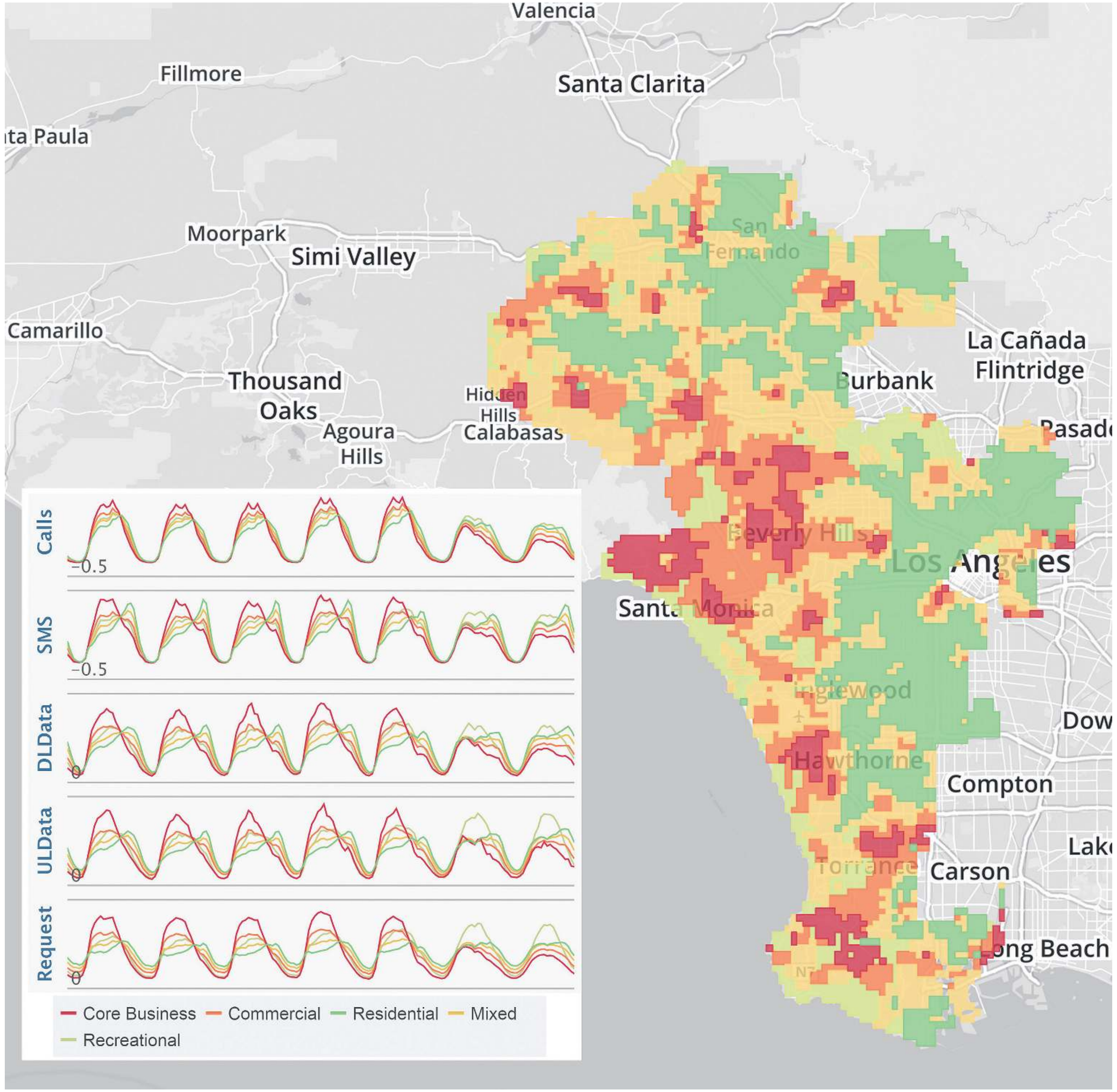


1. 人类活动时间线显示了洛杉矶的功能结构。

1. Timelines of human activity revealing functional structure of Los Angeles.

收稿时间 RECEIVED DATE: 2015-06-10
中图分类号 / TU986.2 文献标识码 / B

多城记
——全球多个城市人类活动特征可视化

A TALE OF MANY CITIES
— VISUALIZING SIGNATURES OF HUMAN ACTIVITY IN CITIES ACROSS THE GLOBE

丹尼尔 · 孔多尔，皮尔利克 · 提博尔特，塞巴斯蒂安 · 格劳文，伊斯特凡 · 格多尔，
西蒙 · 莫瑞茨，斯坦尼斯拉夫 · 索伯勒夫斯基，卡罗 · 拉蒂
/ Dániel KONDOR, Pierrick THEBAULT, Sebastian GRAUWIN, Istvan GÓDOR,
Simon MORITZ, Stanislav SOBOLEVSKY, Carlo RATTI

摘要

大数据在人类活动上的应用正在改变我们看待周围环境的方式。本文介绍了一种技术和可视化工具，其通过移动网络上的聚合活动计量，来获得形成城市结构的人类活动的信息。基于10个月的移动网络数据，我们可以在时间和空间维度上对人类活动模式进行比较，并根据不同地区的具体特征呈现出对应的“城市脉冲”。这项工具还可以依照人类活动的时间轴将社区归类为不同的功能聚类。这种途径及工具提供了从历史的，以及人类行为的实时动态更新记录的双重视角审视城市结构的新方式。这一在线工具呈现了对全球4座都市的分析结果，包括纽约、伦敦、香港和洛杉矶。

关键词

人类活动；大数据；移动网络活动档案；城市空间结构

ABSTRACT

The availability of big data on human activity is currently changing the way we look at our surroundings. In this article, we present a technique and visualization tool which uses aggregated activity measures of mobile networks to gain information about human activity shaping the structure of the cities. Based on ten months of mobile network data, activity patterns can be compared through time and space to unravel the “city’s pulse” as seen through the specific signatures of different locations. Furthermore, the tool allows classifying the neighborhoods into functional clusters based on the timeline of human activity. The approach and the tool provide new ways of looking at the city structure from historical perspective and potentially also in real-time based on dynamic up-to-date records of human behavior. The online tool presents results for four global cities: New York, London, Hong Kong and Los Angeles.

KEY WORDS

Human Activity; Big Date; Mobile Network Activity Profile; City Spatial Structure

译 王璐琦 田乐

TRANSLATED BY Luqi WANG Tina TIAN

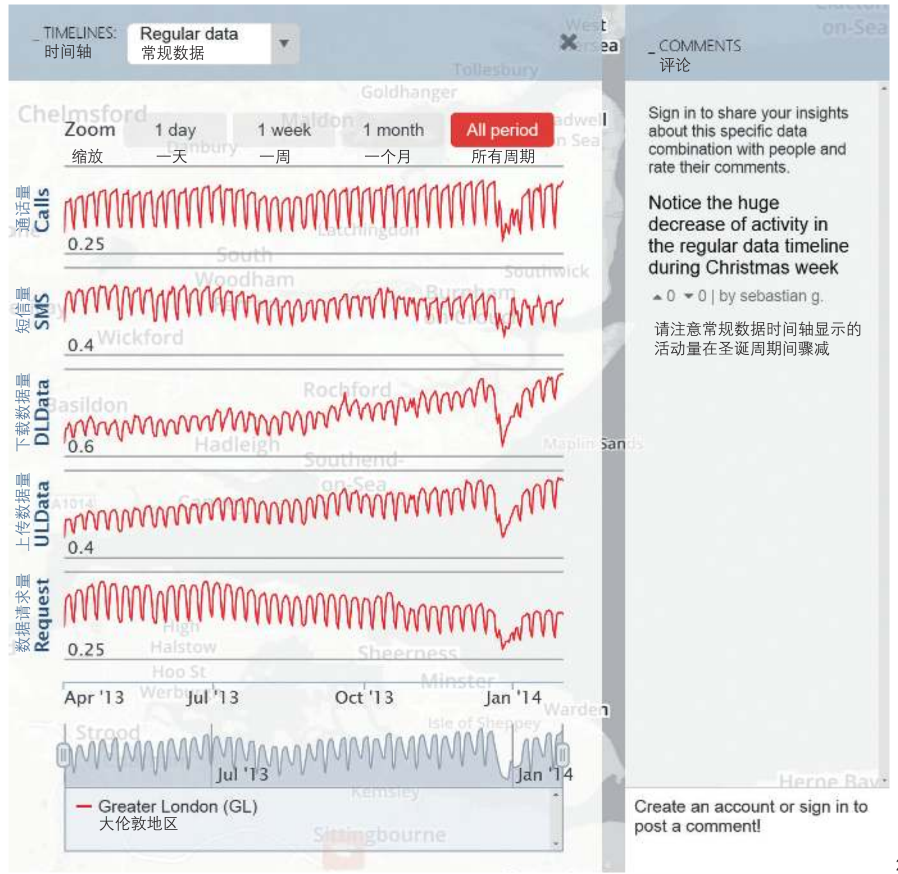
大数据在人类活动上的应用正在改变我们看待周围环境的方式。随着移动电话的高度普及，可以说几乎每一个人都随身携带了一个高精度的传感器，这让我们得以在空前的尺度上对人类活动进行监测和动态分析。在本文中，我们介绍了一种技术和可视化工具，其通过移动网络上的聚合活动计量，来获得塑造城市结构的人类活动的信息。基于10个月的移动网络数据，我们可以在时间和空间维度上对人类活动模式进行比较，并根据不同地区的具体特征呈现出对应的“城市脉冲”。这项工具还可以依照人类活动的时间轴将社区归类为不同的功能聚类，对城市中实际的土地利用格局提出有价值的参考意见。如此一来，这种途径及工具提供了从历史的，以及人类行为的实时动态更新记录的双重视角审视城市结构的新方式。这一在线工具呈现了对全球4座都市的分析结果，包括纽约、伦敦、香港和洛杉矶。

在当前如此飞速发展的世界中，需要恰当的城市规划来适应城市转型中日益增多的动态变量。城市规划者面临的一个重要问题就是要了解不同社区的功能，这对于理解社区的需求来说非常关键。传统意义上讲，城市拥有一个基于官方记录的土地利用分类数据库；尽管这类资料通常具有较高的精度，但这也在本质上制约了这些资料随动态变化而更新的能力。今天，这些挑战迎来了新的解决方案：通过利用匿名的移动手机数据，实现人类活动的动态管理，并以此来感知城市中各个社区的具体特征。随着对这些新的可能性的探索，近年来移动手机数据在研究城市中人们的行为足迹[1][2]、人口流动性[3]-[5]、城市结构[6]-[8]，以及包括区域划分[9][10]在内的其他方面的运用与日俱增。

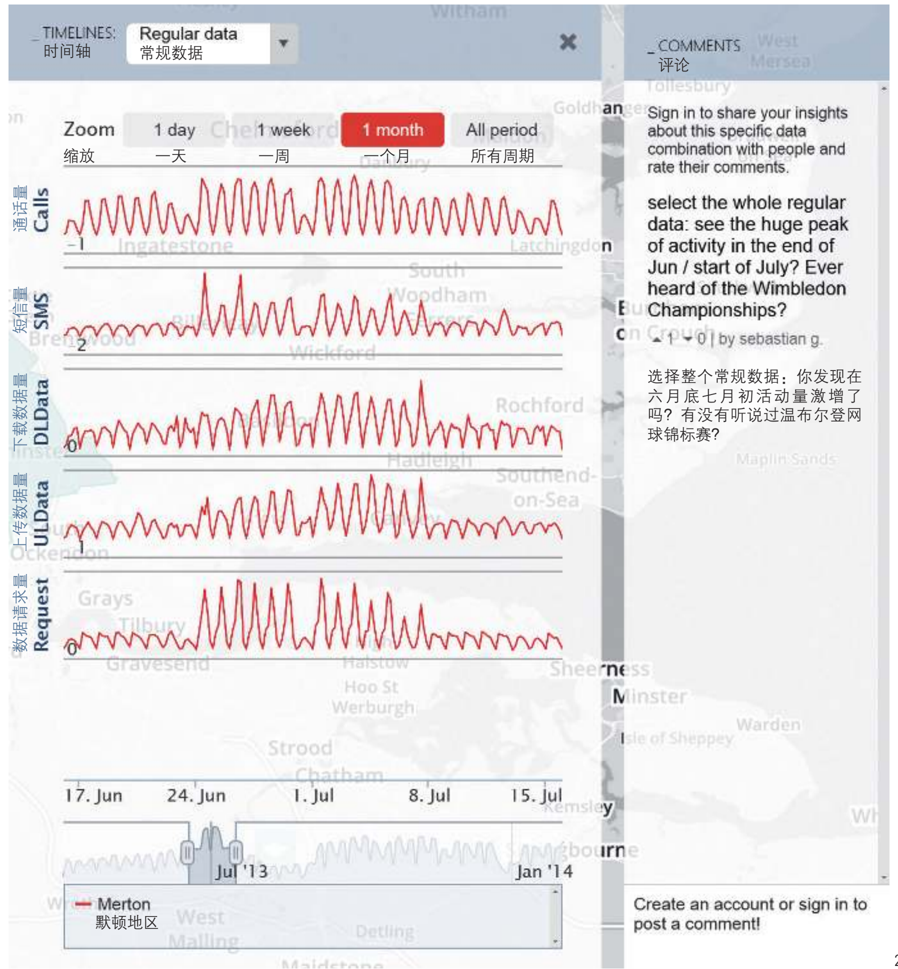
本文中介绍的这项工具①由MIT可感城市实验室和爱立信集团合作研发而成，其可以通过移动网络的聚合活动数据来对纽约、伦敦、香港和洛杉矶这4座城市的结构进行分析比对。②该工具使用了由市场占用率较高的几家移动网络经营商提供的这4个城市2013年4月至2014年1月期间的聚合数据。这些数

① 该工具可经由以下网址获取：http://manycities.org。其主要研发者为Pierrick Thebault和Sebastian Grauwin。

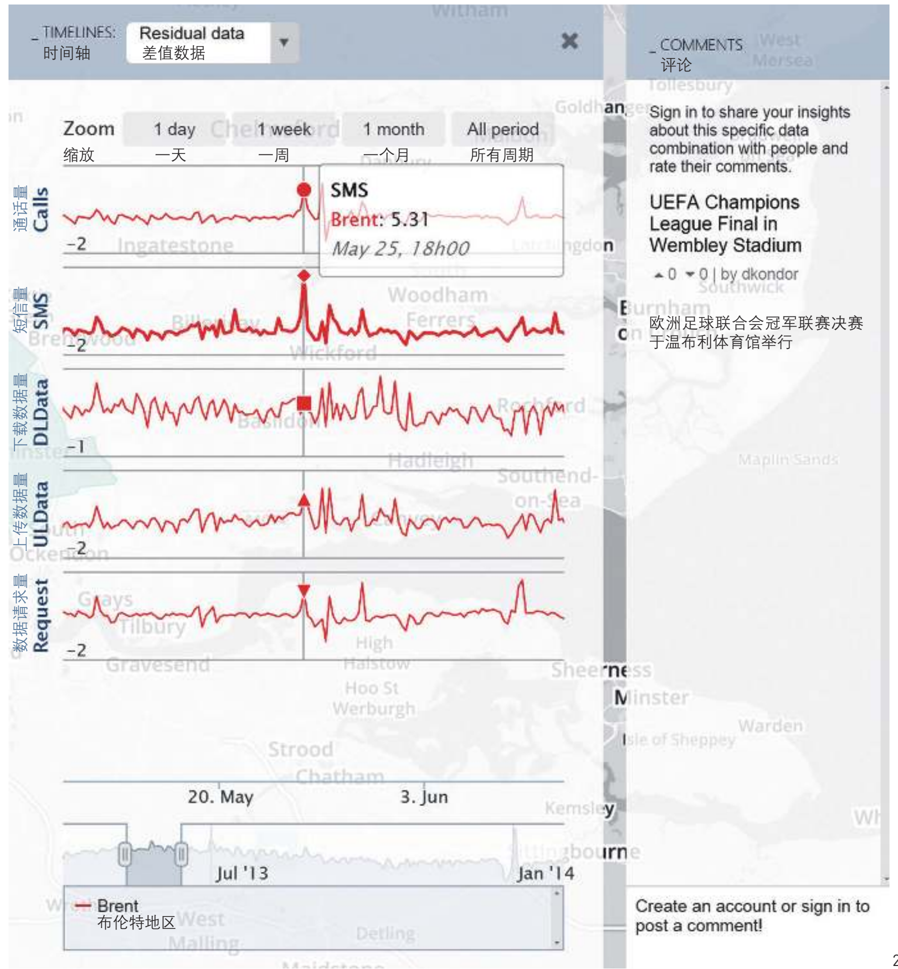
② 更多的研究成果可参见参考文献[11]。



2-1



2-2



2-3

2. 在大伦敦地区使用移动网络活动来辨别长期趋势和特殊事件。
2-1. 全市范围的聚合数据显示了数据流量的平稳增长状态；但是在圣诞期间，所有类型的活动数量都急剧减少。
2-2. 以默顿地区为例，我们可以看到温布尔登网球锦标赛对该地区网络活动的影响。
2-3. 在其他数据中，我们可以通过显著的数据峰值来辨别特殊事件。图中以布伦特地区为例，这里是温布利体育馆——伦敦一个重要的体育赛事场馆——的所在地；时间轴显示了2013年欧洲足球联合会冠军联赛决赛当天的数据信息（2013年5月25日）。

2. Identifying long-term trends and events in mobile network activity in the Greater London area.
2-1. City-wide aggregated data shows a steady increase in data traffic; also notice the large drop in all activity types around Christmas.
2-2. Concentrating on the Merton district, the effect of the Wimbledon Tennis Championship can be seen in network activity.
2-3. Looking at the residual data, special events can be identified by large peaks. The example given here is the Brent district, which contains Wembley Stadium, a major sports venue in London; the timeline is focused on the date of the 2013 UEFA Champions League Final.

项目负责人：
Carlo Ratti
项目团队：
Pierrick Thebault、Sebastian Grauwin、Stanislav Sobolevsky、Istvan Gódor、Simon Moritz、Dániel Kondor
发布时间：
2014年11月21日
相关网站：
http://manycities.org

PROJECT LEADER:
Carlo Ratti
RESEARCH TEAM:
Pierrick Thebault, Sebastian Grauwin, Stanislav Sobolevsky, Istvan Gódor, Simon Moritz, Dániel Kondor
LAUNCH DATE:
November 21, 2014
WEBSITE:
http://manycities.org

据来自于对移动通讯天线层面上聚合的人们的手机活动所进行测量的数据（每隔15分钟测量一次），其中包括每个天线每15分钟时窗内的通话和发送短信的数量、上传和下载的数据量，以及数据请求量（发生数据传输的次数）。在不涉及任何用户的隐私信息的情况下，这些聚合数据能够提供充足的以小型社区为单位的使用格局的具体数据信息。我们对数据进行进一步的空间聚合处理，以消除由于蜂窝网络运用对天线层面上的聚合活动数量的影响。在大部分的分析中，我们对不同地点的时间序列结果进行了规范化处理；我们希望绝大部分信息可以从时间序列形态的对比中获得，例如，一天或一周内的活动是如何分布的。这个工具也可以用于单独观看活动量的空间分布。

在这个在线工具中，我们通过三种模式来展示这些数据的潜在用途：时间序列、空间聚类和密度地图。首先，时间序列模式促进了在城市区域尺度上对人类活动计量的探索，该模式可能的用途如图2~4所示。在这个模式中，人们可以查看计量时段内的单个时间序列，从而确认长期趋势和特殊事件，也可以根据每周的均值来对比不同社区或不同城市中的典型活动格局。长期时间序列的显著特征包括数据流量的稳定增长、节假日数据的显著变化，以及某些重要事件——例如伦敦的温布尔登网球锦标赛或足球比赛——的具体特征。通过观测常规的一周时间序列，可以容易地辨别出城市核心商务区和居住区的典型活动格局。除了查看活动的长期时间序列和每周均值外，这种模式还可用于观察活动之间的差别，例如，某一星期的活动格局与平均值的差异。特殊事件亦可以通过这些时间序列的峰值来判断（图2-3）。

在线工具的第二种模式——空间聚类模式，可以根据移动网络活动时间序列将城市划分为不同的功能聚类，并可以根据分类结果进行交互探索。这项分析将反映网络活动的、以周为单位的常规时间序列聚合在一个规则网格中。每一个功能聚类可以通过该聚类的平均时间序列和一张显示该聚类的网格像素地图来展示。举例来说，纽约城的功能聚类如图5-1所示。根据平均聚类时间序列，我们可以将它们分类——例如，在核心商务区中，活动量在平常工作时段内较高，在非工作时段内较低；在居住区中，活动量的峰值出现在晚上；在休闲场所和公园中，周末的日间时段会出现显著的活动量峰值。我们还可以对不同城市的聚类进行比较。例如，如图5-2所示，我们对纽约和香港的核心商务区聚类与居住区聚类进行了对比。显而易见，两个城市商务区聚类的活动特征相似度比居住区聚类的活动特征相似度高。③

密度地图——该在线工具的第三个模式——可以对城市中活动量的空间分布进行展示。它同样可以用于观察不同类型活动量之间的比例，从而得知城市中不同地区的活动偏好。例如，图6中展示了香港的通话数量的空间分布。

我们相信，本文所介绍的这一工具可以成为城市规划决策支持的宝贵补充，也可以为研究者、城市权益相关者以及公众在理解人们如何使用城市的问题上提供重要的参考意见。LAF

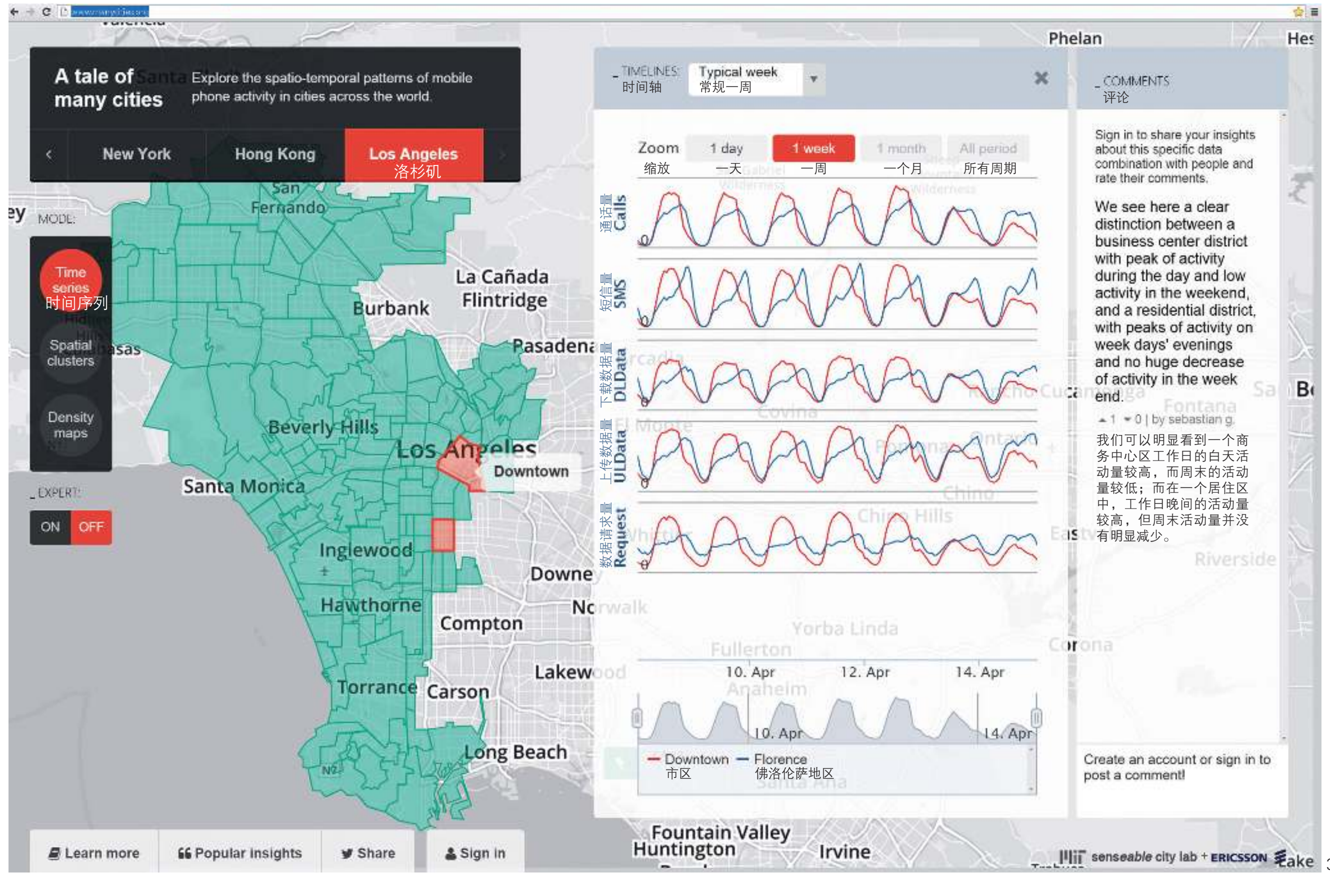
③ 对于聚类结果更为具体的分析，包括对聚类的解读的深入探讨、将分析结果与基于数据普查的传统数据所进行的对比，以及对这4座城市差异性所进行的系统分析可参见参考文献[11]。

致谢

感谢爱立信集团提供了聚集手机活动日志。同时感谢Zsófia Kallus对本项目的启发，以及爱立信集团、MIT SMART计划、埃森哲公司、液化空气集团、西班牙外换银行、可口可乐公司、阿联酋联合电信公司、意大利国家电力公司基金会、2015米兰世博会、法罗里奥集团、美国利宝相互保险公司、伍德法布罗市、大众汽车电子研究室、优步和MIT可感城市实验室的所有团体成员对本研究的支持。

① The tool is available at http://manycities.org. Leading developers: Pierrick Thebault and Sebastian Grauwin.

② Further research insights coming from such analysis are presented in Towards a Comparative Science of Cities: Using Mobile Traffic Records in New York, London, and Hong Kong, by Sebastian Grauwin, Stanislav Sobolevsky, Simon Moritz, Istvan Gódor, and Carlo Ratti.



3

The availability of big data on human activity is currently changing the way we look at our surroundings. With the high penetration of mobile phones, nearly everyone is already carrying a high-precision sensor providing an opportunity to monitor and analyze the dynamics of human movement on unprecedented scales. In this article, we present a technique and visualization tool which uses aggregated activity measures of mobile networks to gain information about human activity shaping the structure of the cities. Based on ten months of mobile network data, activity patterns can be compared through time and space to unravel the "city's pulse" as seen through the specific signatures of different locations. Furthermore, the tool allows classifying the neighborhoods into functional clusters based on the timeline of human activity, providing valuable insights on the actual land use patterns within the city. This way, the approach and the tool provide new ways of looking at the city structure from historical perspective and potentially also in real-time based on dynamic up-to-date records of human behavior. The online tool presents results for four global cities: New York, London, Hong Kong and Los Angeles.

Increasing dynamics of urban transformation in a rapidly developing world calls for adaptive urban planning. One of the central questions for city planners is to know the function of different neighborhoods, which is crucial for understanding their needs. Traditionally, cities maintain a database of land-use classification, which is based on official records. While such data is generally of very high resolution, its nature limits the ability to follow changes dynamically. But today's challenges come together with new solutions to them: the use of anonymized mobile phone data enables dynamic measurement of human activity and sensing the unique signatures of each neighborhood of a city based on that. Exploiting these new possibilities, recent years saw increasing usage of mobile phone data for better understanding people's presence around the city[1][2], human mobility[3]-[5], structure of the city[6]-[8] and many other applications, including regional delineation[9][10].

The presented tool[1], developed in collaboration between MIT SENSEable City Laboratory and Ericsson, allows the comparative analysis of city structure with

3. 通过比较常规一周的活动档案，我们可以看出中央商务区（位于洛杉矶市中心，红线）和某居住区（位于佛洛伦萨地区，蓝线）之间的显著差异。

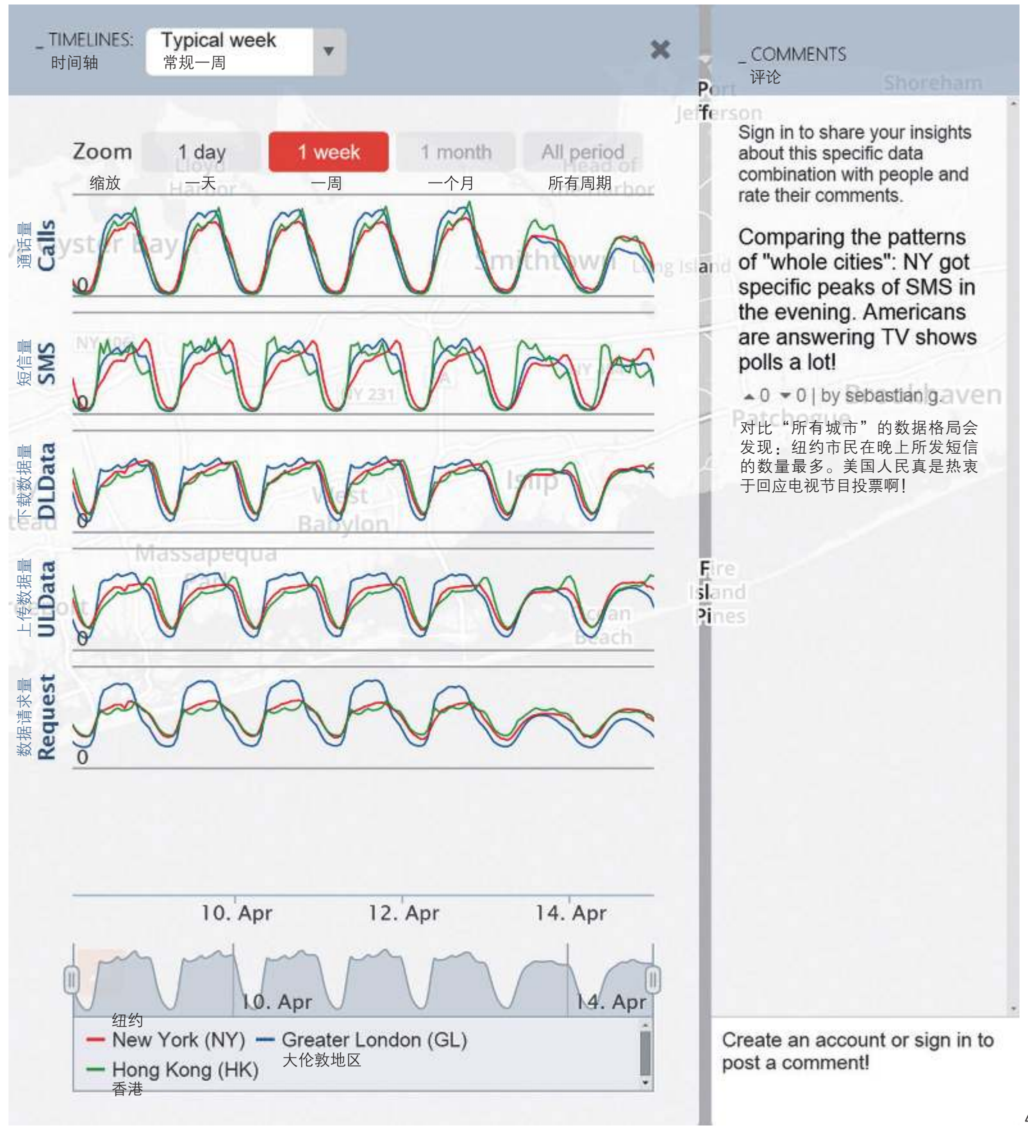
3. Comparing the activity profile of the typical week, there is a significant distinction between a central business area (downtown Los Angeles, the red curve) and a residential neighborhood (Florence, in blue).

4. 纽约、香港、伦敦这三座城市全市范围内的聚集活动档案对比。图中虽然显示出各地人们的作息规律具有很高的共性，但仍存在显著的差异。以短信发送量峰值的出现时间为例，香港出现在早上，纽约出现在晚上，而伦敦则出现在中午——或者说，相较于另两个城市的峰值来说，伦敦的数据流量夜间骤减。我们推测后者的原因是由于伦敦的蜂窝数据流量收费非常高昂，以至于人们在家时会倾向使用价格相对低廉的无线网络。

4. Comparing city-wide aggregated activity profiles in three cities. Apart from showing a high degree of similarity corresponding to the natural circadian rhythm of humans, there are notable differences, e.g. text message activity peaks in the morning in Hong Kong, in the evening in New York and midday in London or the abrupt decrease of data traffic in London in the evenings contrasting the peaks in the other two cities. We speculate that the latter observation is the effect of cellular data traffic being especially expensive in London, prompting people to switch to much cheaper wifi networks when at home.

the use of aggregated mobile network activity data in four global cities: New York, London, Hong Kong and Los Angeles.[2] The tool uses aggregated data collected between April 2013 and January 2014 in the four cities provided by mobile network operators with representative market shares. The data includes measurements of human phone activity aggregated and collected on the level of mobile antennas within 15-minute time intervals. These include the number of calls placed and text messages sent, the amount of data downloaded and uploaded, and the number of data requests (number of individual times data transfer was initiated) for each antenna in each 15-minute time window. This way, the aggregated data does not include any sensitive customer information, but provides enough detail about the typical usage patterns on the scale of small neighborhoods. We proceed by further spatial aggregation so as to account for the variation of antenna-level activity volumes due to technical constraints of cellular network operation. For most of the analysis, we normalize the resulting time series of different locations; we expect that the most information can be gained from comparing the shape of time series, i.e. how activity is distributed over a day or week. Looking at the spatial distribution of activity volumes is possible in a separate usage mode in the tool.

Possible usage of the data is demonstrated via three possible usage modes in the online tool: time series, spatial clusters and density maps. The first, time series mode facilitates the exploration of human activity measures on the level of city districts; possible usage is illustrated in Figures. 2, 3 and 4. In this mode, one can look at individual time series throughout the measurement period making possible to identify long-term trends and special events, and weekly averages enabling the comparison of typical activity patterns in different neighborhoods or in different cities.

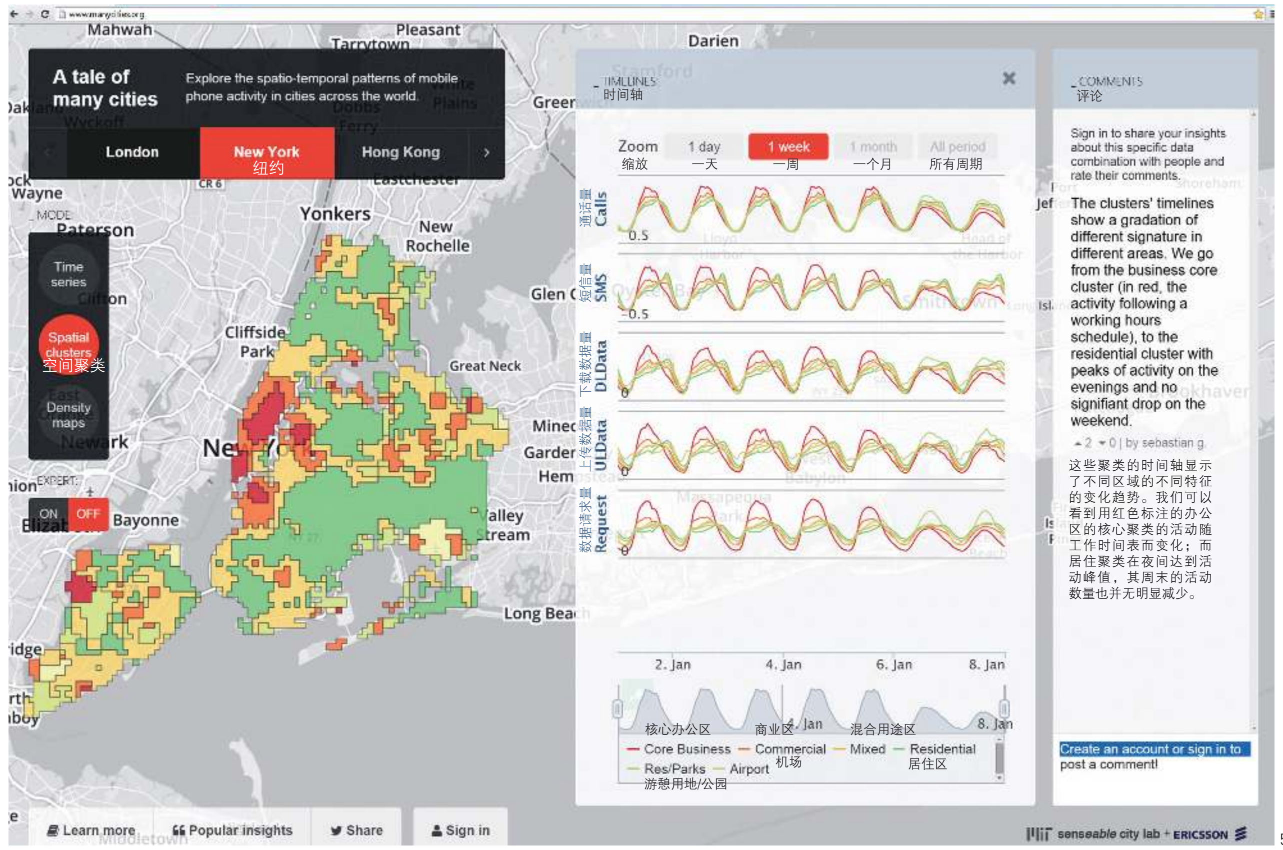


4

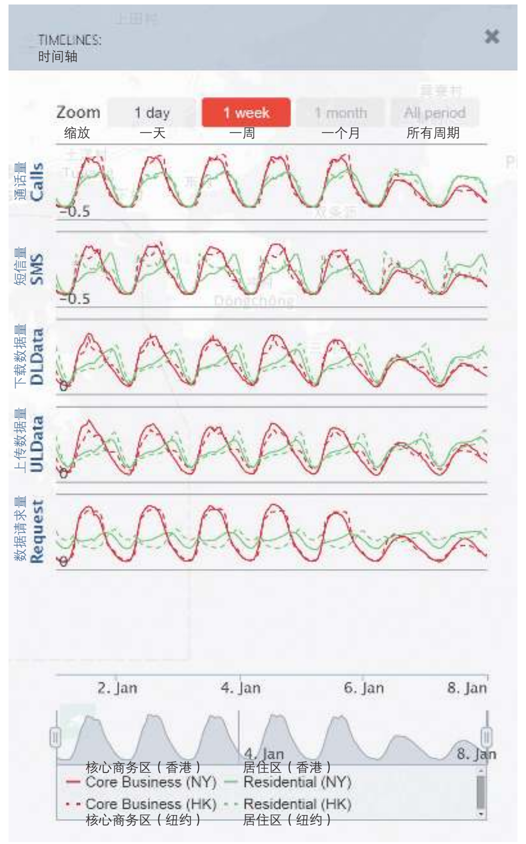
Notable features in long-term time series include the steady increase in data traffic, the distinct effect of holidays, and specific signatures of some important events like the Wimbledon tennis championship or football matches in London. Looking at typical week time series, patterns representative of cities' business centers and residential areas can be easily identified. Apart from looking at the long-term time series of activity and weekly averages, it is also possible to inspect the residual activity, i.e. difference of activity on a specific week from the average. Special events can also be noticed by peaks

① The tool is available at http://manycities.org. Leading developers: Pierrick Thebault and Sebastian Grauwin.

② Further research insights coming from such analysis are presented in Reference [11].



5-1



5-2

in these time series (Fig. 2-3).

The second, spatial clusters mode of the online tool allows the interactive exploration of results on decomposing the cities into functional clusters based on the mobile network activity time series. This analysis was performed on the typical week time series of network activity aggregated into a regular grid. Each of the functional clusters can then be represented with a cluster-wide average time series and a map displaying the grid pixels which belong to it. As an example, the functional clusters of New York City are displayed in Figure 5-1. Looking at the average cluster time series, we can classify them as e.g. core business, where activity is especially high in normal working hours and low otherwise; residential, where the peak of daily activities is in the evenings; leisure and parks, where activity has noticeable peaks in weekends during the day. It is also possible to compare clusters from different cities. For example, in Figure 5-2, we show comparison of clusters identified as core business and residential in New York and Hong Kong. It is remarkable that the activity signatures of the two business clusters seem much more similar than the signatures of the two residential clusters. ③

The third mode of the online tool presents the spatial distribution of activity volumes in the cities. It also allows inspecting the ratio of volumes of different activity types, giving an overview of which type of activity is preferred in different parts of the cities. An example is displayed in Figure 6, showing the spatial distribution of call volumes in Hong Kong.

We believe the presented tool to be a valuable supplement for urban planning decision support as well as for providing important insights on the way people use the city for researchers, urban stakeholders and general public. **LAF**

③ A more detailed analysis of results of clustering, including in-depth discussion of the interpretation of the clusters, comparison of results with traditional, census-based data, and a systematic analysis of differences among the four cities is presented in Reference [11].

ACKNOWLEDGEMENTS

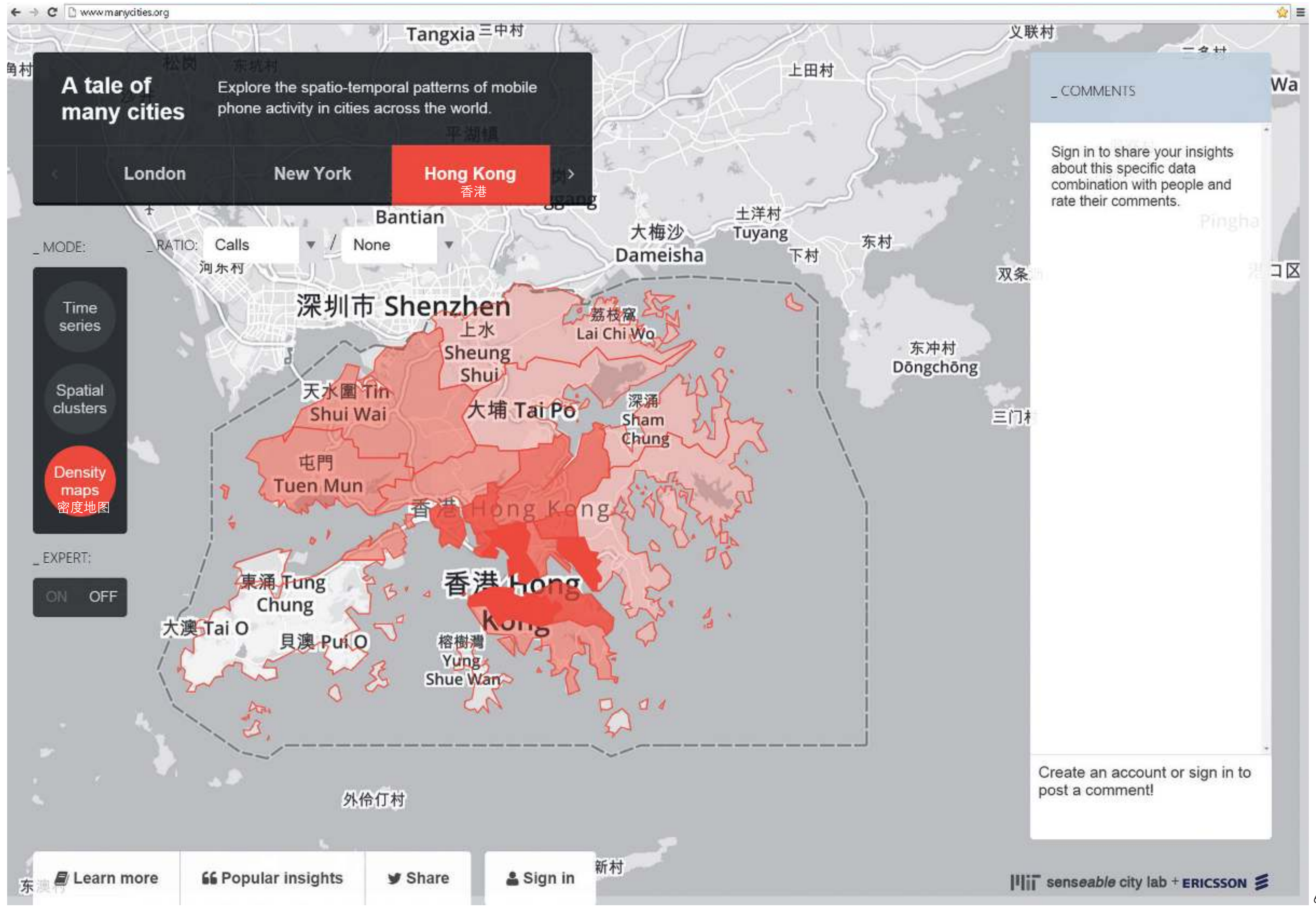
The authors wish to thank Ericsson for providing the aggregated phone activity records. We also thank Zsófia Kallus for stimulating discussions. We further thank Ericsson, MIT SMART Program, Accenture, Air Liquide, BBVA, The Coca Cola Company, Emirates Integrated Telecommunications Company, The ENEL foundation, Expo 2015, Ferrovial, Liberty Mutual, The Regional Municipality of Wood Buffalo, Volkswagen Electronics Research Lab, UBER and all the members of the MIT Senseable City Lab Consortium for supporting the research.

5. 基于常规活动时间序列相似性形成的功能聚类显示了城市重要的结构特征。
5-1. 图中显示了在纽约市的功能聚类及相关评价时间序列。
5-2. 纽约和香港的商务与居住聚类的代表性时间序列对比。
6. 香港平均通话量空间分布

5. Functional clusters formed on the similarity of typical activity time series uncover important structural characteristics of cities.
5-1. Map showing the functional clusters found in New York and a plot of the associated average time series.
5-2. Comparing representative time series of business and residential clusters in New York with the same in Hong Kong.
6. Spatial distribution of average volume of calls placed in Hong Kong.

REFERENCES

[1] Reades, J., Calabrese, F., Sevtsuk, A., Ratti, C. (2007). Cellular census: Explorations in urban data collection. Pervasive Computing, IEEE, 6(3), 30-38.
[2] Deville, P., Linard, C., Martin, S., Gilbert, M., Stevens, F. R., Gaughan, A. E. (2014). Dynamic population mapping using mobile phone data. Proceedings of the National Academy of Sciences, 111(45), 15888-15893.
[3] González, M. C., Hidalgo, C. A., Barabási, A.-L. (2008). Understanding individual human mobility patterns. Nature, 453(7196), 779-782.
[4] Hoteit, S., Secci, S., Sobolevsky, S., Ratti, C., Pujolle, G. (2014). Estimating human trajectories and hotspots through mobile phone data. Computer Networks, 64, 296-307.
[5] Amini, A., Kung, K., Kang, C., Sobolevsky, S., Ratti, C. (2014). The impact of social segregation on human mobility in developing and industrialized regions. EPJ Data Science, 3(1).
[6] Reades, J., Calabrese, F., Ratti, C. (2009). Eigenplaces: analysing cities using the space-time structure of the mobile phone network. Environment and Planning B: Planning and Design, 36(5), 824-836.
[7] Pei, T., Sobolevsky, S., Ratti, C., Shaw, S. L., Li, T., Zhou, C. (2014). A new insight into land use classification based on aggregated mobile phone data. International Journal of Geographical Information Science, 28(9), 1988-2007.
[8] Louail, T., Lenormand, M., Cantú-Ros, O. G., Picornell, M., Herranz, R., Frias-Martinez, E., Ramasco, José J, Barthélemy, M. (2014). From mobile phone data to the spatial structure of cities. Scientific Reports, 4, 5276.
[9] Ratti, C., Sobolevsky, S., Calabrese, F., Andris, C., Reades, J., Martino, M., Claxton, R., Strogatz, S. H. (2010). Redrawing the map of Great Britain from a network of human interactions. PloS one, 5(12), e14248.
[10] Sobolevsky, S., Szell, M., Campari, R., Couronné, T., Smoreda, Z., Ratti, C. (2013). Delineating geographical regions with networks of human interactions in an extensive set of countries. PloS one, 8(12), e81707.
[11] Grauwin, S., Sobolevsky, S., Moritz, S., Gódor, I., Ratti, C. (2015). Towards a Comparative Science of Cities: Using Mobile Traffic Records in New York, London, and Hong Kong. In M. Helbich, J. J. Arsanjani, M. Leitner. Computational approaches for urban environments (pp. 363-387). New York: Springer International Publishing.



6