. In our model, the single free parameter is λ, the strength of memory effect, which affects directly the rate of increase in the number $S(t)$ of locations visited in certain time. For a given empirical data set, the function $S(t)$ can then be used to estimate λ. To accomplish this, we define the following objective function:

$$E(\lambda) = \sum_{t=1}^{L_{\max}} \frac{|S_{\text{real}}(t) - S(t, \lambda)|}{S_{\text{real}}(t)}, \qquad (8)$$

where $L_{\max}$ is the maximum time step, $S_{\text{real}}(t)$ is obtained from the actual data set, and $S(t, \lambda)$ is calculated through the model with parameter λ. The objective function can be minimized to yield an estimated value of λ in the model. We also use the quantities $P(\tau)$ and f_r to estimate λ in addition to that based on $S(t)$, and find little difference in the prediction accuracy for both the individual and population mobility patterns.

References

1. Brockmann, D., Hufnagel, L. & Geisel, T. The scaling laws of human travel. *Nature* **439**, 462–465 (2006).
2. Song, C., Koren, T., Wang, P. & Barabási, A. L. Modelling the scaling properties of human mobility. *Nat. Phys.* **6**, 818–823 (2010).
3. Simini, F., González, M. C., Maritan, A. & Barabási, A. L. A universal model for mobility and migration patterns. *Nature* **484**, 96–100 (2012).

4. Simini, F., Maritan, A. & Néda, Z. Human mobility in a continuum approach. *PLoS ONE* **8**, e60069 (2013).
5. Ren, Y., Ercsey-Ravasz, M., Wang, P., González, M. C. & Toroczkai, Z. Predicting commuter flows in spatial networks using a radiation model based on temporal ranges. *Nat. Commun.* **5**, 5347 (2014).
6. González, M. C., Hidalgo, C. A. & Barabási, A. L. Understanding individual human mobility patterns. *Nature* **453**, 779–782 (2008).
7. Eagle, N., Macy, M. & Claxton, R. Network diversity and economic development. *Science* **328**, 1029–1031 (2010).
8. Hu, Y., Zhang, J., Huan, D. & Di, Z.-R. Toward a general understanding of the scaling laws in human and animal mobility. *Europhys. Lett.* **96**, 38006 (2011).
9. Noulas, A., Scellato, S., Lambiotte, R., Pontil, M. & Mascolo, C. A tale of many cities: universal patterns in human urban mobility. *PLoS ONE* **7**, e37027 (2012).
10. Lenormand, M., Huet, S., Gargiulo, F. & Deffuant, G. A universal model of commuting networks. *PLoS ONE* **7**, e45985 (2012).
11. Goh, S., Lee, K., Park, J. S. & Choi, M. Y. Modification of the gravity model and application to the metropolitan Seoul subway system. *Phys. Rev. E* **86**, 026102 (2012).
12. Yan, X.-Y., Han, X.-P., Wang, B.-H. & Zhou, T. Diversity of individual mobility patterns and emergence of aggregated scaling laws. *Sci. Rep.* **3**, 02678 (2013).
13. Saramäki, J. et al. Persistence of social signatures in human communication. *Proc. Natl. Acad. Sci. USA* **111**, 942–947 (2013).
14. Hou, L., Pan, X., Guo, Q. & Liu, J.-G. Memory effect of the online user preference. *Sci. Rep.* **4**, 06560 (2014).
15. Zhao, Z.-D., Huang, Z.-G., Huang, L., Liu, H. & Lai, Y.-C. Scaling and correlation of human movements in cyber and physical spaces. *Phys. Rev. E* **90**, 050802(R) (2014).
16. Yan, X.-Y., Zhao, C., Fan, Y., Di, Z.-R. & Wang, W.-X. Universal predictability of mobility patterns in cities. *J. R. Soc. Interface* **11**, 20140834 (2014).
17. Šćepanović, S., Mishkovski, I., Hui, P., Nurminen, J. K. & Ylä-Jääski, A. Mobile phone call data as a reginal socio-economic proxy indicator. *PLoS ONE* **10**, e0124160 (2015).
18. Pappalardo, L. et al. Returners and explorers dichotomy in human mobility. *Nat. Commun.* **6**, 8166 (2015).
19. Gallotti, R., Bazzani, A., Rambaldi, S. & Barthelemy, M. A stochastic model of randomly accelerated walkers for human mobility. *Nat. Commun.* **7**, 12600 (2016).
20. Zhao, Y.-M., Zeng, A., Yan, X.-Y., Wang, W.-X. & Lai, Y.-C. Unified underpinning of human mobility in the real world and cyberspace. *New J. Phys.* **18**, 053025 (2016).
21. Frías-Martínez, E., Williamson, G. & Frías-Martínez, V. *IEEE 3rd Conference on Privacy, Security, Risk and Trust (PASSAT) and 2011 IEEE Third Inernational Conference on Social Computing (SocialCom)* 57–64 (IEEE, 2011).
22. Belik, V., Geisel, T. & Brockmann, D. Nature human mobility patterns and spatial spread of infectious diseases. *Phys. Rev. X* **1**, 011001 (2011).
23. Wesolowski, A. et al. Quantifying the impact of human mobility on malaria. *Science* **338**, 267–270 (2012).
24. Jiang, S et al. In *Proceedings of the 2nd ACM SIGKDD International Workshop on Urban Computing (UrbComp'13)* 57–64 (ACM Press, New York, 2013).
25. Tizzoni, M. et al. On the use of human mobility proxies for modeling epidemics. *PLoS ONE* **10**, e1003716 (2014).
26. Deville, P. et al. Dynamic population mapping using mobile phone data. *Proc. Natl. Acad. Sci. USA* **111**, 15888–15893 (2014).
27. Berlingerio, M. et al. *Machine Learning and Knowledge Discovery in Databases Lecture Notes in Computer Science 8190* 663–666 (2013).
28. Lima, A., Domenico, M. D., Pejovic, V. & Musolesi, M. Disease containment strategies based on mobility and information dissemination. *Sci. Rep.* **5**, 10650 (2015).
29. Song, C., Qu, Z., Blumm, N. & Barabási, A. L. Limits of predictability in human mobility. *Science* **327**, 1018–1021 (2010).
30. Deville, P. et al. Scaling identity connects human mobility and social interactions. *Proc. Natl. Acad. Sci. USA* **113**, 7047–7052 (2016).
31. Zipf, G. K. The $p_1 p_2 / d$ hypothesis:on the intercity movement of persons. *Am. Sociol. Rev.* **11**, 677–686 (1946).
32. Han, X.-P., Hao, Q., Wang, B.-H. & Zhou, T. Origin of the scaling law in human mobility: hierarchy of traffic systems. *Phys. Rev. E* **83**, 036117 (2011).
33. Szell, M., Sinatra, R., Petri, G., Thurner, S. & Latora, V. Understanding mobility in a social petri dish. *Sci. Rep.* **2**, 00457 (2012).
34. Zhao, Z.-D. et al. Emergence of scaling in human-interest dynamics. *Sci. Rep.* **3**, 3472 (2013).
35. Levandoski, J. J., Sarwat, M., Eldawy, A. & Mokbel, M. F. *IEEE 28th International Conference on Data Engineering* 450–461 (2012).
36. Blondel, V. D., Decuyper, A. & Krings, G. A survey of results on mobile phone datasets analysis. *EPJ Data Science* **4**, 10 (2015).
37. Cho, E., Myers, S. A. & Leskovec, J. *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* 1082–1090 (2011).
38. Schneider, C. M., Belik, V., Couronné, T., Smoreda, Z. & González, M. C. Unravelling daily human mobility motifs. *J. R. Soc. Interface* **10**, 20130246 (2013).
39. Appleby, S. Multifractal characterization of the distribution pattern of the human population. *Geogr. Anal.* **28**, 147–160 (1996).
40. Wyllys, R. E. Empirical and theoretical bases of Zipf's law. *Libr. Trends* **30**, 53–64 (1981).
41. Balcan, D. & Vespignani, A. Phase transitions in contagion processes mediated by recurrent mobility patterns. *Nat. Phys.* **7**, 581–586 (2011).

Author contributions

X.-Y.Y., W.-X.W., Z.-Y.G. and Y.-C.L. designed the research; X.-Y.Y. and W.-X.W. performed the research; X.-Y.Y., W.-X.W. and Z.-Y.G. contributed analytic tools; W.-X.W., Z.-Y.G. and Y.-C.L. analyzed the data; and X.-Y.Y., W.-X.W., Z.-Y.G. and Y.-C.L. wrote the paper.

Additional information

Supplementary Information accompanies this paper at doi:10.1038/s41467-017-01892-8.

Competing interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
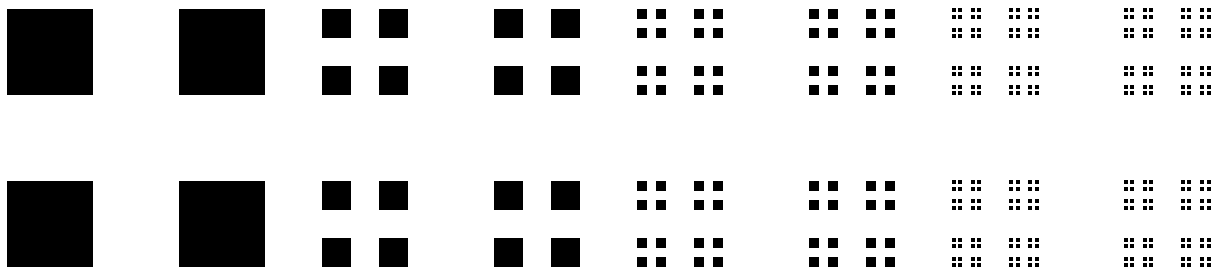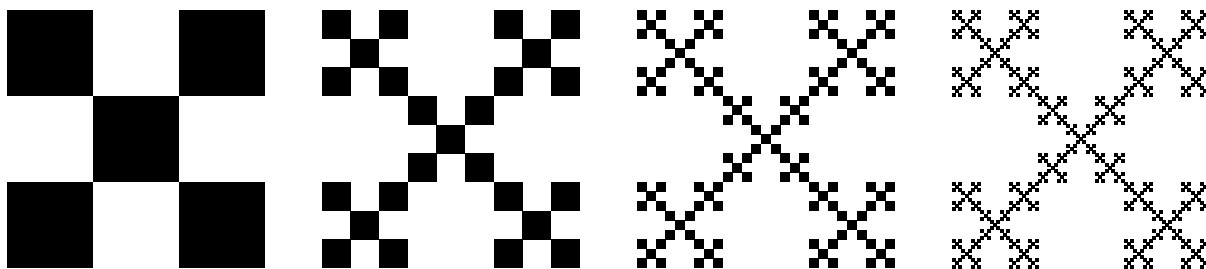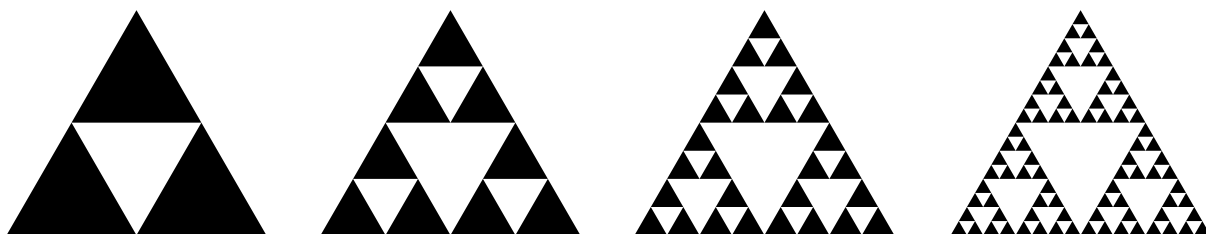
SUPPLEMENTARY INFORMATION
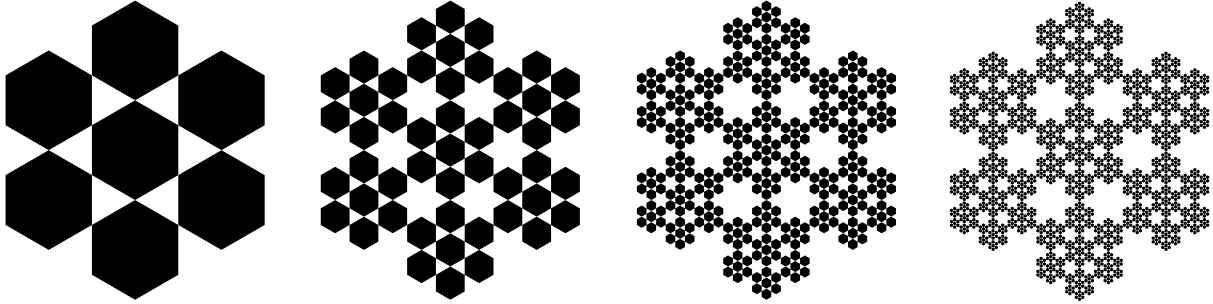
Supplementary Figures

Supplementary Figure 1 | The first 4 iterations in the generation of a 2D Cantor dust.
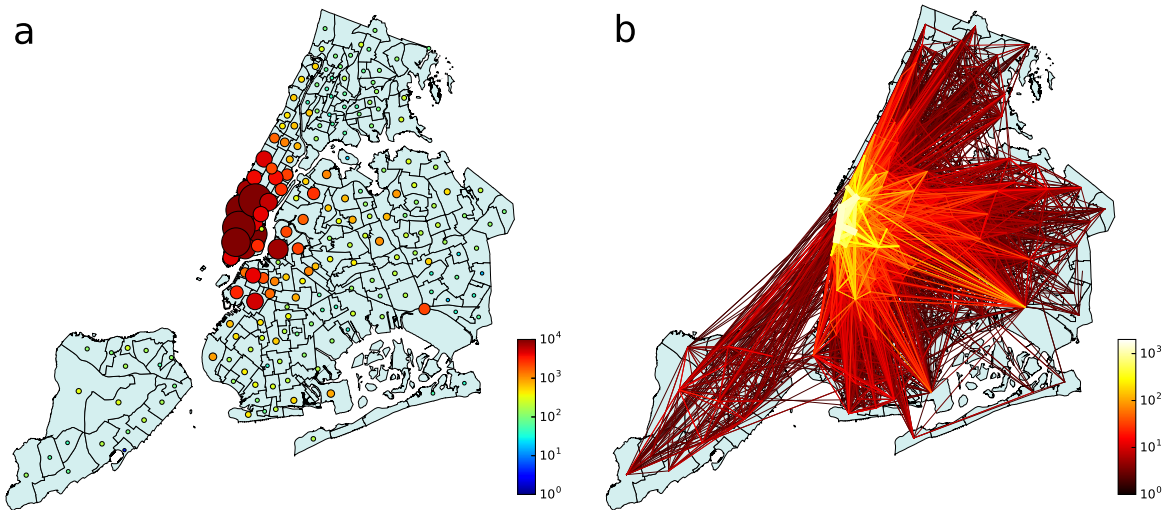

Supplementary Figure 2 | The first 4 iterations of the Vicsek fractal.

Supplementary Figure 3 | The first 4 iterations of the Sierpinski triangle.

**Supplementary Figure 4 | The first 4 iterations of the hexaflake.**

Supplementary Figure 5 | Foursquare users check-ins and mobility data in New York city. (a) Check-in times distribution. The zones are U.S. 2010 census blocks. (b) Traveling steps between locations.
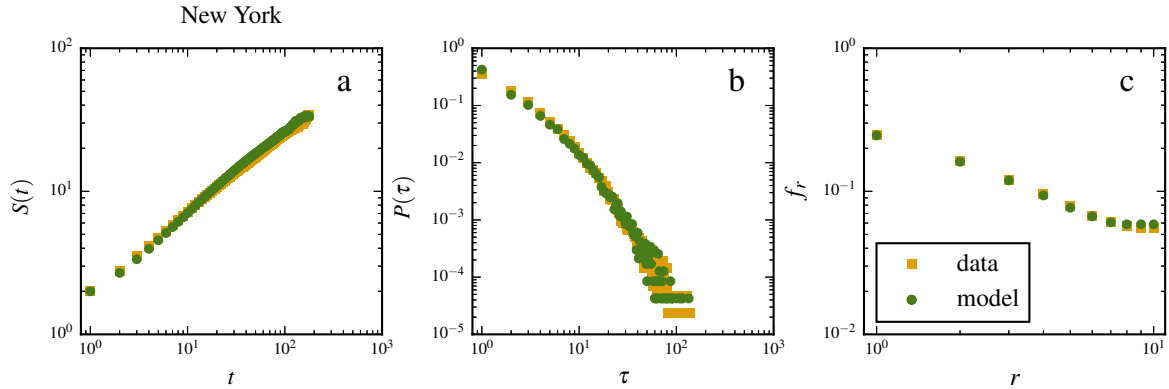
Supplementary Figure 6 | Model predicted and empirical statistical patterns of individual mobility in New York city. (a) The total number $S(t)$ of locations visited as an increasing function of time t, (b) the return time distribution $P(\tau)$, and (c) Zipf-like frequency distribution of location visits. In all panels, the green color specifies results predicted from our model, and orange denotes the empirical results obtained directly from data.
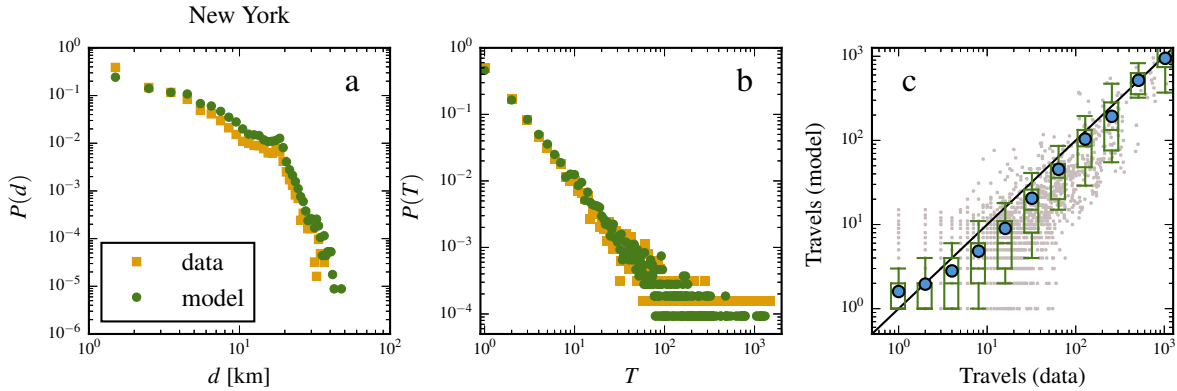
Supplementary Figure 7 | Model predicted and empirical statistical patterns of collective mobility in New York city. Model predicted (green circles) and real (orange squares) distributions of (a) traveling distance d and (b) the number of traveling steps, T, between two locations. (c) Model predicted versus real values of T.
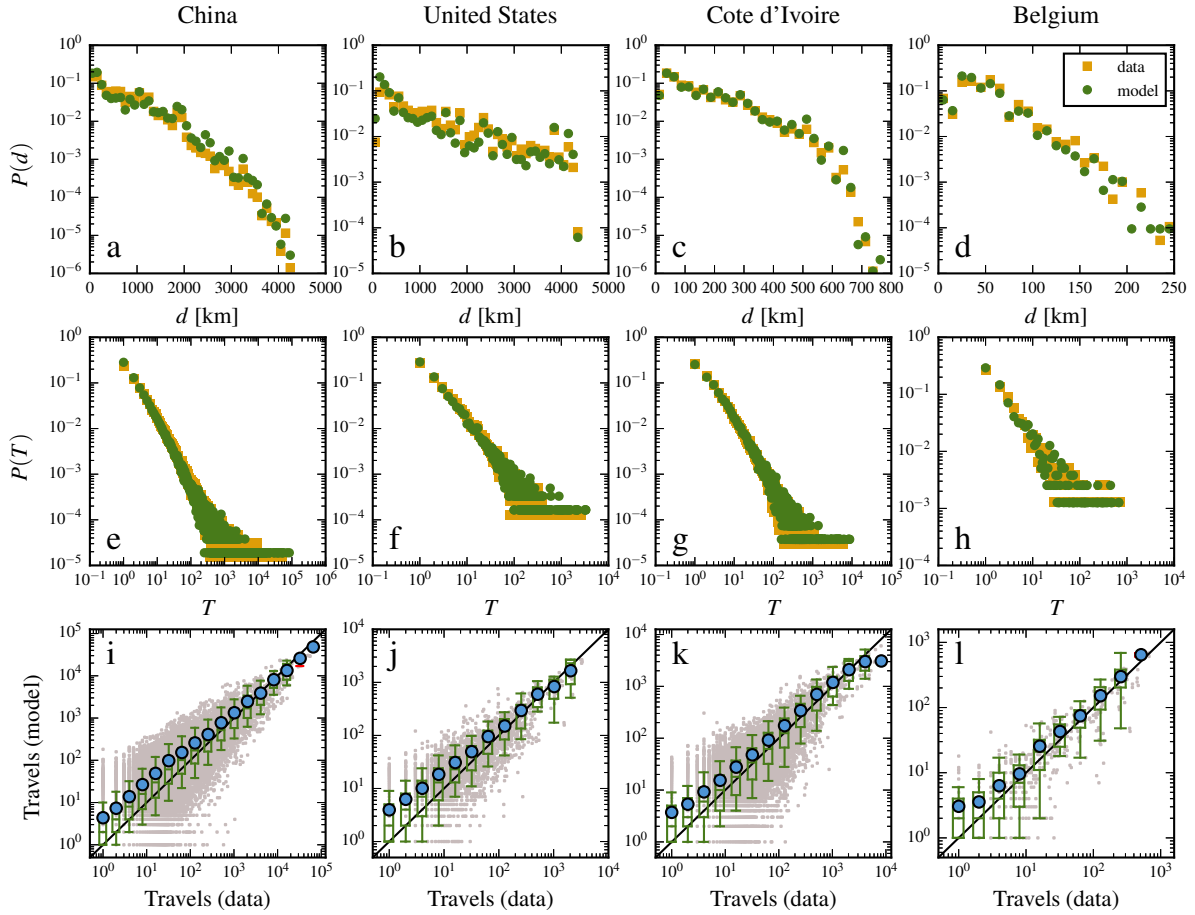
Supplementary Figure 8 | Collective mobility patterns predicted by the memory-free model.
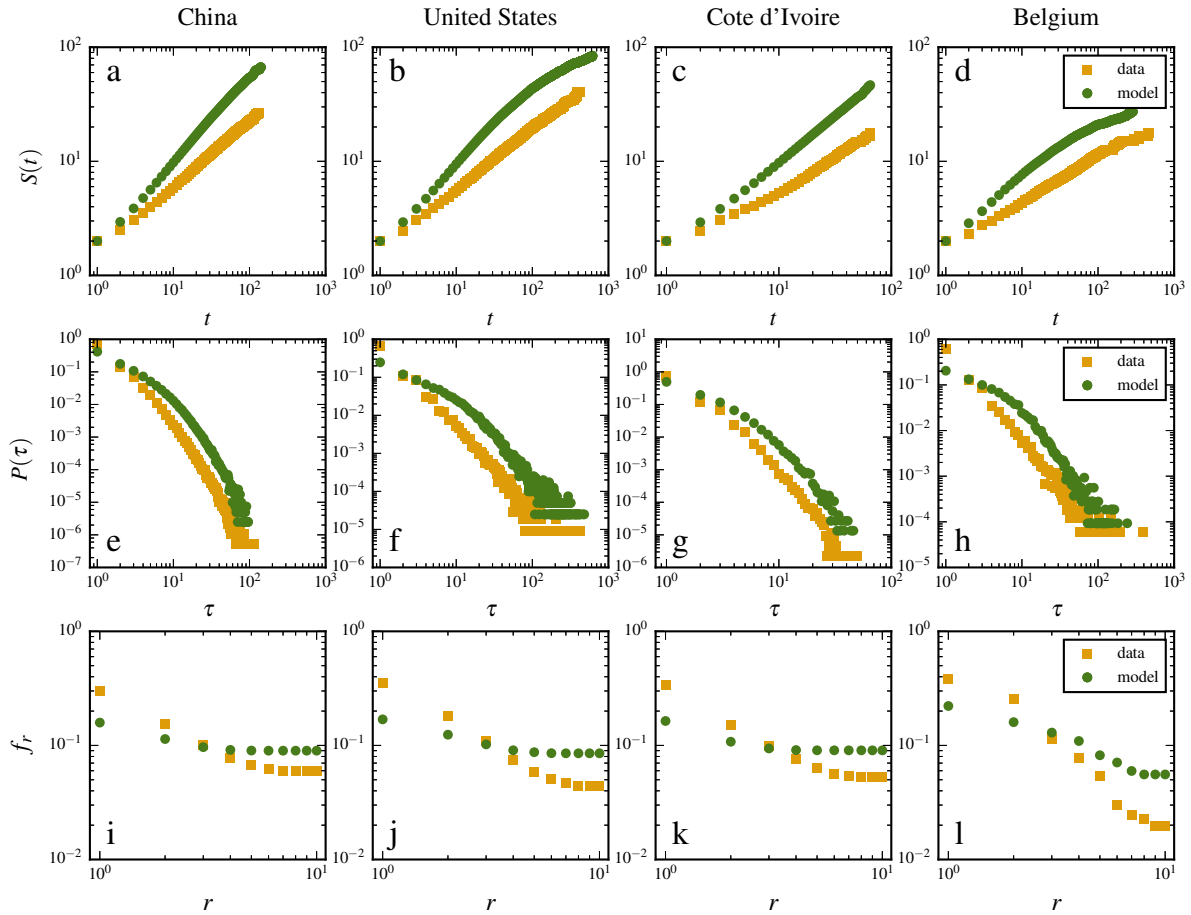
Supplementary Figure 9 | Individual mobility patterns predicted by the memory-free model.

5
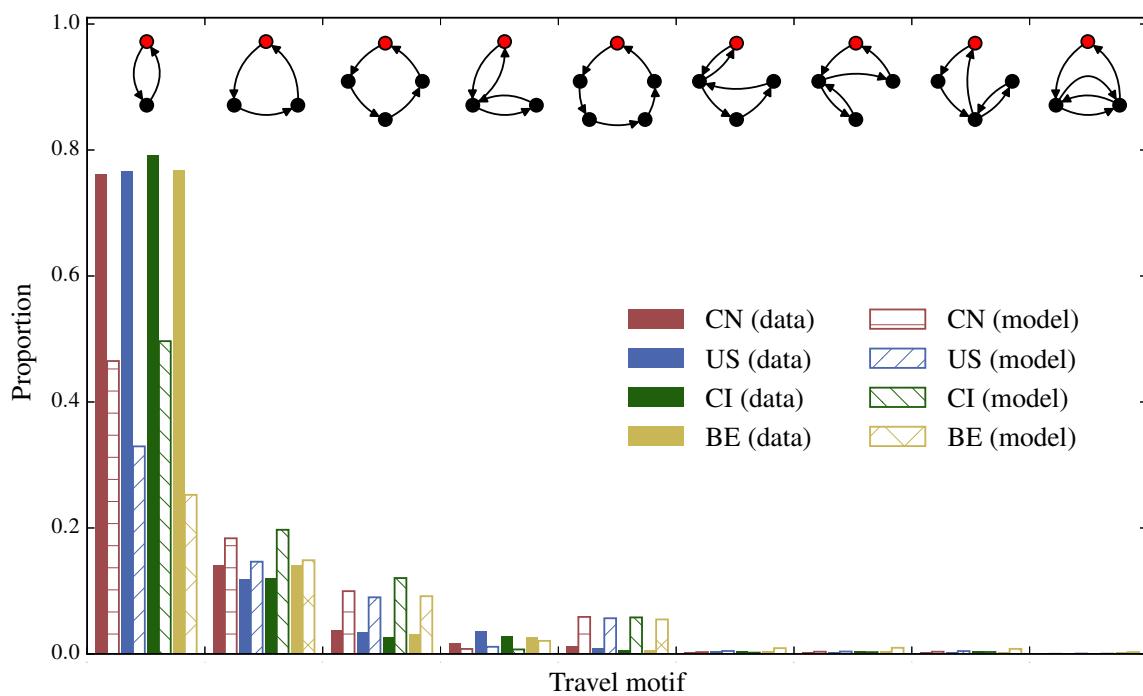
Supplementary Figure 10 | Frequency of occurrence of motifs in the memory-free model.

Supplementary Figure 11 | Collective mobility patterns predicted by the competition-free model.
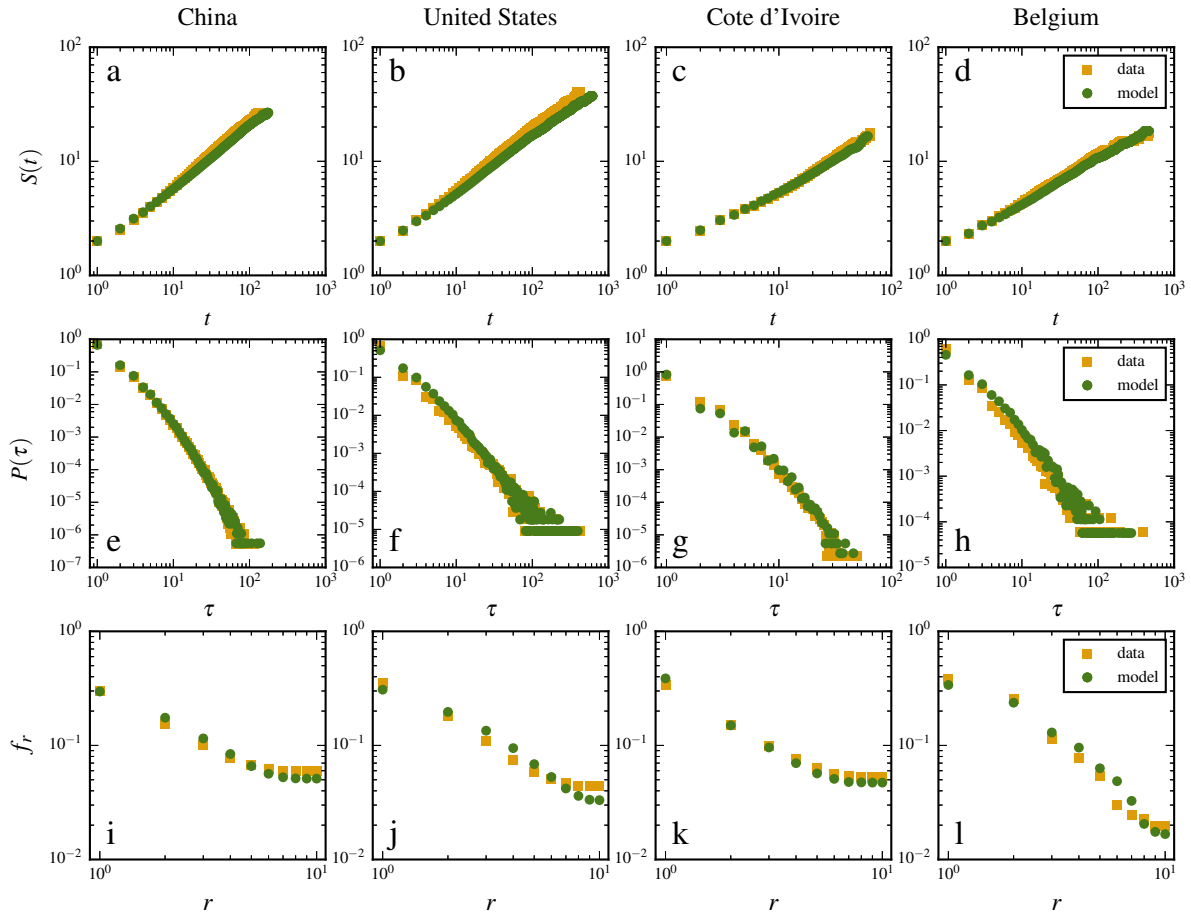
Supplementary Figure 12 | Individual mobility patterns predicted by the competition-free model.

8
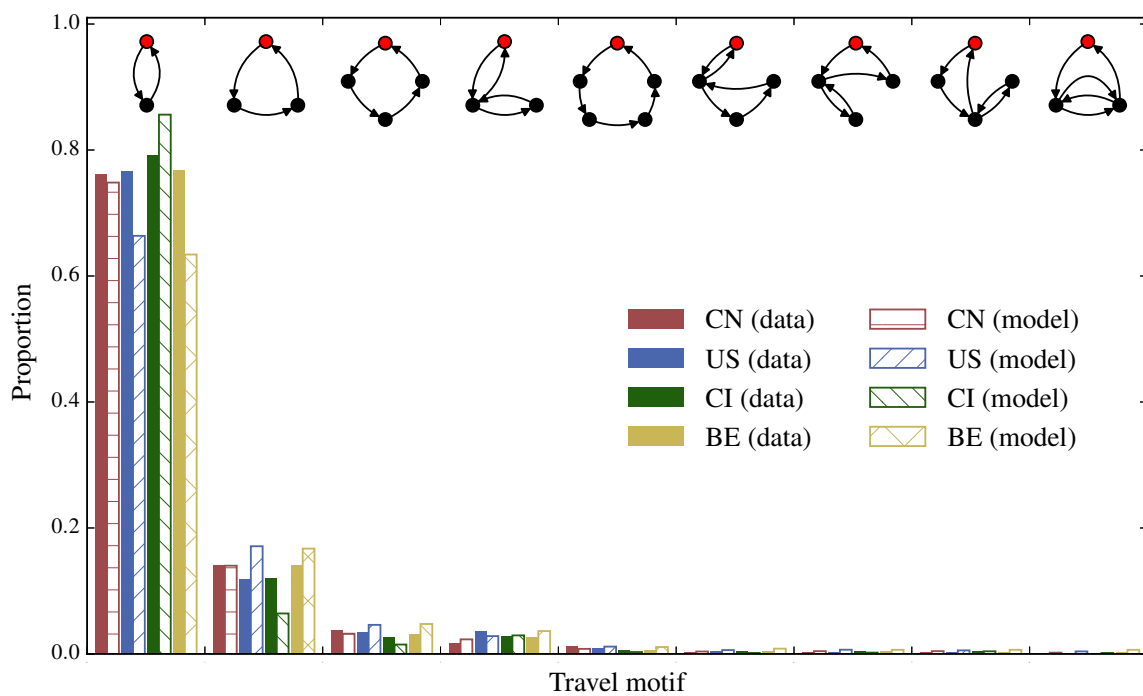
Supplementary Figure 13 | Frequency of occurrence of motifs in the competition-free model.

Supplementary Table

Supplementary Table 1 | Comparison of models prediction accuracy. SSI is the Sørensen similarity index between real observations of collective travel flow and model results. RMSE is the root-mean-square error of 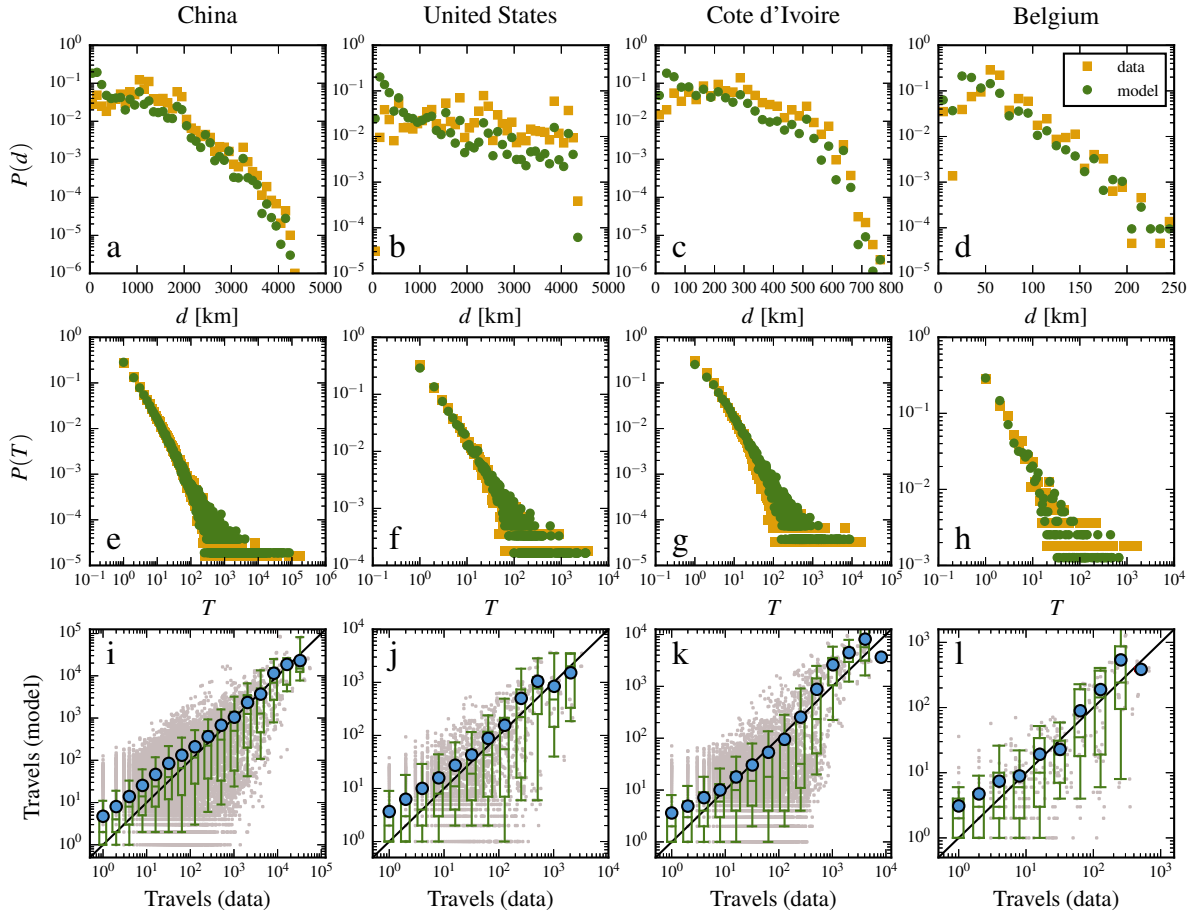predicted $S(t)$. OM, MF, and CF stand for the original model, the memory-free model, and the competition-free model, respectively.

Country	SSI-OM	SSI-MF	SSI-CF	RMSE-OM	RMSE-MF	RMSE-CF
China	0.7126	0.6759	0.5394	1.3556	24.8674	1.6098
US	0.7013	0.7225	0.5156	1.8376	29.3604	4.7381
Cote d'Ivoire	0.7136	0.6914	0.5333	0.6738	16.7012	0.6926
Belgium	0.6936	0.7739	0.5481	0.7141	9.2657	0.8308

Supplementary Notes

Supplementary Note 1: Analysis of individual mobility patterns

1.1 Individual mobility model based on Zipf's law

We consider a simplified version of our human mobility model, which is obtained by imposing the approximations that the population is uniformly distributed among all available locations and the effects of population induced competition are negligible. The simplified model is effectively a random walk with memory in a finite space, in which the transition probability p_{ij} of an individual traveling form location i to j is

$$p_{ij} \propto 1 + \frac{\lambda}{r_j}, \tag{1}$$

where λ is the memory strength parameter and r_j denotes that j is the rth newly visited location for the individual. Insofar as λ is much larger than unity, the transition probability approaches that given by the basic Zipf's law: $p_{ij} \propto \lambda^{-1} + r_j^{-1} \approx r_j^{-1}$.

We first consider the function $S(t)$, the number of locations that an individual has visited in time t. The probability of choosing a new city at the next time step is

$$P_{\text{new}} = \frac{N - S}{N + \sum_{r=1}^{S} \lambda N/r}, \tag{2}$$

where N is the total number of cities and r signifies that the corresponding location is the rth newly visited city. Treating S as continuous, for $S \ll N$, we have

$$P_{\text{new}} \approx \frac{1}{1 + \lambda \int_1^S \frac{dr}{r}} = \frac{1}{1 + \lambda(\ln S + C)}, \tag{3}$$

where C is the Euler's constant and $C \approx 0.577$. We thus get

$$\frac{dS}{dt} = P_{\text{new}} = \frac{1}{1 + \lambda(\ln S + C)}, \tag{4}$$

which gives

$$t = (1 + \lambda C)S + \lambda S(\ln S - 1) - A. \tag{5}$$

11

Since $S(t = 0) = 1$, we obtain $A = 1 + \lambda C - \lambda$. Substituting this relation into Eq. (5) yields

$$t = AS + \lambda S \ln S - A, \tag{6}$$

which determines the function $S(t)$ implicitly. For $\lambda = 0$ (i.e., no memory effect), we have $S = t+1$. For $\lambda = 1/(2-C)$, we have $\lambda = A$, so Eq. (6) can be written as $t = \lambda[S+S\ln(S)-1]$. Using the approximation $\ln(S) \approx S - 1$, we get

$$S \sim \sqrt{(2 - C)t + 1}, \tag{7}$$

indicating that S increases sublinearly with t.

We next derive the frequency distribution f_r of visited locations. Since $S(t)$ distinct locations have been visited by time t, the probability for an individual to choose an already visited location is

$$P_{\text{old}} = 1 - P_{\text{new}} = \frac{S + \sum_{r=1}^{S} \lambda N/r}{N + \sum_{r=1}^{S} \lambda N/r}. \tag{8}$$

Among all the $S(t)$ old locations, the probability of visiting the rth location, i.e., the rate of increase of g_r, the number of times that the location has been visited, is

$$\frac{\mathrm{d}g_r}{\mathrm{d}t} = P_{\text{old}} \frac{1 + \lambda N/r}{S + \sum_{r=1}^{S} \lambda N/r} = \frac{1 + \lambda N/r}{N + \sum_{r=1}^{S} \lambda N/r} \approx \frac{\lambda}{r[1 + \lambda(\ln S + C)]}. \tag{9}$$

From Eq. (6), we have $\mathrm{d}t = [A + \lambda(\ln S + 1)]\mathrm{d}S$, which when being substituted into Eq. (9) yields

$$\frac{\mathrm{d}g_r}{\mathrm{d}S} = \frac{\lambda}{r}. \tag{10}$$

We thus obtain

$$g_r = \lambda S/r + B. \tag{11}$$

For $S = 1$, we have $r = 1$ and $g_r = 1$, so the constant B in Eq. (11) is given by $B = 1 - \lambda$. Since f_r is proportional to g_r, we have

$$f_r \propto \lambda S/r + B. \tag{12}$$

We finally analyze the return time distribution, $P(\tau)$, the probability that a traveler returns to a previously visited location after τ steps. In our model, for the rth first visited location, $P_r(\tau)$ is the joint probability of the following three probabilities: (1) q_r, the probability of visiting the location at the present, (2) $(1 - q_r)^{\tau-1}$, the probability that the rth location will

not be visited in the remaining $\tau - 1$ time steps, and (3) q_r, the probability that the location is visited exactly at time τ. We have

$$P_r(\tau) = q_r^2(1 - q_r)^{\tau-1}.$$ (13)

From Eq. (9), the probability that the rth location is visited at each time step is given by

$$q_r = \frac{H}{r},$$ (14)

where $H = \lambda/[1 + \lambda(\ln S + C)]$. Substituting Eq. (14) into Eq. (13), we get

$$P_r(\tau) = \frac{H^2}{r^2}\left(1 - \frac{H}{r}\right)^{\tau-1}.$$ (15)

For all the already visited locations, the return time distribution is given by

$$P(\tau) = \int_1^S P_r(\tau)\mathrm{d}r = \int_1^S \frac{H^2}{r^2}\left(1 - \frac{H}{r}\right)^{\tau-1}\mathrm{d}r = \frac{H}{\tau}\left[\left(1 - \frac{H}{S}\right)^\tau - (1 - H)^\tau\right],$$ (16)

which is approximately a mixture of two algebraic terms of the identical exponent -1.

1.2 Individual mobility model based on the generalized Zipf's law

The above simplified model assumes that individuals visit locations with the probability given by the basic Zipf's law. However, there are situations where individuals choose their visiting locations by following the generalized Zipf's law [1]: $f \propto 1/r^\zeta$ ($\zeta > 1$). We are thus led to incorporate the generalized Zipf's law into our model to gain a better understanding of the real mobility behavior at the individual level. A straightforward way to generalize our basic model is to modify the transition probability from $p_{ij} \propto 1 + \lambda/r_j$ to $p_{ij} \propto 1 + \lambda/r_j^\zeta$, where ζ is an adjustable parameter. To be able to derive analytical results on the individual mobility patterns, we further simplify the transition probability to be $p_{ij} \propto 1/r^\zeta$. In addition, to obtain theoretical results associated with the generalized Zipf's law, it is necessary to exploit the exploration and preferential return (EPR) model proposed by Song et al. [2]. In the EPR model, an individual has two choices when he/she chooses a location: (1) exploration - the individual moves to a new location with the probability $P_{\text{new}} = \rho S^{-\gamma}$, where S is the total number of visited locations with ρ and γ being parameters; (2) preferential return - the individual returns to a previously visited location with the probability $1 - P_{\text{new}}$, and the

probability to visit an old location is proportional to the number of travels to this location. The EPR model can lead analytically to the generalized Zipf's law from the relation $P_{\text{new}} = \rho S^{-\gamma}$. As an inverse process of the EPR model, the basic assumption of our model is the generalized Zipf's law, and we aim at analytically deriving the relationship between P_{new} and S. Our generalized model can be constructed, as follows.

There are two options for an individual at each step: (1) travel to a new location with the probability P_{new} that decreases with the increase of S and approaches 0 for large values of S; (2) returns to an old location with the probability $1 - P_{\text{new}}$, and the probability of selecting the i-th old location is proportional to the number g_i of travels to location i. We assume that the visiting frequency f_r to the rth visited location follows the generalized Zipf's law

$$f_r \propto r^{-\zeta}, \tag{17}$$

where $\zeta > 1$ is an adjustable parameter. The frequency f_i of visiting the location i can be calculated by the number of visits to all the locations:

$$f_i = \frac{g_i}{t}, \tag{18}$$

where t is the total number of travel steps (time steps), which is equal to the total number of visits to all locations $t = \sum_{i=1}^{S} g_i$. When a location i was first visited, we have

$$g_i(t_i) = 1, \tag{19}$$

where t_i is the time at which location i was discovered. For a large value of S, the probability P_{new} approaches 0, so the rate of increase in the number of visits to the i-th location is approximately given by

$$\frac{\mathrm{d}g_i}{\mathrm{d}t} = \frac{g_i}{t}, \tag{20}$$

for which the solution is

$$g_i = C_i t, \tag{21}$$

where C_i is an arbitrary constant. Combining Eqs. (19) and (21), we have

$$g_i(t_i) = C_i t_i = 1, \tag{22}$$

and

$$C_i = \frac{1}{t_i}. \tag{23}$$

14

Inserting Eq. (23) into Eq. (21), we obtain

$$g_i = \frac{t}{t_i}. \tag{24}$$

Combining Eqs. (17), (18) and (24), we get

$$t_r \propto r^\zeta, \tag{25}$$

where t_r denotes the travel time during which location r was first visited.

For a trajectory with t jumps and S distinct locations, we can rank all visited locations $r = 1, 2, ..., S$ in the order of the discovery time t_r. We then have

$$S(t_r) = r. \tag{26}$$

Combining Eqs. (25) and (26), we obtain

$$S(t_r) \propto t_r^{\frac{1}{\zeta}}, \tag{27}$$

or

$$S(t) \propto t^{\frac{1}{\zeta}}, \tag{28}$$

which gives the scaling relation between time t and the visited location number $S(t)$. Note that $S(t) \propto t^{\frac{1}{\zeta}}$ is the solution of

$$\frac{\mathrm{d}S}{\mathrm{d}t} \propto S^{1-\zeta}. \tag{29}$$

Since the increase rate of S equals the probability of visiting the new location, i.e.,

$$P_{\text{new}} = \frac{\mathrm{d}S}{\mathrm{d}t}, \tag{30}$$

we obtain

$$P_{\text{new}} = \rho S^{-(\zeta-1)}, \tag{31}$$

where ρ is a parameter and can be determined from empirical data [2]. These results demonstrate that, according to the assumption that the visiting frequency follows the generalized Zipf's law, our generalized model can give rise to a power-law relationship between P_{new} and S. If we denote $\zeta = 1 + \gamma$, Eq. (31) becomes

$$P_{\text{new}} = \rho S^{-\gamma}, \tag{32}$$

which is exactly the same as the basic assumption - the relationship between P_{new} and S in the EPR model [2]. This analytical result indicates that the generalized Zipf's law and the

15

power-law relation between P_{new} and S have a mutually causal relationship, so the model for individual mobility based on the former is equivalent to that based on the latter.

We also derive the return time distribution $P(\tau)$. For a sufficiently large value of S, the probability that the rth location is visited at each time step is

$$q_r \propto f_r \propto \frac{1}{r^\zeta} \tag{33}$$

or

$$q_r = \frac{L}{r^\zeta}, \tag{34}$$

where L is a constant and $\sum_1^S L/r^\zeta = 1$. Combining Eqs. (13), (15) and (34), we can formulate the probability that an individual returns to the rth visited location after τ steps as

$$P_r(\tau) = \frac{L^2}{r^{2\zeta}}\left(1 - \frac{L}{r^\zeta}\right)^{\tau-1}. \tag{35}$$

For $\zeta \approx 1$, we have $P_r(\tau) \approx \frac{L^2}{r^2}(1 - \frac{L}{r})^{\tau-1}$, so

$$P(\tau) = \int_1^S P_r(\tau)\mathrm{d}r \approx \frac{L}{\tau}\left[\left(1 - \frac{L}{S}\right)^\tau - (1 - L)^\tau\right]. \tag{36}$$

For $\zeta \gg 1$, $P(\tau)$ cannot be solved analytically but numerical solutions can be readily obtained.

Supplementary Note 2: Analysis of collective mobility patterns

2.1   Uniform population distribution

To derive the mobility patterns at the population level, we simplify the model further by assuming that each individual move only one step. In this case, the individual memory effect can be neglected, so the transition probability p_{ij} of population traveling form location i to j is

$$p_{ij} \propto \frac{m_j}{W_{ji}}. \tag{37}$$

If there are m individuals at each location and all locations are uniformly distributed in a 2-D domain, we have $W_{ji} = \rho \pi d_{ij}^2$, where ρ is the population density. In addition, using $T_{ij} = m_i p_{ij}$, we can rewrite Eq. (37) as

$$T_{ij} \propto \frac{m_i m_j}{d_{ij}^2}, \tag{38}$$

which is the gravity model with a square distance function. In the real world, the spatial distribution of cities in a region often exhibits fractal properties [3]: $W_{ji} \propto d_{ij}^D$, where D is the fractal dimension. We thus have

$$T_{ij} \propto \frac{m_i m_j}{d_{ij}^D}, \tag{39}$$

which is the gravity model with an algebraic distance function.

Since the populations at different locations are assumed to be equal, Eq. (39) can be simplified as

$$T(d) \propto d^{-D}, \tag{40}$$

which is the number T of traveling steps between two locations of distance d apart. We then obtain the travel distance distribution as

$$P(d) \propto d^{-D}, \tag{41}$$

which is an algebraic distribution with exponent $-D$. In the real world the fractal dimension D ranges from 1 to 2 [3]. Indeed a previous work [4] demonstrated that the empirical scaling exponents from some real travel distance distributions are consistent with the value of the fractal dimension.

From Eq. (40), we obtain the distance between two locations as a function of T as

$$d(T) \propto T^{-\frac{1}{D}}. \tag{42}$$

Moreover, the number of location pairs $N(d)$ with distance $\leq d$ in a fractal space is

$$N(d) \propto Q(y \leq d) \propto d^D, \tag{43}$$

where $Q(y \leq d)$ is the fraction of location pairs with distance $\leq d$ in the fractal space. The quantity $Q(y \leq d)$ is larger than or equal to the fraction of location pairs with traveling steps $T(d)$, for the reason that the distance d is a decreasing function of T, as indicated by Eq. (40). We thus have

$$P(x \geq T) = Q[y \leq d(T)]. \tag{44}$$

Combining Eqs. (42), (43) and (44), we obtain the distribution of T as

$$P(T) \propto \frac{\mathrm{d}d(T)^D}{\mathrm{d}T} \propto \frac{\mathrm{d}(T^{-\frac{D}{D}})}{\mathrm{d}T} \propto T^{-2}, \tag{45}$$

which is an algebraic distribution with exponent -2 - a universal value independent of the fractal dimension D of the domain.

2.2   Heterogeneous population distribution

In the real world, population distributions among cities can be highly heterogeneous, which has a significant effect on the distribution of T. To treat the heterogeneity analytically, we consider a simple scenario in which the travelers are only allowed to travel from the central city c to other cities. The central city has the largest population m_c, and the population m_j of another city decreases from m_c algebraically as a function of the distance d_{cj} to the central city:

$$m_j \propto d_{cj}^{-\xi}, \tag{46}$$

with $\xi > 0$. Combining Eqs. (37) and (46), we obtain

$$T_{cj} \propto \frac{m_c m_j}{W_{jc}} \propto \frac{m_c d_{cij}^{-\xi}}{\bar{m} d_{cj}^D} \propto d_{cj}^{-D-\xi}, \tag{47}$$

where $\bar{m}$ is the average population of the cities. Analogously, Eqs. (43), (44) and (47) lead to

$$P(T) \propto \frac{\mathrm{d}(T^{-\frac{D}{D+\xi}})}{\mathrm{d}T} \propto T^{-1-\frac{D}{D+\xi}}, \tag{48}$$

where the scaling exponent ranges from -1 to -2, covering most empirically observed exponent values in the real-world travel flow distributions [see Fig. 5(e-h) in the main text].

Supplementary Note 3: Simulation setup of mobility behaviors in a fractal domain

To validate our analytical results, we simulate our simplified model in four typical fractal domains: 2D Cantor dust, Vicsek fractal, Sierpinski triangle, and hexaflake. The results are

18

shown in Fig. 6(d) in the main text. Here we describe the four fractal sets and the simulation setup.

3.1 2D Cantor dust

The Cantor dust [5] is a 2D version of the Cantor set. It is obtained by starting with a basic unit square, scaling its size by 1/3, then placing four scaled copies at the four corners, respectively. Recursively applying the procedure *ad infinitum*, we generate a 2D Cantor set, as shown in Supplementary Fig. 1. Let N_n be the number of black boxes and L_n be the length of a side of a box after the nth iteration. The fractal dimension of the Cantor set is

$$D = -\lim_{n\to\infty} \frac{\ln N_n}{\ln L_n} = -\frac{\ln 4^n}{\ln 3^{-n}} = \frac{\ln 4}{\ln 3} \approx 1.262. \tag{49}$$

3.2 Vicsek fractal

The Vicsek fractal [6], also known as the box fractal, can be constructed using a procedure similar to that for the 2D Cantor dust. In particular, the basic unit square is decomposed into nine smaller squares in a 3-by-3 grid. The four squares at the corners and the middle square are kept, while the other squares are removed. The process is repeated recursively for each of the five remaining subsquares. The Vicsek fractal is the set obtained at the limit of this procedure, as shown in Supplementary Fig. 2. Its fractal dimension is

$$D = -\lim_{n\to\infty} \frac{\ln N_n}{\ln L_n} = -\frac{\ln 5^n}{\ln 3^{-n}} = \frac{\ln 5}{\ln 3} \approx 1.465. \tag{50}$$

3.3 Sierpinski triangle

The Sierpinski triangle [7], or Sierpinski gasket, is obtained by starting from an equilateral triangle of unit side length (in fact any triangle can be used), subdividing it into four smaller congruent equilateral triangles, and removing the central one. This procedure is repeated *ad infinitum*, generating the Sierpinski triangle fractal, as shown in Supplementary Fig. 3. Its fractal dimension is

$$D = -\lim_{n\to\infty} \frac{\ln N_n}{\ln L_n} = -\frac{\ln 3^n}{\ln 2^{-n}} = \frac{\ln 3}{\ln 2} \approx 1.585. \tag{51}$$

19

3.4 Hexaflake

Hexaflake [8] is a fractal constructed starting from a hexagon, which is replaced by a flake of smaller hexagons such that a scaled hexagon is placed at each vertex and at the center. This procedure is repeated recursively to generate the hexaflake, as shown in Supplementary Fig. 4, with its fractal dimension given by

$$D = -\lim_{n \to \infty} \frac{\ln N_n}{\ln L_n} = -\frac{\ln 7^n}{\ln 3^{-n}} = \frac{\ln 7}{\ln 3} \approx 1.771. \tag{52}$$

3.5 Simulation setup

We iterate the construction procedure for each of the four fractal geometries and set the central position of the shapes in each fractal set (black blocks in Supplementary Fig. 1-4) to be the location coordinates. To avoid an idealized fractal domain and better mimic empirically observed fractal features in the real world, we add a small random offset to the coordinates of each location. The iterations are 5, 4, 6, 3 for the 2D Cantor dust, Vicsek fractal, Sierpinski triangle and hexaflake, respectively, and the numbers of existent shapes (locations) are 1024, 625, 729 and 343, respectively.

We set population m at each location and use Eq. (37) to calculate the number of traveling steps T_{ij} between any location pairs. We use $\min(T)$ to rescale all T, i.e., we let the minimum number of traveling steps be one, and finally obtain the travel flow distribution $P(T)$. Similarly, we can obtain the rescaled travel distance distribution $P(d)$, as shown in Fig. 6(d) in the main text.

Supplementary Note 4: Model validation with empirical mobility data at the city scale

In the main text we have demonstrated that our model is universally applicable to countries with diverse spatial scales. Here we show that the model can also characterize individual and population mobility patterns at the city scale accurately.

We use the foursquare check-ins data set [9] in New York city as the proxy data of human mobility for model validation. The data set contains 42035 individuals, in which 23520 have travels among different locations (here the locations are defined as the 2010 census blocks [10], see Supplementary Fig. 5), and the total number of traveling steps is 113279.

For this data set, we first estimate the memory strength parameter and obtain $\lambda \approx 8.0$. We then simulate our model using this parameter value and compare the results with those from the real mobility data, as shown in Supplementary Figs. 6 and 7. We see that both the individual and collective mobility patterns in New York city can be predicted, suggesting that our model is capable of characterizing human mobility patterns at small (intra-city) spatial scales.

Supplementary Note 5: Comparison with alternative models

Our mobility model contains two essential ingredients: individual memory and population induced competition effects. They jointly determine the attractiveness of a location, as well as the transition probability for an individual to move from one location to another. Here we consider two alternative models, a memory-free model and a competition-free model, to show that the two ingredients are indispensable for modeling human mobility patterns.

5.1 Memory-free model

To construct a memory-free model, we modify Eq. (3) in the main text by writing the transition probability p_{ij} of traveling from location i to j as

$$p_{ij} \propto \frac{m_j}{W_{ji}}. \tag{53}$$

The model prediction results are shown in Supplementary Figs. 8, 9 and 10. We see that, while the memory-free model can reproduce the collective mobility patterns to certain extent, it fails to capture the individual movement patterns.

To offer a better comparison of the results obtained from different models, we use two indices to measure the model prediction accuracy. The first is the Sørensen similarity index (SSI) [11], which is a statistical tool to identify the similarity between two samples. Here we use SSI to quantify the degree of similarity between real observations of collective travel flow between locations and model prediction results. The SSI is defined as

$$\text{SSI} \equiv \frac{1}{N^2} \sum_i^N \sum_j^N \frac{2\min(T'_{ij}, T_{ij})}{T'_{ij} + T_{ij}}, \tag{54}$$

where T'_{ij} is the model predicted traveling steps from location i to j and T_{ij} is the observed number. Apparently, if each T'_{ij} is equal to T_{ij} the index should be unity, whereas if all T'_{ij}s are far from the real values, the index should be close to 0.

The second index is the root-mean-square error (RMSE) [12], which is a frequently used quantity to measure the differences between the model predicted and actually observed values. We use RMSE to measure the prediction errors of the models in the total number of locations $S(t)$ visited within time t, which is an important characteristic of individual movement patterns. The RMSE of $S(t)$ is defined as

$$\text{RMSE} = \sqrt{\frac{\sum_{t=1}^{L_\text{max}} [S_\text{real}(t) - S(t)]^2}{L_\text{max}}}, \tag{55}$$

where $S_\text{real}(t)$ is obtained from the actual data set, $S(t)$ is calculated from the model, and L_max is the maximum time step.

We calculate the SSI between observed and predicted travel flow and RMSE of the predicted $S(t)$ for the original model and memory-free model, and list the results in Supplementary Table 1. These results suggest that the memory effect mainly affects the individual movement patterns but has a little effect on the collective mobility patterns.

5.2 Competition-free model

The competition-free model is a memory-based mobility model without any population induced competition effect. In this model, the transition probability p_{ij} of traveling from location i to j is

$$p_{ij} \propto m_j(1 + \frac{\lambda}{r_j}), \tag{56}$$

meaning that individuals only consider the inherent attractiveness of locations and their own preferences induced by memory when selecting travel destinations.

Supplementary Figs. 11, 12, 13 and Supplementary Table 1 show the prediction results from the competition-free model. We see that, in contrast to the memory-free model, the competition-free model can reasonably reproduce the individual mobility patterns, but the accuracy of predicting collective movement patterns is poor. *These results further demonstrate that both the memory effect and population induced competition effect are essential for modeling and predicting human mobility patterns simultaneously at the individual and population levels.*

Supplementary References

[1] R. E. Wyllys, Empirical and theoretical bases of Zipf's law. *Library Trends* **30**, 53-64 (1981).

[2] C. Song, T. Koren, P. Wang, A. L. Barabási, Modelling the scaling properties of human mobility. *Nat. Phys.* **6**, 818-823 (2010).

[3] P. Frankhauser, The fractal approach. A new tool for the spatial analysis of urban agglomerations. *Population* **10**, 205-240 (1998).

[4] R. Gallotti, A. Bazzani, S. Rambaldi, M. Barthelemy, A stochastic model of randomly accelerated walkers for human mobility. *Nat. Commun.* **7**, 12600 (2016).

[5] G. Helmberg, *Getting Acquainted with Fractals* (Walter de Gruyter, 2007).

[6] T. Vicsek, M. Shlesinger, M. Matsushita, *Fractals in Natural Sciences* (World Scientific, 1994).

[7] A. A. Lale, B. V. Khiste, G. Burshe, S. Khobragade, Study of Sierpinski triangle gasket. In 2011 3rd International Conference on Electronics Computer Technology (ICECT), Vol. 3, pp. 386-390 (2011).

[8] S. M. Choudhury, M. A. Matin, Effect of FSS ground plane on second iteration of hexaflake fractal patch antenna, In 2012 7th International Conference on Electrical Computer Engineering (ICECE), pp. 694-697 (2012).

[9] J. Bao, Y. Zheng, F. M. Mokbel, Location-based and Preference-Aware Recommendation Using Sparse Geo-Social Networking Data. In ACM SIGSPATIAL (GIS 2012), Redondo Beach, CA, US, 2012

[10] Data can be accessed via http://gis.nyc.gov/census/

[11] T. Sørensen, A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on danish commons. *Biol. Skr.* **5** 1-34 (1948).

[12] R. J. Hyndman, A. B. Koehler, Another look at measures of forecast accuracy. *Int. J. Forecast.* **22**, 679-688 (2006).