Understanding metropolitan patterns of daily encounters

Lijun Sun[a,b], Kay W. Axhausen[a,c,1], Der-Horng Lee[b], and Xianfeng Huang[a,d]

[a]Future Cities Laboratory, Singapore–ETH Centre for Global Environmental Sustainability, Singapore 138602; [b]Department of Civil and Environmental Engineering, National University of Singapore, Singapore 117576; [c]Institute for Transport Planning and Systems, Swiss Federal Institute of Technology, CH-8093 Zürich, Switzerland; and [d]State Key Lab of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China

Edited by Susan Hanson, Clark University, Worcester, MA, and approved July 3, 2013 (received for review April 5, 2013)

Understanding of the mechanisms driving our daily face-to-face encounters is still limited; the field lacks large-scale datasets describing both individual behaviors and their collective interactions. However, here, with the help of travel smart card data, we uncover such encounter mechanisms and structures by constructing a time-resolved in-vehicle social encounter network on public buses in a city (about 5 million residents). Using a population scale dataset, we find physical encounters display reproducible temporal patterns, indicating that repeated encounters are regular and identical. On an individual scale, we find that collective regularities dominate distinct encounters' bounded nature. An individual's encounter capability is rooted in his/her daily behavioral regularity, explaining the emergence of "familiar strangers" in daily life. Strikingly, we find individuals with repeated encounters are not grouped into small communities, but become strongly connected over time, resulting in a large, but imperceptible, small-world contact network or "structure of co-presence" across the whole metropolitan area. Revealing the encounter pattern and identifying this large-scale contact network are crucial to understanding the dynamics in patterns of social acquaintances, collective human behaviors, and—particularly—disclosing the impact of human behavior on various diffusion/spreading processes.

human mobility | behavioral rhythms | social networks | social sciences

Highlighting their importance in various temporal spreading processes (1–3), recent studies of human contact networks demonstrate an increasing interest in physical encounters (4–9). Contrary to nonphysical social contacts initiated by mobile phone calls, e-mails, and online social networks (10–14), human subjects' physical encounters take place with heterogeneous prior knowledge of each other, from acquaintances to unknowns, linking two individuals by their copresence in both spatial and temporal dimensions (15). On the other hand, with increasing human interactions, communities may also emerge from social contagion enabled by physical proximity: from not noticing each other, to unintentionally interacting, to intentional communicating, to mutual trust (9). We are tracing—for a large population—in this case for all of Singapore's bus users, how these encounters are structured by both individual behavior and institutional structures. Bus use is a small slice of urban life, but one where "familiar strangers" will emerge (16, 17)—strangers who have been encountered frequently in daily life, but might never have been addressed. This context is one of many, which, in their totality, give residents the social background against which they construct their more intense relationships. We believe the joint encounter pattern is influenced by individual regularity (18, 19). A previous study based on the dispersal of bank notes suggested that human trajectories follow continuous-time random walks (20); however, considering the inherent regularity in individual behaviors, recent analyses of large-scale trajectories from mobile phone data and travel diaries indicate, on the contrary, that individual mobility patterns display significant regularity and remarkable predictability (18, 19, 21).

With the help of sensors and online networks, data describing close proximity in real-world situations sheds light on encounter patterns and spreading dynamics in contact networks other than diary-based surveys (4). However, these data collection systems are generally embedded in limited samples in spatially small-scale settings such as schools (6), conferences and exhibitions (5, 7), and even in prostitution (8). On a large scale, we still lack empirical data describing examples of both individual regularity and joint encounter patterns (other than simulating mobility and behavior patterns individually, relying on computational and agent-based models) (22–24). Thus, given data limitations, studies on individual mobility regularity and collective interactions are traditionally conducted separately: the mechanisms driving our daily encounters remain unclear.

Therefore, with the increasing quantity and range of human mobility, a central task is to explore social interaction patterns along with mobility regularity. However, previous data collection techniques fail to offer high-resolution information on collective interactions on a large scale (across the population). In this context, individual-based passive data collections embedded in our daily life, such as credit cards and smart cards transactions, can be advantageous. At present, transit use might be the best proxy to capture the patterns of both individual mobility and collective interactions in an urban environment (25). Here, we use public transit transaction records to uncover encounter patterns (see *SI Appendix*, section I for a detailed description). This dataset consists of more than 20 million bus trips from 2,895,750 anonymous users over 1 wk (Fig. 1*A*) (about 55% of the resident population) in Singapore. The high spatial-temporal resolution of this dataset allows us to extract time-resolved in-vehicle encounters, defined as two individuals occupying the same vehicle simultaneously (Fig. 1*B*). Using these, a city-scale empirical temporal contact network is created.

Results

Use of transit service in general, and buses in particular, is differentiated by ethnicity, sex, age, and income, meaning that daily transit use might exhibit social segregation as well. To address dependency and segregation of bus use on social demographic attributes, we incorporated two additional datasets aside from smart card transactions: population census and national household interview travel survey (HITS). For our case, although transit use in Singapore shows dependency on age and income, public transit is still the most important transport means for daily commuting trips across all Singaporeans (*SI Appendix*, section

Author contributions: L.S., K.W.A., and D.-H.L. designed research; L.S. and K.W.A. performed research; L.S., K.W.A., and X.H. analyzed data; and L.S. and K.W.A. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. E-mail: axhausen@ivt.baug.ethz.ch.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1306440110/-/DCSupplemental.

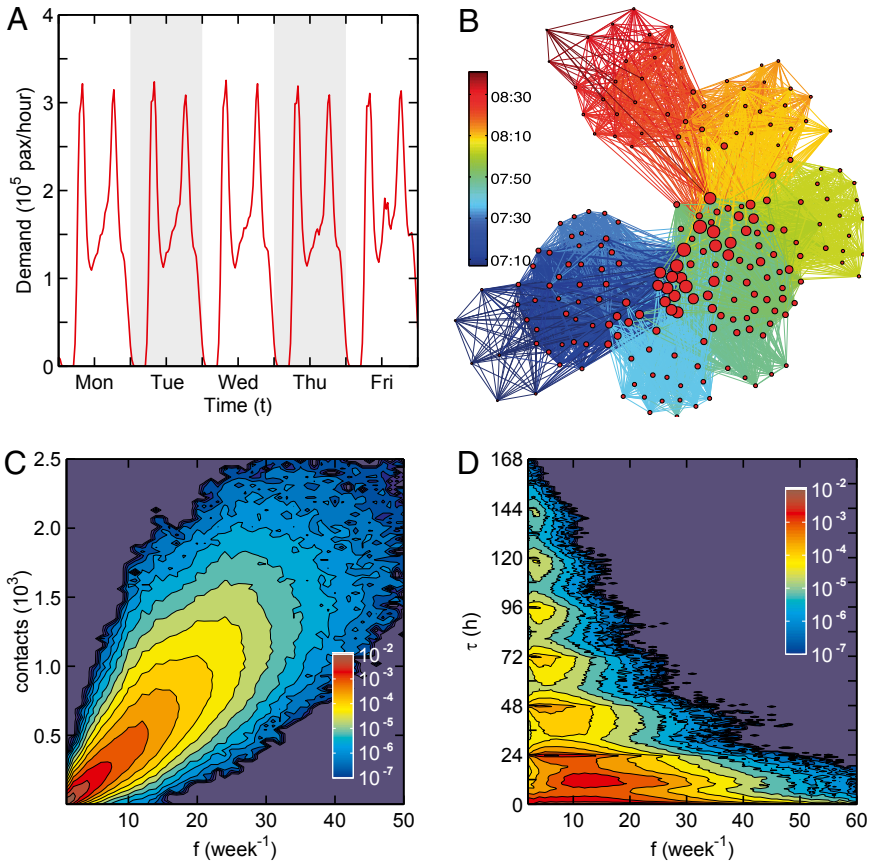

Fig. 1. Characteristics of transit behaviors. (*A*) Departure rate of city bus trips. Demand shows similar shapes from Monday to Friday. Two prominent peaks can be identified every day, indicating collective commuting behavior. (*B*) Time-resolved encounter network on one vehicle service (214 passengers) created using transaction records. Size of each node indicates number of encountered people; color and thickness of each edge indicate start time and duration, respectively, of each encounter activity. This figure explains use of smart card data to create an empirical encounter network in city-scale. (*C*) Probability density function $P(f,n)$ of trip frequency f and number of encountered people n of all individuals. Symmetric pear-shaped pattern suggests linear relationship of $n=\sum_{i=1}^{f} n_i \approx 50f$. (*D*) Probability density function $P(f,\tau)$ of trip frequency f and interevent interval τ between consecutive bus trips. Although diversity of individual transit use obscures the distribution, a remarkable area with $(5\le f\le 25, 0\le \tau\le 25)$ can be identified through similarity in regular commuters' transit behavior.

II). By studying transit use, we found that both trip duration and trip frequency can be accurately characterized by exponentially decaying tails, showing that people's transit activities are limited in number and duration during 1 wk (*SI Appendix*, sections III.2 and IV.1). To understand individuals' transit use patterns, we first measured the number of encountered people n against trip frequencies f for each individual. We find that the joint distribution $P(f,n)$ has a symmetric pattern against $n/f\approx 50$, indicating the substantial number of people encountered in each trip (Fig. 1*C*). We then measured the interevent time τ between consecutive bus trips for the population, finding that $P(\tau)$ shows clear temporal patterns with prominent peaks (*SI Appendix*, section IV.3), which is contrary to the non-Poissonian nature of $P(\tau)\sim\tau^{-\beta}\exp(-\tau/\tau_c)$ in communication activities observed from digital networks (10–14), suggesting the periodicity of transit use on the population scale. To explore the pattern of τ at the individual level, we grouped people according to their trip frequencies. In Fig. 1*D*, we added f to each τ as an attribute and measured the joint probability distribution $P(f,\tau)$. Observing the heterogeneity rooted in individual behaviors helps us distinguish regular travelers from other passengers. Therefore, given the transit use variation (and the analysis in *SI Appendix*, sections III and IV), we suggest that transit use is a good example for capturing both individual mobility patterns and collective encounter patterns, certainly for our case study.

As mentioned, an important phenomenon triggered by collective regularities is the familiar stranger, which is also crucial in explaining how likely it is that the same persons will be encountered again (Fig. 2*A*). To explore the pattern of repeated encounters on the population scale, we created an aggregated encounter network over weekdays and measured the interevent time τ_e between consecutive encounters of paired individuals, capturing 27,892,055 intervals from 18,724,388 pairs. We found that the distribution $P(\tau_e)$ is characterized by prominent peaks at 24 h, 48 h, 72 h, and 96 h, displaying a strong tendency of periodic encounters covering about 75% of all cases, suggesting that the joint regularity also displays significant temporal periodicity (Fig. 2*B*). We also observed a decreasing pattern of $P(\tau_e)$, as the observed numbers decrease when interval τ_e is longer. To avoid bias, we extracted and grouped the timings of current and the next encounters (T_c,T_n) over all pairs, so that for one pair with three encounters, two records are created $[(T_{1st},T_{2nd})$ and $(T_{2nd},T_{3rd})]$. We then measured the joint distribution $P(T_c,T_n)$ (Fig. 2*C*). Strikingly, we found that the joint distribution presents reproducible patterns on the same diagonal with $D_n-D_c=\{1d,2d,3d\}$, respectively, where D_c and D_n represent the days of T_c and T_n, respectively, suggesting the homogeneity of daily encounters. To measure reproducibility, we summed the probability for each day-of-the-week pair, excluding the diagonal cases with $D_c=D_n$, obtaining the density matrix:

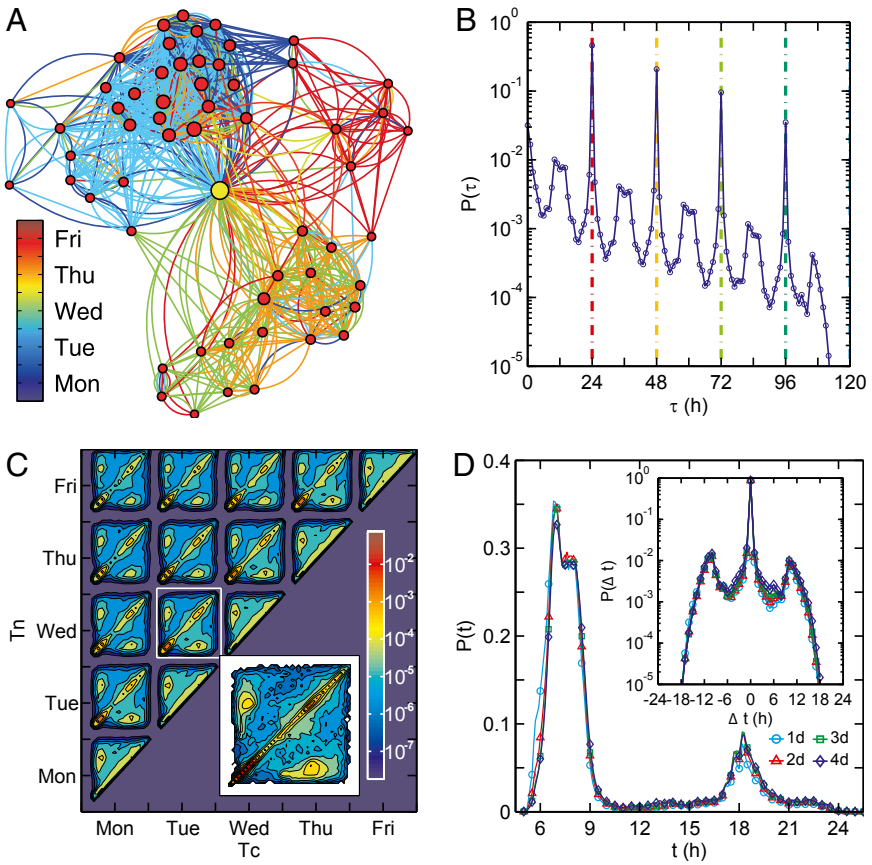

Fig. 2. Temporal patterns of repeated encounters. (*A*) A typical temporal contact network of one individual with his/her familiar strangers (encountered more than once). Edge colors indicate start timings during the week. As observed, neighbors are also highly connected over time, suggesting network's strong social connection and small-world property. (*B*) Probability density function $P(\tau_e)$ of interevent time between two consecutive encounters of paired individuals. The prominent peaks suggest that paired individuals tend to meet each other daily. A general decreasing trend is observed as well. On average, pairs of individuals with repeated encounters can meet each other about $27,892,055/18,724,388+1\approx 2.5$ times during the week. (*C*) Joint probability density $P(T_c, T_n)$ of timings of two consecutive encounters. (*Inset*) Zoom-in of one density square. (*D*) Distribution $P(t)$ of encounter time for four groups with $D_n - D_c = \{1d, 2d, 3d, 4d\}$, respectively. Inset of figures shows the distribution $P(\Delta t)$ of intervals between rescaled timing of paired encounters $\Delta t = t_n - t_c$. The prominent peak at $\Delta t = 0$ suggests most repeated encounters happen at the same time.

$$P(D_c, D_n) = \begin{bmatrix} 0 & 0.141 & 0.086 & 0.061 & 0.046 \\ 0 & 0 & 0.148 & 0.091 & 0.062 \\ 0 & 0 & 0 & 0.145 & 0.085 \\ 0 & 0 & 0 & 0 & 0.135 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

Considering the values in $P(D_c, D_n)$, we found that repeated encounters over the population can be modeled well as a Bernoulli process with probability of success $P_{encounter} \approx 0.33$, which is another factor behind the decrease of $P(\tau_e)$.

To reveal the homogeneity of day-to-day encounters, we further studied $P(T_c, T_n)$. As the *Inset* of Fig. 2*C* shows, we found a strong diagonal on $23h \le T_n - T_c \le 25h$, which covers 85% of the cases, suggesting that most recurring encounters happen at about the same time of day. Despite this, we also observed two areas representing cross-period encounters, such as the first in the afternoon and the following in the morning on the next day. To compare the distributions of different day-of-the-week pairs (D_c, D_n), we grouped the pairs according to $D_n - D_c = \{1d, 2d, 3d, 4d\}$, respectively, and rescaled both T_c and T_n to time of day t_c and t_n. Then, we measured the probability density $P(t)$ of encounter time by merging t_c and t_n (Fig. 2*D*) and the distribution of interevent intervals $\Delta t = t_n - t_c$ (Fig. 2*D*, *Inset*). Taken together, we found both $P(t)$ and $P(\Delta t)$ for different groups share indistinguishable shapes, indicating that the daily encounters can be characterized by a general temporal pattern. Furthermore, although afternoon peaks are longer and higher than morning peaks in daily transit use (Fig. 1*A* and *SI Appendix*, Fig. S6*B*), we confirmed that repeated encounters tend to happen more often in the morning, suggesting that collective regularity is more pronounced in the morning than in the afternoon (21). In this contact network, it is also implied that spreading via repeated interactions is more likely in the morning than afternoon. In addition, the prominent peaks at $\Delta t = 0$ indicate that the most probable time for a recurring encounter is the same as the previous one.

Still, until now, the field lacked analyses to uncover mechanisms that drive encounters on an individual level. Therefore, we began to measure the encounter frequency of paired individuals, finding that the distribution $P(f_e)$ of encounter frequencies can be well characterized by a heavy-tailed distribution, even though the network is very dense, explaining the emergence of familiar strangers (*SI Appendix*, section VI.3). Although f_e is a good approximation of connection strength, it fails to capture the actual overlapping of collective behaviors when considering other external factors such as transfers. To avoid misinterpreting the data, we used total encounter duration $d(i,j) = \sum_{k=1}^{f_e(i,j)} t_{d,k}(i,j)$ to better quantify the connection strength of individual pairs (i,j), where $t_{d,k}(i,j)$ is the duration of their k^{th} encounter. In Fig. 3*A*, we show the distributions $P(d)$ (for all individual pairs) and $P(t_d)$

(over all trips) in both log–log scale and semilog scale. We found that $P(t_d)$ can be well captured by an exponentially decaying tail, whereas $P(d)$ displays a power-law tail $P(d) \sim d^{-\beta}$ with exponent $\beta \approx 4.8 \pm 0.1$. The significant degree of heterogeneity indicates that collective d has overtaken the exponential bounded t_d. To summarize the observation of pairs of travelers, collective regularities do occur, suggesting that encounter patterns are influenced by paired individuals' behavior patterns.

Next, on an individual level (see *SI Appendix*, section VII for modeling details), we propose a personal weight w_i of individual i proportional to their paired encounter frequency:

$$w_i \equiv \sum_{j \in N(i)} (f_e(i,j) - 1), \qquad [1]$$

where $N(i)$ is the set of encountered people and $f_e(i,j)$ is the frequency of encounters between individual i and j observed. Thus, w_i captures the likelihood of encountering familiar strangers. In Fig. 3*B*, we chose individuals who had recurring encounters ($w_i \geq 0$) and plotted the probability density functions $P(w)$ of personal weight and $P(k)$ of the number of familiar strangers, respectively. We find that both distributions can be approximated well by power-laws with high cutoffs, with the same exponent $\beta \approx 1.08 \pm 0.02$. It is important to note the great variation in person weights—that is, encounter likelihoods—suggesting that encounter patterns might be influenced by individual behavior patterns.

To explore how individual behavioral regularity impacts collective encounter patterns, we tried to quantify both individual encounter capability and transit use. First, to better measure individual encounter likelihood beyond travel time influence, we rescaled w_i to $r_i = w_i / T_i (hour^{-1})$ for each individual, where T_i is the total travel time in hours. Inspired by the k-means clustering method—for individual i—we used the absolute difference m_i of morning and evening trips to quantify behavioral regularity:

$$m_i = \sum_{k=1}^{2} \sum_{t_j \in S_k} |t_j - \mu_k| / n, \qquad [2]$$

where $t_j = (t_{start} + t_{end})/2$ is the mean of start and end times of the j^{th} $(j = 1, 2, \cdots, n)$ trip and μ_k is the mean of $\{t_j | t_j \in S_k\}$. Therefore, m_i captures the time variation of daily transit use (the lower m_i is, the more repetitive the individual will be; *SI Appendix*, section VII.2). To reveal the relation between behavioral regularity and encounter likelihood, we grouped the individuals according to

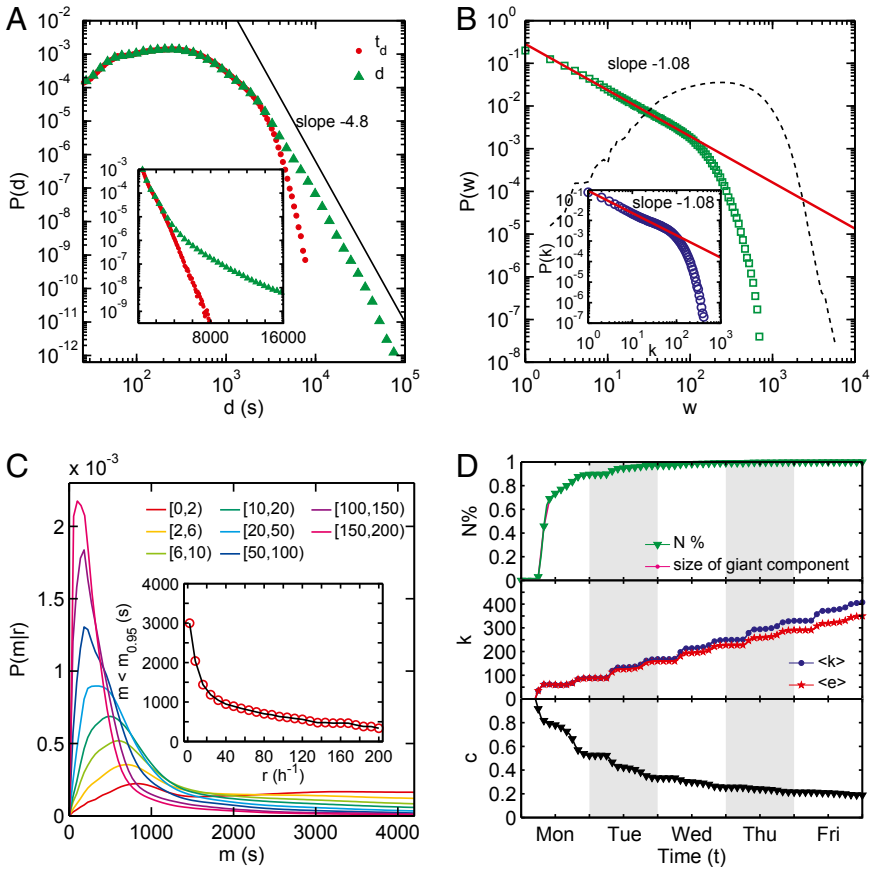

Fig. 3. Individual's impact on collective encounters. (*A*) Probability density functions $P(d)$ of sum of durations of repeated encounters and $P(t_d)$ of duration of individual encounters on weekdays; the same distributions are shown in semilog scale as *Inset*. Solid line shows power-law with exponent $\beta = 4.8$, suggesting that although distinct encounters exhibit exponentially bounded nature, paired total encounter duration presents scale-free property. Maximum values of d and t_d are $9.22 \times 10^4 s$ (25.6 h) and $7.93 \times 10^3 s$ (2.2 h), respectively. (*B*) Distributions $P(w)$ of personal weight and $P(k)$ of number of familiar strangers (*Inset*) of 1,626,040 individuals with $w_i > 0$. Solid lines show power-laws and dashed curve shows distribution $P(u)$ of number of total encounters $u_i \equiv \sum_{j \in N(i)} f_e(i,j)$ as guides. (*C*) Probability density function $P(m)$ of absolute trip difference m for grouped individuals with different r (colored lines). (*Inset*) Mean value of $m < m_{0.95}$ for people with certain r, suggesting a direct correlation: the higher one's encounter capability, the more stable one's behavior. (*D*) Evolution of time-aggregated network of 783,247 individuals with personal weight $w_i \geq 10$. (*Top*) The fraction of presented nodes and size of largest component. We found these two curves indistinguishable, suggesting high network connectivity. (*Middle*) Evolutions of average number of encounters $\langle k \rangle$ and average number of encountered people $\langle e \rangle$. (*Bottom*) Evolution of average clustering coefficient c during the week, suggesting small-world property of this aggregated network.

r_i and measured the distribution $P(m|r)$ for each group. As Fig. 3*C* shows, individuals with higher r_i tend to have less skewed $P(m)$, whereas those with low r_i display a more skewed distribution. In addition, we took the average of $m \le m_{0.95}$ over people with certain r, where $m_{0.95}$ is the 95th percentile of their m, finding that the average absolute difference reaches about 50 min for $r < 4h^{-1}$, whereas for those with $r \ge 80h^{-1}$, the value decreases to less than 15 min, significantly shorter considering that the expected headway (i.e., service interval) of public buses is around 10 min (Fig. 3*C*, *Inset*). In summary, we found that a larger encounter likelihood of an individual is strongly rooted in his/her behavioral regularity.

With these common daily physical interactions, the resulting regularity-rooted encounter network plays an important role in various urban environment dynamics like epidemic spreading. Given that most contact-network–based spreading models still focus on network topology (26, 27) or small-scale contact processes (5–8), identifying this real-world physical contact network is potentially important in studying encounter patterns and diffusion/spreading dynamics in large populations. To model the dynamical evolution of this contact network, we extracted time-aggregated networks of people with $w_i \ge 10$. At the top of Fig. 3*D*, we plotted the fraction of those individuals over the population, finding a rapid increase from 0% to about 90% on Monday, followed by slower growth afterward. We next checked the variation of average number of encounters $\langle k \rangle$ and average number of encountered people $\langle e \rangle$, finding linear increases of both $\langle k \rangle$ and $\langle e \rangle$ without saturation, indicating that the people one may encounter differ from day to day, result in weak and random connections compared with social relations (Fig. 3*D*, *Middle*). However, we note that $\langle k \rangle$ increases faster than $\langle e \rangle$, suggesting that "stronger" connections with familiar strangers are formed gradually through random encounters. At the bottom of the figure, we plotted the evolution of average clustering coefficient c, finding that the encounter network displays strong small-world property network with characteristic path length $\langle l \rangle \ge l_{rand}$ and $c \gg c_{rand}$ (28) ($\langle l \rangle = 2.95$, $l_{rand} = \ln(n)/\ln(k) = 2.63$ and diameter $l_{max} = 6$; $c = 0.19$ and $c_{rand} = \langle e \rangle / n = 4.5 \times 10^{-4}$). Viewed as a whole, the empirical encounter network we illustrate here is a well-connected small-world graph, in which individuals are no longer confined to local encounters in one vehicle, but interact strongly with increasing number of people across the whole city from day to day.

Discussion

It has been assumed that human behavior and social interaction/contagion were connected for a long time; however, it is difficult to identify the link between them in observational studies (29, 30). Taking advantage of the availability of metropolitan data collection offered by smart cards in Singapore (this exercise could be very useful for other cities around the world), we tie together thinking on individual mobility with collective interaction patterns. Although the specific results are certainly embedded within the social profile and data of our study, questions remain: how can smart card data give such insight on social interactions, and how much do these patterns vary from context to context? As a result of various preferences and constraints on individual behavior, spatial-temporal patterns and collective regularity can be found in daily life, such as morning/evening peaks in transportation, degree of crowding in shopping malls and supermarkets at weekends and in restaurants during dining hours, and so forth. Transit use is only one of these social activities with a limited time allocation and specific locations. This study scrutinized only one of a whole spectrum of metropolitan patterns, although it is a critical one. We took transit users as our subjects and the definition of encounter in our study is limited to physical proximity or copresence, which is characterized by individuals occupying the same vehicle simultaneously. Considering vehicle configuration and loading profile (*SI Appendix*, section I), physical proximity does not necessarily indicate a more intense social contact, such as talking to each other, but does imply diverse interactions, from not noticing each other, to fleeting eye contact and close observation. As a core of social psychology, social contagion deals with thoughts and behaviors of others by innovation adoption, rumor spreading, and decision making. Although the similarity between social contagion and epidemic spreading was recognized a long time ago (31), in the context of physical proximity, social contagion depends more on familiarity than epidemic spreading. Thus, future questions are raised: how to measure the familiarity in the passive familiar strangers networks, and how to define the threshold of familiarity on social diffusion processes. Nevertheless, we know that social communities emerge from the increasing familiarity of individuals into collective forms: from unintentional and passive interactions to intentional and active communication, from mere physical proximity to mutual trust (9). We have shown the existence of city-wide structures of copresence in Singapore; the extent to which the bus users are aware of these structures has yet to be determined.

Unlike other social networks, where people interact within a circle of friends and acquaintances, we show an often-ignored type of social link: weak, passive, and indirectly enabled by daily encounters. Moreover, the structuring of physical encounters is demonstrated on a metropolitan scale. As a result of deep-rooted individual behavior patterns, our results also present the collective regularity of people with their recurring encounters as evidence, explaining the familiar strangers phenomenon in daily life. With the rapid development of information and communication technologies, richer data will be generated throughout our daily life (32). Although the role of such data is limited by an inherent tradeoff, it does not portray everything in detail. However, the emergence of such data provides us with considerable opportunities to enhance our understanding of the social world and its processes. Our work should serve as a base to better understand collective human behaviors, dynamical evolution of social networks (33, 34), and especially the impact of collective regularity on various diffusion/spreading processes (4, 35, 36).

ACKNOWLEDGMENTS. We thank Peng Gong and Song Liang for discussion and comments on the manuscript, Karen Ettlin for copyediting the paper, Margaret Grieco for giving us important comments on the text, and Singapore's Land Transport Authority for providing the smart card data. This study was supported by the National Research Foundation of Singapore, the funding authority of the Future Cities Laboratory.

1. Holme P, Saramäki J (2012) Temporal networks. *Phys Rep* 519(3):97–125.
2. Perra N, Gonçalves B, Pastor-Satorras R, Vespignani A (2012) Activity driven modeling of time varying networks. *Sci Rep* 2:469.
3. Krings G, Karsai M, Bernhardsson S, Blondel VD, Saramäki J (2012) Effects of time window size and placement on the structure of an aggregated communication network. *EPJ Data Sci* 1(1):1–16.
4. Read JM, Eames KTD, Edmunds WJ (2008) Dynamic social networks and the implications for the spread of infectious disease. *J R Soc Interface* 5(26):1001–1007.
5. Stehlé J, et al. (2011) Simulation of an SEIR infectious disease model on the dynamic contact network of conference attendees. *BMC Med* 9(1):87.
6. Salathé M, et al. (2010) A high-resolution human contact network for infectious disease transmission. *Proc Natl Acad Sci USA* 107(51):22020–22025.
7. Isella L, et al. (2011) What's in a crowd? Analysis of face-to-face behavioral networks. *J Theor Biol* 271(1):166–180.
8. Rocha LEC, Liljeros F, Holme P (2011) Simulated epidemics in an empirical spatiotemporal network of 50,185 sexual contacts. *PLOS Comput Biol* 7(3):e1001109.
9. Grannis R (2009) *From the Ground Up: Translating Geography into Community Through Neighbor Networks* (Princeton Univ Press, Princeton, NJ).
10. Eckmann JP, Moses E, Sergi D (2004) Entropy of dialogues creates coherent structures in e-mail traffic. *Proc Natl Acad Sci USA* 101(40):14333–14337.

11. Barabási A-L (2005) The origin of bursts and heavy tails in human dynamics. *Nature* 435(7039):207–211.
12. Onnela JP, et al. (2007) Structure and tie strengths in mobile communication networks. *Proc Natl Acad Sci USA* 104(18):7332–7336.
13. Golder SA, Macy MW (2011) Diurnal and seasonal mood vary with work, sleep, and daylength across diverse cultures. *Science* 333(6051):1878–1881.
14. Rybski D, Buldyrev SV, Havlin S, Liljeros F, Makse HA (2012) Communication activity in a social network: Relation between long-term correlations and inter-event clustering. *Sci Rep* 2:560.
15. Crandall DJ, et al. (2010) Inferring social ties from geographic coincidences. *Proc Natl Acad Sci USA* 107(52):22436–22441.
16. Milgram S (1974) The frozen world of the familiar stranger. *Psychol Today* 17: 70–80.
17. Paulos E, Goodman E (2004) The familiar stranger: Anxiety, comfort, and play in public places. *CHI Conf Proc* (ACM Press, New York), 223–230.
18. González MC, Hidalgo CA, Barabási A-L (2008) Understanding individual human mobility patterns. *Nature* 453(7196):779–782.
19. Song C, Qu Z, Blumm N, Barabási A-L (2010) Limits of predictability in human mobility. *Science* 327(5968):1018–1021.
20. Brockmann D, Hufnagel L, Geisel T (2006) The scaling laws of human travel. *Nature* 439(7075):462–465.
21. Schönfelder S, Axhausen KW (2010) *Urban Rhythms and Travel Behaviour: Spatial and Temporal Phenomena of Daily Travel* (Ashgate, Farnham, UK).
22. Eubank S, et al. (2004) Modelling disease outbreaks in realistic urban social networks. *Nature* 429(6988):180–184.
23. Wang P, González MC, Hidalgo CA, Barabási A-L (2009) Understanding the spreading patterns of mobile phone viruses. *Science* 324(5930):1071–1076.
24. Smieszek T, et al. (2011) Reconstructing the 2003/2004 H3N2 influenza epidemic in Switzerland with a spatially explicit, individual-based model. *BMC Infect Dis* 11:115.
25. Pelletier MP, Trépanier M, Morency C (2011) Smart card data use in public transit: A literature review. *Transp Res, Part C Emerg Technol* 19:557–568.
26. Rohani P, Zhong X, King AA (2010) Contact network structure explains the changing epidemiology of pertussis. *Science* 330(6006):982–985.
27. Pastor-Satorras R, Vespignani A (2001) Epidemic spreading in scale-free networks. *Phys Rev Lett* 86(14):3200–3203.
28. Watts DJ, Strogatz SH (1998) Collective dynamics of 'small-world' networks. *Nature* 393(6684):440–442.
29. Christakis NA, Fowler JH (2013) Social contagion theory: Examining dynamic social networks and human behavior. *Stat Med* 32(4):556–577.
30. Aral S, Muchnik L, Sundararajan A (2009) Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks. *Proc Natl Acad Sci USA* 106(51): 21544–21549.
31. Goffman W (1966) Mathematical approach to the spread of scientific ideas—The history of mast cell research. *Nature* 212(5061):449–452.
32. Vespignani A (2009) Predicting the behavior of techno-social systems. *Science* 325(5939):425–428.
33. Nowak MA (2006) Five rules for the evolution of cooperation. *Science* 314(5805): 1560–1563.
34. Palla G, Barabási A-L, Vicsek T (2007) Quantifying social group evolution. *Nature* 446(7136):664–667.
35. Ancel Meyers L, Newman MEJ, Martin M, Schrag S (2003) Applying network theory to epidemics: Control measures for Mycoplasma pneumoniae outbreaks. *Emerg Infect Dis* 9(2):204–210.
36. Christakis NA, Fowler JH (2007) The spread of obesity in a large social network over 32 years. *N Engl J Med* 357(4):370–379.