Impact of Human Mobility on Social Networks

Dashun Wang and Chaoming Song

Abstract: **Mobile phone carriers face challenges from three synergistic dimensions: Wireless, social, and mobile. Despite significant advances that have been made about social networks and human mobility, respectively, our knowledge about the interplay between two layers remains largely limited, partly due to the difficulty in obtaining large-scale datasets that could offer at the same time social and mobile information across a substantial population over an extended period of time. In this paper, we take advantage of a massive, longitudinal mobile phone dataset that consists of human mobility and social network information simultaneously, allowing us to explore the impact of human mobility patterns on the underlying social network. We find that human mobility plays an important role in shaping both local and global structural properties of social network. In contrast to the lack of scale in social networks and human movements, we discovered a characteristic distance in physical space between 10 and 20 km that impacts both local clustering and modular structure in social network. We also find a surprising distinction in trajectory overlap that segments social ties into two categories. Our results are of fundamental relevance to quantitative studies of human behavior, and could serve as the basis of anchoring potential theoretical models of human behavior and building and developing new applications using social and mobile technologies.**

Index Terms: **Clustering, heterogeneous network, human mobility, mobile phones, percolation, scale-free network, social network.**

I. INTRODUCTION

SOCIAL networks have attracted significant interest across multiple disciplines in recent years, largely due to their critical role in a wide range of applications [1]–[7]. Despite recent explosion of research on social networks, the bulk of work has primarily focused on the social space, leaving its interplay with the physical space largely underexplored. Yet, in an accelerating number of settings, we are witnessing emerging convergence in social and mobile technologies, fueling rapid advances in areas as broad as marketing, security and communications. For example, location-based social networking services offer information sharing that enables new ways in marketing, connecting with friends, and recommending services [8], [9]. Mobile phone carriers face challenges from three synergistic dimensions: wireless, social, and mobility [10]–[12]. Together these dimensions are imperative to determine their key service safety, operability and profitability, from data transmission efficiency and service reliability to vulnerability and security. Therefore, as our society becomes globally interconnected, the social and physical space—social connections between individuals and their mobility—no longer exist in isolation. Rather they increasingly interact with and depend on each other. To truly harness and unleash the potential of social and mobile technologies, we need to develop a quantitative framework of the interplay between social networks and human mobility patterns.

Our knowledge about the interplay between social networks and human mobility patterns is limited, partly due to the difficulty in obtaining large-scale dataset that could offer at the same time social and mobile information across a substantial population over an extended period of time. This situation is changing drastically, however, thanks to the ever-increasing availability of detailed traces of human behavior [13], [14], from mobile-phone records to global-positioning-system (GPS) data to location based social networking services. Take mobile phones as an example: On one hand, mobile phones are carried by individuals during their daily routines, acting as an excellent proxy to capture individual human trajectories; On the other hand, with almost 100% mobile-phone penetration in industrial countries, mobile communications offer us a comprehensive picture of social interactions within a society [10], [11], [15]–[24]. Therefore, these large-scale datasets, capturing time resolved locations of individuals and their interactions, are not only fueling rapid advances in our understanding of individual mobility patterns, from empirical analyses to modeling tools and frameworks, but also offer unique opportunities to jointly explore individual trajectories and social communications [8], [20], [25], [26]. Indeed, there has been growing evidence examining physical space as a determinant of social ties. By assuming that individuals have a fixed time-independent location in space, like their home or work addresses, previous research examined the probability of a social tie between them given the distance between their locations, finding a robust power law decay in such probabilities across different contexts [27]–[30]. In the dynamic dimension of physical space, researcher discovered that, not only are individual trajectory overlaps predictive for the formation of social ties [9], [20], [31] location information of social neighbors is also revealing about individuals' whereabouts [8]. While these new results focused on dyadic, if microscopic, social interactions, they document first few connections between human mobility and social networks, portending to a quantitative framework between these two areas.

Here, we take advantage of a massive, longitudinal mobile phone dataset and ask the question: to what degree do our mobility patterns impact the dynamical and structural properties of our social networks? On both areas of social networks and human mobility, we establish a quantitative framework by carrying out a series of analyses that attend to both static and dynamic patterns. Through these analyses, we offer new evidence and insights of how social networks and human mobility correlate and interplay with each other in multiple levels of granularity. We

Manuscript received January 9, 2015.

D. Wang is with the College of Information Sciences and Technology, Pennsylvania State University, University Park, PA 16802 USA, email: dwang@ist.psu.edu.

C. Song is with the Department of Physics, University of Miami, Coral Gables, FL 3314 USA, email: c.song@physics.miami.edu.

Digital object identifier 10.1109/JCN.2015.000023

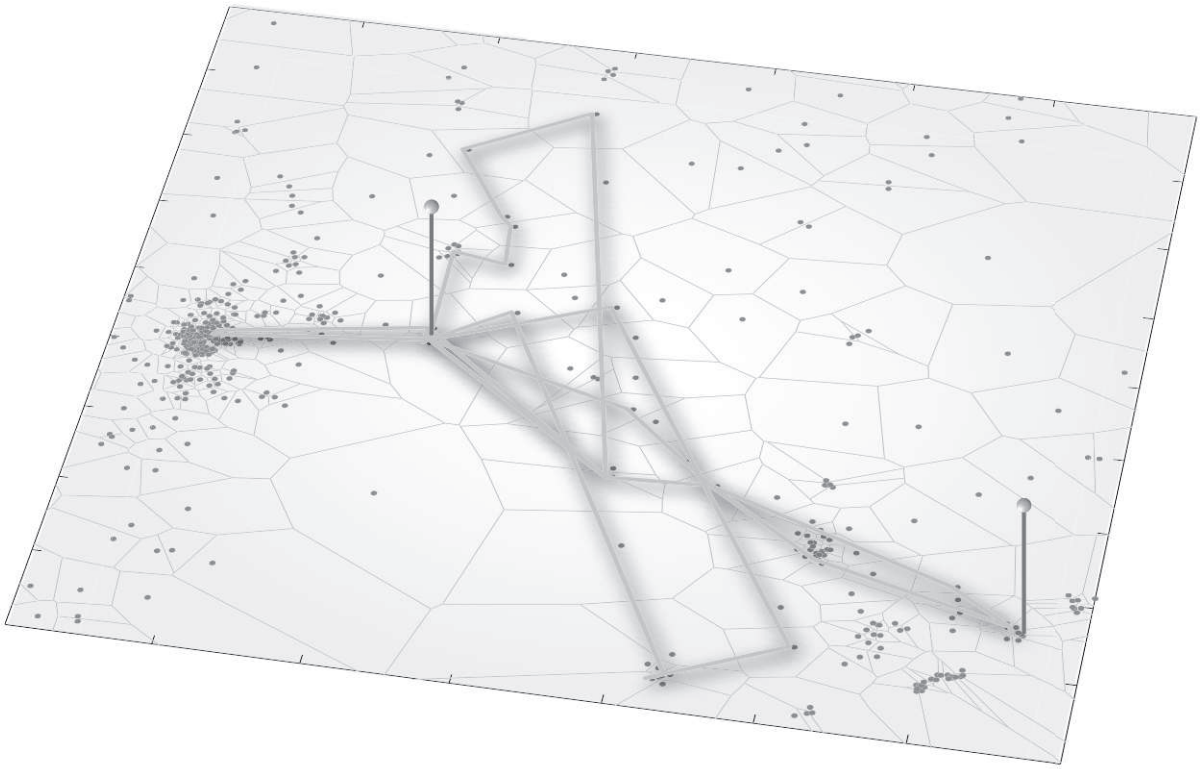

Fig. 1. Three-month trajectory of one mobile phone users.

find that human mobility plays an important role in shaping both local and global structural properties of social network. We find the emergence of spatial segmentation that reveals higher order regularities behind previously well-established results by studying social space alone, affecting both structural and dynamical properties of social network. Our results offer new and significant empirical evidence in connecting social networks and human mobility.

The remainder of this paper is organized as follows. Section II provides detailed description of the mobile phone dataset. In Section III we explore the basic characteristics of social networks and human mobility pattern, respectively. Section IV studies the co-locations between trajectories of pair of friends. Section V exams the proximity between human mobility and social networks in orders of increasing sophistication, from lower to higher order connections (a social tie to a local clustering). Sections VI investigates how human mobility impact underlying social network, for static and dynamical properties respectively. Section VII concludes this paper.

II. DATA DESCRIPTION

In the past few years, a collaboration with an European mobile communication company grants us access to anonymized country wide mobile phone dataset, containing call records of their customers and collected for billing purposes. Being mobile communication records, the dataset naturally offers information on wireless social communications through phone calls and text messages between individuals. At the same time, whenever a mobile communication was initiated, the dataset also records the information of the mobile tower that routed the phone call or text message, whose geographic coordinates allow us to pinpoint the location of the individual at the time when s/he initiated the communication. Hence, this data allows us to reconstruct the daily trajectory of each mobile phone user for an extended period of time. The dataset is massive and of excellent Longitudinality: It covers activities of 10 million customers for more than five years. In this study, we used data from two consecutive years. The database offers detailed empirical observation of human mobility. For example, in Fig. 1 we show the three month trajectory of one users in our database, illustrating the nature of the mobility patterns that can be extracted from the dataset. The different cell phone towers are denoted as grey dots, and the Voronoi cell in grey marks the approximate reception area of each tower. Note that the trajectories are time resolved. That is, each time a user makes a call, the closest tower that routes the call is recorded, serving as a proxy of the user's appropriate location at the moment.

There are two caveats about the dataset that are particularly worth mentioning: (a) The user location is subject to the spatial resolution of the mobile phone towers, which on average results in uncertainties around 1-3 km. (b) User's location is only recorded when the user uses the phone, hence we have no knowledge of the user's whereabouts between these active sessions. To cope with such caveats, improve temporal resolution of location sampling, and ensure our results are not affected by them, we constructed a second dataset that contains 50,000 individuals, chosen from the 10 million mobile phone users based

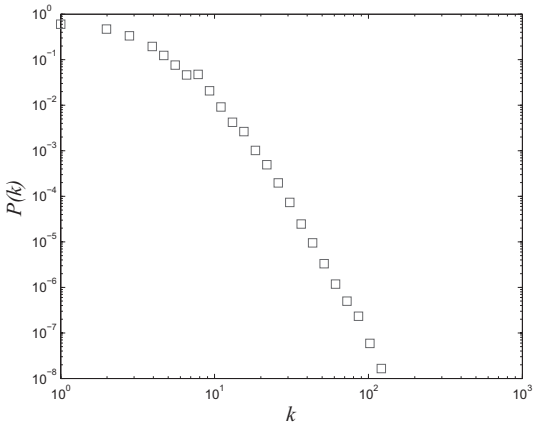

Fig. 2. Degree distribution of social network, where degree k measures the number of contacts of each individual has.

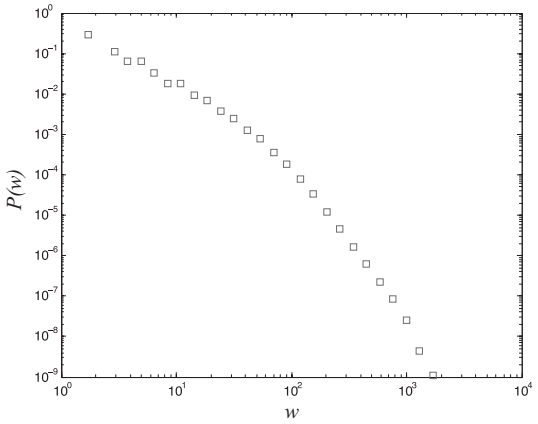

Fig. 3. Weight distribution of social network, where weight w measures the frequency of contacts between two individuals.

on their activity levels. We followed the same criteria as in [19], [20] and focus our analysis on this sample of individuals.

III. BASIC CHARACTERISTICS OF SOCIAL NETWORKS AND HUMAN MOBILITY

The dataset offers simultaneously individuals' communication patterns and their trajectories. In the following we describe basic characteristics in these two folds.

A. Social Network Characteristics

Fig. 2 shows the degree distribution $P(k)$ of the underlying social network where degree k measures the number of contacts of each individual has. We found that the degree distribution follows a power law tail

$$P(k) \sim k^{-\gamma}. \tag{1}$$

Similarly, the weight w that counts the frequency of contacts between two individuals (a social tie) is also broadly distributed (Fig. 3), following

$$P(w) \sim w^{-\beta}. \tag{2}$$

These two quantities indicate the scale-free nature of the social graph [2], [5], [32].

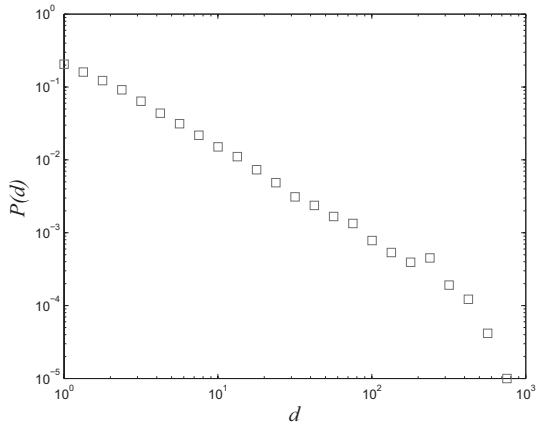

Fig. 4. Distance distribution between social ties, where distance d measures the geographic separation between two individuals' most likely locations.

B. Human Mobility Characteristics

To characterize spatial aspects of mobile users, we define individual's "home" as his/her most frequent location during the three-month records. For each social tie observed in the dataset we measured the distance d separated the corresponding pair. Fig. 4 plots the distance distribution $P(d)$, showing that

$$P(d) \sim d^{-\alpha} \tag{3}$$

indicating that it is more likely to form a wireless communication in a local region compared to a distant one, e.g., social ties are not uniformly distributed on the space [27]–[30].

IV. CO-LOCATION OF MOBILE PHONE USERS

The trajectories of two pairs of friends are shown in Fig. 5. We choose a pair of friends whose whereabouts are mostly in the same city. Their weeklong trajectories are shown in Fig. 5(a) in red and green, respectively. Although they communicate with each other frequently on the phone, their paths seldom intersect, indicating that they never, (if ever) have face-to-face interactions. However, a different pair of friends (Fig. 5(b)) displays widely different mobility patterns from the one in Fig. 5(a). Despite of some unique places visited by each user, their paths share a large number of towers in common. Furthermore, among the commonly visited locations, the percentage of time each user spent in the vicinity of those particular towers is uneven. These examples indicate that for pairs of individuals, being connected in the social network, their trajectories may share dramatically different degrees of overlap, corresponding to different extent of face-to-face interactions. In order to capture the degree of face-to-face interactions imposed by two individuals' trajectories, we define the co-location Rate, C, as the probability for two users to appear at the same location during the same hour [20]. A co-location rate of 0 indicates two individuals move in different neighborhoods and probably never meet (Fig. 5(a)). The two trajectories shown in Fig. 5(b) have a co-location rate of 0.44, capturing the concurrent movements of the two individuals around the same neighborhood.

This raises an important question: What is the distribution of C for all the social ties across the society? In Fig. 6 we

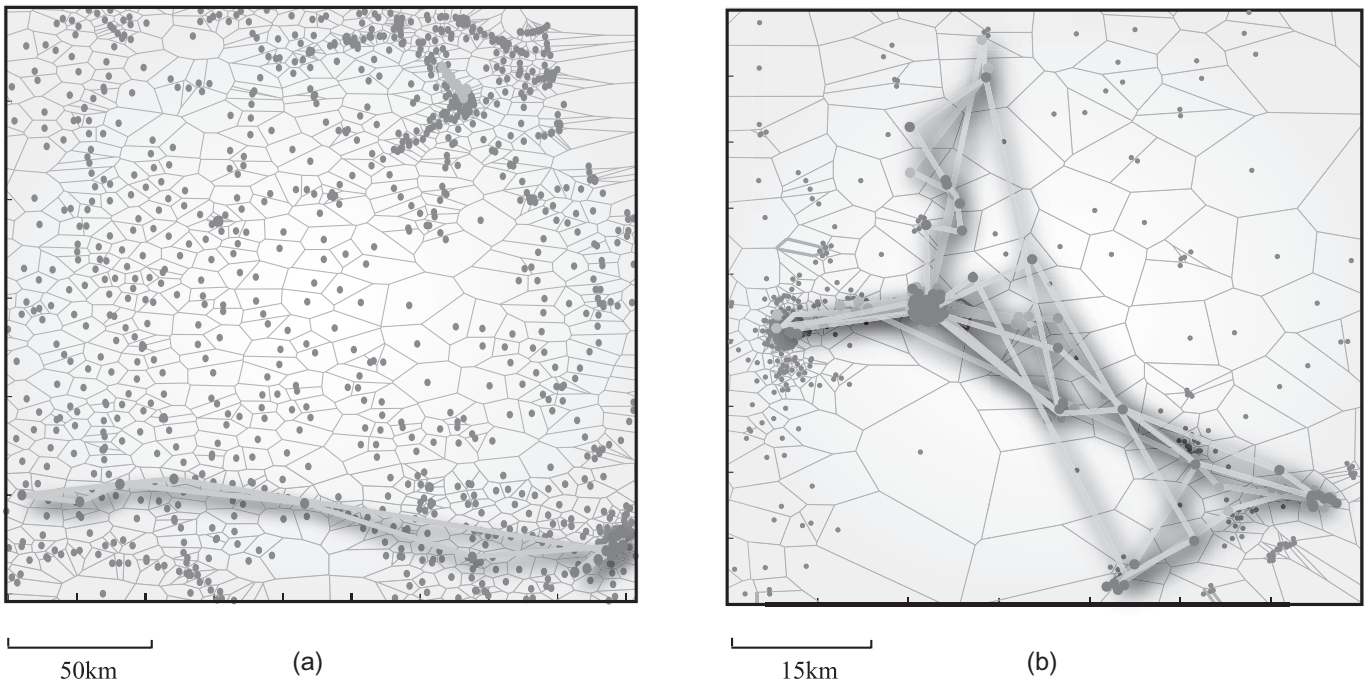

Fig. 5. Trajectories of two pairs of friends with co-location rate: (a) $C = 0$ and (b) $C \approx 0.44$ during three-months time period. The grey dots correspond to mobile phone towers and the grey lines represent the reception areas of towers. The color dots represent locations visited by the users and the lines correspond to users' trajectories.

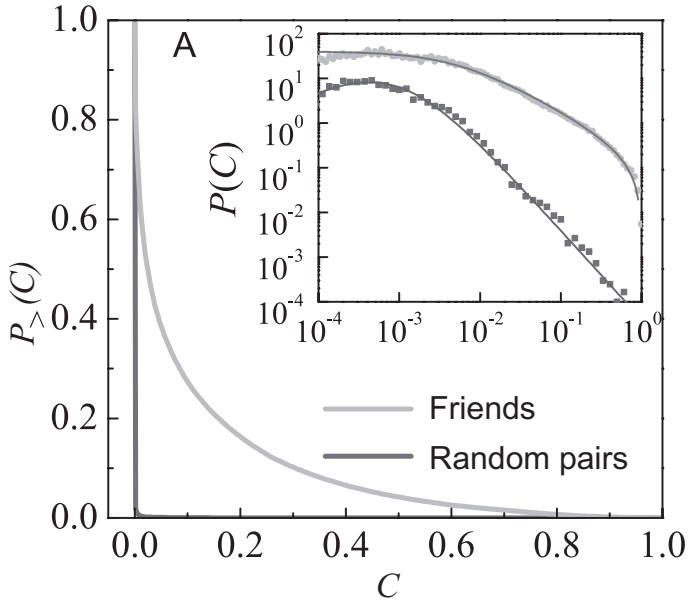

Fig. 6. The complementary cumulative distribution $P_>(C)$ of the co-location rate C for friend (blue) and random pairs (purple). Inset: The co-location rate distribution $P(C)$ for friend pairs in log-log plot.

plot the complimentary cumulative distribution, $P_>(C)$, which measures the probability of finding a social tie of co-location rate greater than C. We find, first of all, 15% of links have a co-location rate of 0, corresponding to social ties where two friends' whereabouts do not overlap. Despite a relatively small fraction of the existence of these social ties, we check our dataset to confirm that they are not artificial (e.g., customer service or salesperson). Second, for the ties that share some degree of overlap in their trajectories, such similarities in two individuals' mobility patterns are rather heterogeneous across the population. Indeed, as shown in the inset of Fig. 6, the probability distribution, $P(C)$, can be well approximated by lognormal, indicating that while most friends' paths only intersect very little, there are notable amount of pairs of individuals who move concurrently during their daily routines. Such pervasive overlap in individuals' trajectories raises an interesting question: whether the co-location phenomenon is merely the privilege of being as friends? Indeed, people often go about their lives when they are with friends, including job- and family-imposed restrictions, which may contribute to the observed co-location rates; yet they also often appear at massively public places in which they are members of large crowds, corresponding to potential co-location with a large number of strangers. To answer this question, we randomly selected 10M pairs of individuals that are not linked in the social network, i.e., they never communicate with each other via mobile phones. We then measure the co-location rate for these pairs of strangers, the distribution of which is shown as the purple curve in Fig. 6. We find the vast majority of pairs (more than 98%) have a co-location rate of 0. For those rare pairs that co-locate, the rates are mostly very small, corresponding to the rapid decay in probability.

V. SOCIAL AND SPATIAL PROXIMITY

The surprising distinctions in the degree of trajectory overlap between two individuals, characterized by pervasive overlap of friends' trajectories and barely any intersection for randomly selected two individuals, prompts us to investigate how social and spatial proximity interacts, correlations, and reinforces each other. Here we use two measures for each of both social and

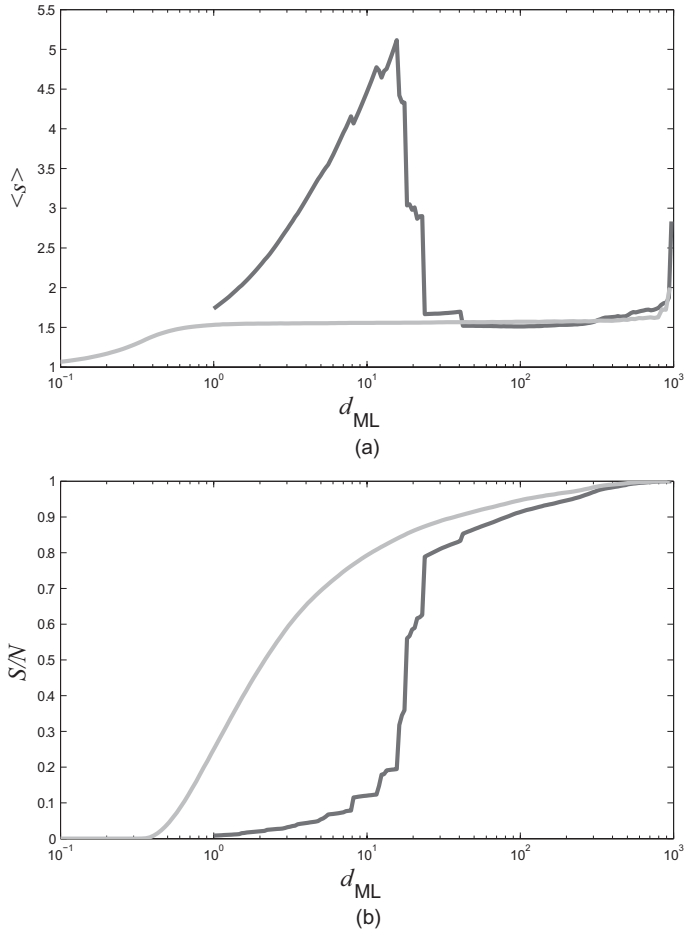

Fig. 7. The geometrical restricted networks for different cutoff distance d, showing: (a) How the average size of the isolated clusters $\langle s \rangle$ and (b) the relative GCC size S change with d (red lines), suggesting the overall social communication develops a percolation phase transition around $d \approx 10-20$ km, where each peaks in (a) associated with one jump in (b), corresponding to emergence of certain size clusters at four different stages.

spatial fronts, attending to two different levels of granularity. To quantify spatial proximity on the physical space, we not only measured the aforementioned colocation rate C to capture the dynamical aspect of trajectory overlap, but also calculated the geographic distance between the most likely locations for each pair of friends. To quantify social proximities we measure on the local level the likelihood of social ties. To further uncover the local clustered structure, for each pair of friends i and j we calculated the topological overlap $O_T(i,j) = J_n(i,j)/\min(k_i,k_j)$, where $J_n(i,j)$ denotes the number of common friends both user i and j share and $\min(k_i,k_j)$ represents the smaller degree of them [33], [34]. The topological overlap measures the strength of clustering between two friends, where $O_T = 1$ implies that they have same set of friends, whereas a zero value indicates that i and j do not share any common friends.

A. Geographic Distance Fragments the Structure of Social Network

To understand the effect of distance on social network, we adopt a measurement framework in percolation theory [35]–[37] by introducing a cutoff distance d to measure the farthest distance one can reach his friends to spread information. Varying d and removing all links with distance less than d results in a truncated communication network, where the links between two friends are allowed only if their pair distance is less than the cutoff d, representing individual passes information to his or her friends only inside the individual's vicinity with radius less than d. As individuals can share information inside the same connected components, we measure the giant connected component (GCC) to quantify the capacity of the global information sharing. In Fig. 7, we show the relative GCC size S and the average size of the isolated clusters $\langle s \rangle$ change as a function of d. When d is small, only the links between two friends who live closer to distance d are allowed. Hence the resulting social network consists of small fragments that are separated from each other, characterized by a small S, and $\langle s \rangle$ increases slowly ad d increases. When d becomes large enough, we recover the whole communication network as all social ties are preserved. As we vary the parameter d, we observe a very interesting transition occurring around $d \approx 10-20$ km, where the societal wide social communication network undergoes a notably change from localized fragments to globally connected components, suggesting a typical communication distance separating the local and global phases. Furthermore, there exists several minor transitions around this typical distance, corresponding to peaks observed in $\langle s \rangle$, indicating several different stages of network formation through which individual clusters are merged into larger connected components and eventually forming one single giant component. This result reveals inherent modular structures inside the social communication system based on distance, which is rather unexpected, as previous results suggest distance between social communications lack any characteristic scale (Fig. 4), raising an important question: Could have this result emerged due to a random process? To answer this question we randomize the spatial distance d_{ij} and break its correlation with the underlying social tie between individuals i and j. That is, we keep the social network structure intact, but assign a new distance to each social tie drawing from the same distance distribution. We find that similar to the random scale-free network the phase transition is absent [38], indicating that an underlying correlation between spatial allocations and social network structure is responsible for observed localizations (green curves in Fig. 7). It is also interesting to note that the unusual sharpness of transition points to a discontinuous 1st-order phase transition in contrast to the common second order phase transition, which aligns with recent work on the explosive percolation [38].

B. Geographic Distance Affects Local Clustering

To uncover the local clustered structure, for each pair of friends i and j we calculated the topological overlap $O_T(i,j) = J_n(i,j)/\min(k_i,k_j)$, measuring the degree to which user i and j share common friends [33], [34]. Indeed, O_T measures the strength of clustering between two friends, where $O_T = 1$ implies that they have the same set of friends, whereas a zero value indicates that i and j do not share any friends in common. This raises an interesting question: how does the topological overlap between two friends depend on how far they live from each other? In Fig. 8, we measure $O_T(d)$ as a function of distance d between two friends. We find that O_T deceases with d, im-

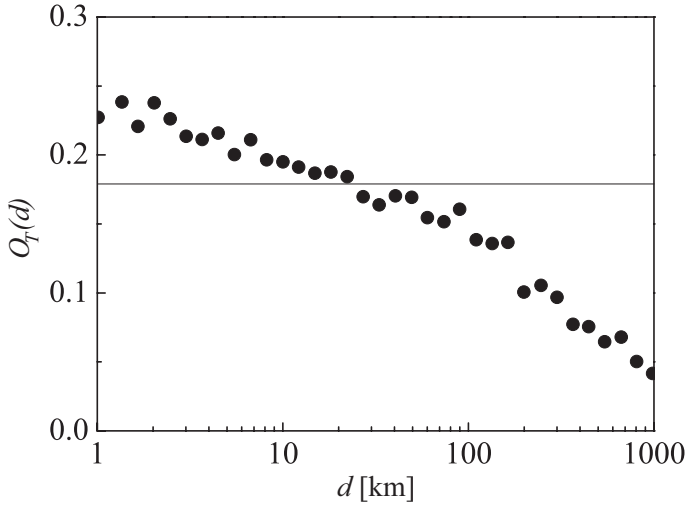

Fig. 8. The topological overlap O_T vs. pair distance d, revealing the fact that high O_T correlates with small distance region. The black bar represents the average O_T measured in uncorrelated case.

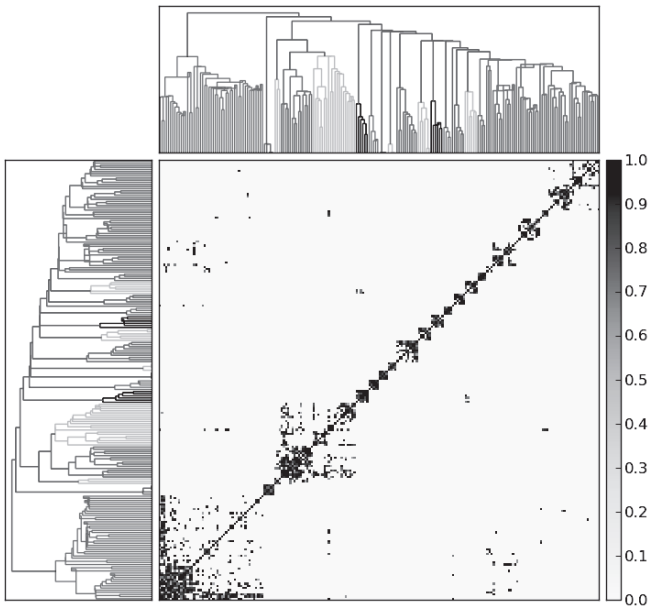

Fig. 9. The co-location matrix corresponding to a sample of users within same city obtained through a snow-ball sampling for illustration purposes. The dendrogram represents the hierarchical structure based on topological overlap on social network.

plying that the closer two users live, more common friends they share. The black line in Fig. 8 measures what we would have expected in the random case. That is, when the two quantities are uncorrelated with each other. Therefore, at the local distance region ($d < 20$ km), $O_T(d)$ is greater than the random expectation, indicating that a positive correlation between topological clustering and distance. That is, social ties that are close to each other on the physical space (small d) tend to have higher local clustering on the social space. In contrast, when $d > 20$ km, $O_T(d)$ reveals a negative correlation between topology clustering and distance. It is also interesting to note the crossover point ($d \approx 10 \sim 20$ km) is in good agreement with previous finding, further supporting the existence of a characteristic distance in wireless communication networks around $10 \sim 20$ km.

C. Trajectory Overlap Correlates with Topological Overlap

Preceding results indicate trajectory overlap highly correlates with the existence of social ties, raising the question of whether such correlation also hold for local clustering measured by topological overlap. We find the answer is yes. To illustrate this, we perform a snow-ball sampling by randomly selecting a user in a major city, sampling up to three degrees and extracting all users within the same city. In the sample studied in this section, we obtained a sample of 768 users. We computed the co-location rates (C) for any two users. This results in a 768 by 768 mobility matrix with the elements measuring trajectory overlap in their mobility patterns. We also computed the social matrix of topological overlap on the social network for any two individuals, serving as the proximity measure in social space. By applying an average-linkage hierarchical clustering algorithm [33] to the social matrix, we obtain a dendrogram, where nodes that are proximate on the social network are placed closely to each other. The rows and columns of mobility matrix are then reordered according to the order of leaves in the dendrogram. The mobility matrix after reordering alongside the social dendrogram is shown in Fig. 9. The non-zero co-location rates form blocks along the diagonal of the mobility matrix, corresponding to the modular structures of the social network. The off-diagonal elements, i.e., individuals that are far away in the social space, are characterized by mostly zero trajectory overlap. Fig. 9 illustrates strong correlations between patterns of human mobility and the structures of social network.

VI. IMPACT OF MOBILITY OVERLAP ON WIRELESS COMMUNICATION NETWORK

Being proximate is thought to encourage chance encounters and opportunities for interaction, which can lead to the formation of new relationships and the maintenance of existing ones [39]. Indeed, a number of studies have shown that the distribution of geographical distances between friends follows a power law. That is, people are far more likely to make friends with individuals who are nearby than those that are far away.

A. Static Properties

To reconcile the interplay between physical propinquity and similarity in mobility patterns, we group links based on their co-location rates, and measure $P(d|C) \sim C^{-\alpha}$, the conditional distribution of distances between friends for different values of C,. Inspired by previous results on the fat-tailed nature of the spatial distribution, one would anticipate that as well. However, despite of qualitative agreement with the expectations, Fig.10a raises a surprising classification based on mobility patterns. That is, the distributions over different C values are well separated into only two groups: the $C = 0$ group (blue circles) with $\alpha \approx 1.5$ and the $C > 0$ group (red dots) with $\alpha \approx 2.0$. Indeed, the curves collapse for groups of any co-location rates greater than zero (inset of Fig. 10(a)). That is, the distribution of distances, $P(d)$, characterized by a power law distribution with a larger exponent, is independent of the degree of overlap in two individuals'

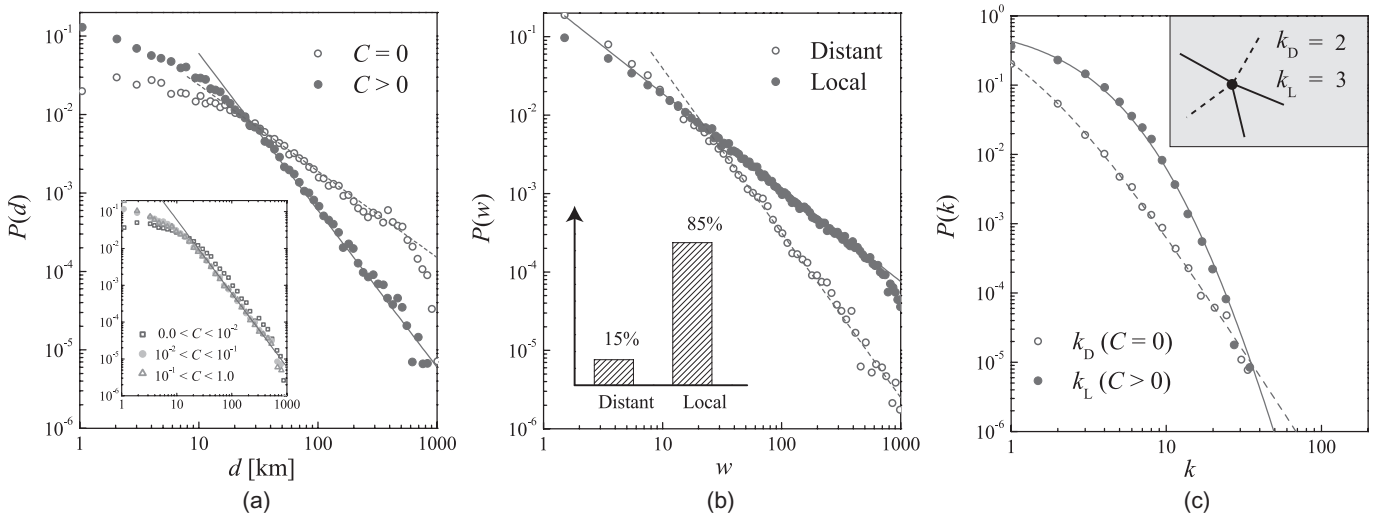

Fig. 10. (a) The distance distribution $P(d)$ measuring the probability of find a pair of friends with distance d apart from each other for two different pair groups ($C > 0$ and $C = 0$), respectively. This plot indicates that $P(d) \sim d^{-\alpha}$ decays as a power law for both cases, with different scaling exponents $\alpha_{C=0} \approx 1.5$ and $\alpha_{C>0} \approx 2.0$. Inset: $P(d)$ for different C groups for $C \in [0, 0.01)$, $[0.01, 0.1)$ and $[0.1, 1.0)$, indicating the distance exponents α are independent of the C values as far as $C > 0$. Together with the main panel, these plots demonstrate that $C = 0$ (local) and $C > 0$ (distant) are two major classes of human mobility interactions. (b) The $P(w)$ distribution of the strength of tie w for local and distant pairs respectively, showing $P(w) \sim w^{-\beta}$ with different exponents $\beta_L \approx 1.2$ and $\beta_D \approx 2.2$. (c) The degree distribution $P(k)$ for local and distant friends, showing again different power law exponents 7.0 and 3.5, respectively. Both lines are guides to the eye with power law decays. Inset: An illustration demonstrates an individual with $k_L = 3$ and $k_D = 2$, where the solid and dashed lines represent local and distant ties, respectively.

trajectories as long as their trajectories intersect. Even the group of social ties with very small C, i.e., friends do co-locate with each other yet very rarely, differentiates remarkably from the group of friends that do not share any foci along their paths (see the inset of Fig. 10(a)). These significant deviations offer strong evidence that the social ties can be in fact categorized into two types: friends, whose trajectories overlap ($C > 0$), and those with $C = 0$. We therefore refer the first type as local ties, and the latter type as distant ties hereafter.

The intriguing discovery of two types of social ties raises an interesting question: how will such categorization of social ties based on mobility patterns contribute to our current understanding of social networks? When it comes to structural properties of links, the link weight distribution is probably the most fundamental one. We therefore plot the link weight distribution, $P(w)$, separately for distant ties and local ties. As we show in Fig. 10(b), the tie strength distribution of two groups both follow a power law distribution yet with radical different exponents. Indeed, the strength of the local ties exhibits more heterogeneity than that of the distant ties. This not only corroborates that these two types of social ties are fundamentally different, but also indicates that the local ties are more closely connected than the distant ties, further proving that spatial co-location induces proximity on the social space.

Degree is another fundamental quantity in networks. Indeed, every node is connected with its neighbors through either local (L-) or distant (D-) ties, depending on its co-location rate with them. To inspect the impact of these two types of ties on the degree distribution, we break down the degree of each individual into his/her local and distant degrees, k_L and k_D, which are the total number of local and distant ties s/he has, respectively. Note, the summation of local and distant degrees equals to the degree of node, i.e., $k = k_L + k_D$. If there is no correlation between these two types of social ties, the distributions of local and distant degree should follow

$$P(k_L) = \sum_{k \le k_L} \binom{k}{k_L} p_L^{k_L} (1 - p_L)^{k - k_L} P(k), \qquad (4a)$$

$$P(k_D) = \sum_{k \le k_D} \binom{k}{k_D} P_L^{k - k_D} (1 - p_L)^{k_D} P(k) \qquad (4b)$$

where $P(k)$ is the overall degree distribution and $p_L = 85\%$ is the ratio of local ties out of all social ties in the network. Together with the fat-tailed degree distribution (1), Eq. (4) above predicts that both $P(k_L)$ and $P(k_D)$ will be fat-tailed with exponents $\gamma_L = \gamma_D = \gamma$, independent of p_L. In Fig. 10c, we test this prediction by empirically measuring the distributions of local and distant degrees. Different from what was predicted for the uncorrelated case, the distributions of these two types of degrees are associated with evidently different exponents ($\gamma_D \approx 3.5$ and $\gamma_L \approx 7$).

B. Temporal Dynamics

The fact that γ_D is much smaller than γ_L indicates that the distant ties, rather than distributed randomly, tend to cluster together, favoring individuals who have larger k_D and therefore resulting in a more heterogeneous distribution $P(k_D)$. In this section, we further exam how the the underlying correlation and clustering between similar types of ties evolves dynamically. Indeed, clustering is crucial for describing the propensity that two individuals who share common friends often communicate with each other, characterized by the total number of 3-cliques (or triangles). We therefore classify all the topological triangles based on the distinction of $L-$ and $D-$ties. There are overall four types of triangles: LLL, DDD, LLD and LDD, as illustrated

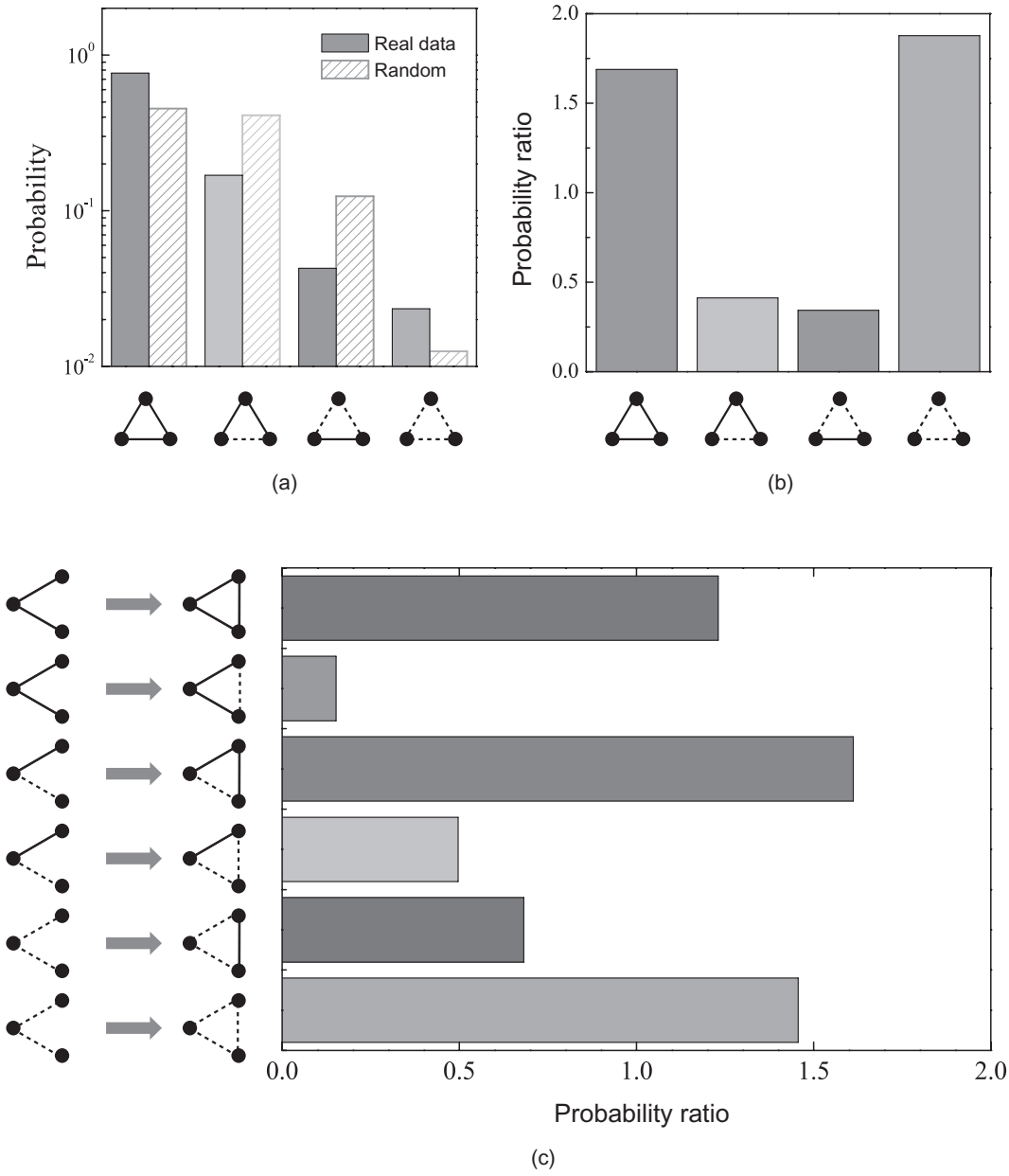

Fig. 11. (a) The probability of finding four different triangles (from left to right: LLL, LLD, LDD and DDD) for real-world dataset (solid bars) and randomized cases (striped bars). (b) The probability ratio $R = P/P_{\rm rand}$ for four triangles observed in (a), quantifying the deviation from the randomized case, indicating that the probability of finding triangles with similar ties is notably greater than random case. (c) The transition probability ratio for six different triadic closures (from top to bottom: $LL \to LLL$, $LL \to LLD$, $LD \to LLD$, $LD \to LDD$, $DD \to LD$ and $DD \to DDD$), in respective of random case, showing that $LL \to LLL$, $LD \to LLD$ and $DD \to DDD$ are three major pathways for triangle creations.

in the bottom of Fig. 11(a). If $L-$ and $D-$ties were distributed uniformly, the probability of finding a certain type of triangles should follow a binomial distribution, indicated by striped bars in Fig. 11(a). However, as Fig. 11(a) shows, the probability of finding different types of triangles all deviates dramatically from predictions (with p-values less than 10^{-10}), indicating remarkable underlying correlations between different types of ties. To explore these correlations, in Fig. 11(b), we measure the probability ratio, $R = P/P_{\rm rand}$, for these four types of triangles, quantifying the deviation from the randomized case. A probability ratio $R = 1$ means the number of triangles in this type is just about what you would expect randomly. $R > 1$ indicates a more favorable and populated type, and vice versa. Fig. 5(b) demonstrates that triangles are more probable to be formed by ties of the same type, i.e., individuals are mainly grouped together by either all LLL or DDD communications. Instead, the triangles formed by complementary ties (LLD or LDD) are underrepresented compared with the random case. These results raise an important question: How do these different clustered groups emerge? That is, how do these two types of ties, induced from mobility patterns, affect triadic closure in the social network? To answer this, we need to look at the creation of triangles under temporal evolution. We first construct networks, G_1 and G_2, from two consecutive months, respectively. We then filtered triangles where a new link in G_2 attaches to an open triplet in G_1 and closes a triangle. (An open triplet is a sub-graph of

three nodes connected by two links). As illustrated in Fig. 11(c), there are six possible types of such closures of triads, namely, $LL \rightarrow LLL$, $LL \rightarrow LLD$, $LD \rightarrow LLD$, $LD \rightarrow LDD$, $DD \rightarrow LDD$, $DD \rightarrow DDD$. The probability ratio of forming each type of triangles is shown in Fig. 11(c). Interestingly, we find that $L-$tie is more likely to close a triangle when there pre-exists at least one $L-$tie in the open triplet. This indicates, although it is widely believed that physical co-presences and social foci are of fundamental importance in forming social ties, to the extent co-location affects triadic closure, such hypothesis fails when the other two ties in the triangle are distant ties, corresponding to the opposing forces between face-to-face interaction and homophily effect.

VII. CONCLUSION

In summary, our knowledge of how individual mobility patterns impact the social network is essential for a deeper understanding of the wireless network structure.In this paper, we analyzed the mobile phone network includes social network topology, spatial distribution and individual's movements, exploring the underlying correlations between these features.

We find that the social network is spatially clustered, forming a hierarchical structure based on the spatial separations between individuals. The strong correlation between social ties and their spatial distribution results in a geographical localization within a vicinity with size around 10–20 km. We further find that the co-location between individuals' movements significantly contributes to the clustering phenomena in social network as well as triadic closure. Most important, the social ties can be classified into two distinct categories based on the co-location rate between two individuals. When two samples, both following a fat-tailed distribution with yet different exponents, are mixed and measured together, the smaller exponent dominates the overall distribution . Therefore, if one takes a mean-field approach, i.e., without taking into account the distinction between local and distant ties, to measure the distributions of network quantities, from link weights and degrees to distances between friends, one type of ties will inevitably be masked from the other type, corresponding to overestimating the overall distributions.

As human mobility and social networks is largely relevant to a wide array of applications, these results based on their mobility correlations are expected to revolutionize our strategies when leveraging our current understanding of both human behavior and social networks, helping improve existing processes and create new applications including better understandings of information spreading, improvement of cybersecurity and wireless capability, and better designs of routing protocols, etc. For instance, our findings of social tie localization suggests more realistic spatial wireless network models. It also allows to make a better prediction of the network dynamics [20], potentially leading to efficient routing protocols. Furthermore, the social network graph is a key determinant of social influence and how fast can a worm propagate between social neighbors of a node, which contributes to its vulnerability. Humans interact with the social dimension at many levels such as by making mobility decisions and by determining how information is shared on the social network. Therefore, the observed two classifications of social ties based on human mobility and their dynamical triadic closure will lead to new information diffusion models across wireless complex networks on the geographical space, potentially help our understanding of malware propagation and information spreading.

REFERENCES

[1] A.-L. Barabasi, "Linked: How everything is connected to everything else and what it means," *Plume Editors*, 2002.

[2] R. Albert and A.-L. Barabási, "Statistical mechanics of complex networks," *Reviews of modern physics*, vol. 74, no. 1, p. 47, 2002.

[3] M. E. Newman, "The structure and function of complex networks," *SIAM review*, vol. 45, no. 2, pp. 167–256, 2003.

[4] D. J. Watts, *Six degrees: The science of a connected age*. WW Norton & Company, 2004.

[5] G. Caldarelli, "Scale-free networks: Complex webs in nature and technology," *OUP Catalogue*, 2007.

[6] M. Newman, *Networks: an introduction*. Oxford University Press, 2010.

[7] D. Easley and J. Kleinberg, *Networks, crowds, and markets: Reasoning about a highly connected world*. Cambridge University Press, 2010.

[8] E. Cho, S. A. Myers, and J. Leskovec, "Friendship and mobility: user movement in location-based social networks," in *Proc. ACM SIGKDD*, 2011, pp. 1082–1090.

[9] S. Scellato, A. Noulas, and C. Mascolo, "Exploiting place features in link prediction on location-based social networks," in *Proc. ACM SIGKDD*, 2011, pp. 1046–1054.

[10] N. Eagle, A. Pentland, and D. Lazer, "Inferring friendship network structure by using mobile phone data," in *Proc. National Academy of Sciences*, vol. 106, no. 36, 2009, p. 15274.

[11] M. González, C. Hidalgo, and A. Barabási, "Understanding individual human mobility patterns," *Nature*, vol. 453, no. 7196, pp. 779–782, 2008.

[12] N. D. Lane *et al.*, "A survey of mobile phone sensing," *IEEE Commun. Mag.*, vol. 48, no. 9, pp. 140–150, 2010.

[13] D. Lazer *et al.*, "Computational social science," *Science*, vol. 323, no. 5915, pp. 721–723, 2009.

[14] A. Halevy, P. Norvig, and F. Pereira, "The unreasonable effectiveness of data," *IEEE Intell. Syst.*, vol. 24, no. 2, pp. 8–12, 2009.

[15] C. Ratti *et al.*, "Mobile landscapes: using location data from cell phones for urban analysis," *Environment and Planning b Planning and Design*, vol. 33, no. 5, p. 727, 2006.

[16] M. C. González and A.-L. Barabási, "Complex networks: From data to models," *Nature Physics*, vol. 3, no. 4, pp. 224–225, 2007.

[17] J. Candia *et al.*, "Uncovering individual and collective human dynamics from mobile phone records," *J. Physics A: Mathematical and Theoretical*, vol. 41, no. 22, p. 224015, 2008.

[18] P. Wang *et al.*, "Understanding the spreading patterns of mobile phone viruses," *Science*, vol. 324, no. 5930, pp. 1071–1076, 2009.

[19] C. Song *et al.*, "Limits of predictability in human mobility," *Science*, vol. 327, no. 5968, pp. 1018–1021, 2010.

[20] D. Wang *et al.*, "Human mobility, social ties, and link prediction," in *Proc. ACM SIGKDD*, 2011, pp. 1100–1108.

[21] J. P. Bagrow, D. Wang, and A.-L. Barabasi, "Collective response of human populations to large-scale emergencies," *PloS one*, vol. 6, no. 3, p. e17680, 2011.

[22] J. P. Bagrow and Y.-R. Lin, "Mesoscopic structure and social aspects of human mobility," *PloS one*, vol. 7, no. 5, p. e37676, 2012.

[23] F. Simini *et al.*, "A universal model for mobility and migration patterns," *Nature*, vol. 484, no. 7392, pp. 96–100, 2012.

[24] L. Gao *et al.*, "Quantifying information flow during emergencies," *Scientific reports*, vol. 4, 2014.

[25] F. Calabrese *et al.*, "Interplay between telecommunications and face-to-face interactions: A study using mobile phone data," *PloS one*, vol. 6, no. 7, p. e20814, 2011.

[26] J.-P. Onnela *et al.*, "Geographic constraints on social network groups," *PLoS one*, vol. 6, no. 4, p. e16939, 2011.

[27] L. Adamic and E. Adar, "How to search a social network," *Social Networks*, vol. 27, no. 3, pp. 187–203, 2005.

[28] D. Liben-Nowell *et al.*, "Geographic routing in social networks," in *Proc. National Academy of Sciences of the United States of America*, vol. 102, no. 33, 2005, pp. 11623–11628.

[29] R. Lambiotte *et al.*, "Geographical dispersal of mobile communication networks," *Physica A: Statistical Mechanics and its Applications*, vol. 387, no. 21, pp. 5317–5325, 2008.

[30] G. Krings *et al.*, "Urban gravity: a model for inter-city telecommunication flows," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2009, no. 07, p. L07003, 2009.
[31] D. J. Crandall *et al.*, "Inferring social ties from geographic coincidences," in *Proc. National Academy of Sciences*, vol. 107, no. 52, 2010, pp. 22436–22441.
[32] R. Cohen and S. Havlin, *Complex networks: structure, robustness and function*. Cambridge Univ. Press, 2010.
[33] E. Ravasz *et al.*, "Hierarchical organization of modularity in metabolic networks," *science*, vol. 297, no. 5586, pp. 1551–1555, 2002.
[34] J. Onnela *et al.*, "Structure and tie strengths in mobile communication networks," in *Proc. National Academy of Sciences*, vol. 104, no. 18, 2007, p. 7332.
[35] R. Albert, H. Jeong, and A.-L. Barabási, "Error and attack tolerance of complex networks," *Nature*, vol. 406, no. 6794, pp. 378–382, 2000.
[36] R. Cohen *et al.*, "Resilience of the internet to random breakdowns," *Physical review lett.*, vol. 85, no. 21, p. 4626, 2000.
[37] D. S. Callaway *et al.*, "Network robustness and fragility: Percolation on random graphs," *Physical review letters*, vol. 85, no. 25, p. 5468, 2000.
[38] D. Achlioptas, R. M. D'Souza, and J. Spencer, "Explosive percolation in random networks," *Science*, vol. 323, no. 5920, pp. 1453–1455, 2009.
[39] M. T. Rivera, S. B. Soderstrom, and B. Uzzi, "Dynamics of dyads in social networks: Assortative, relational, and proximity mechanisms," *annual Review of Sociology*, vol. 36, pp. 91–115, 2010.

Dashun Wang is Assistant Professor of Information Sciences and Technology at the Pennsylvania State University and Adjunct Assistant Professor of Physics at Northeastern University. Prior to joining Penn State, he was a Research Staff Member at the IBM T.J. Watson Research Center. He received his Ph.D. in Physics from Northeastern University in 2013, where he was a Member of the Center for Complex Network Research. From 2009 to 2013, he had also held an affiliation with Dana-Farber Cancer Institute, Harvard University as a Research Associate. Dashun received his B.S. degree in Physics from Fudan University in 2007.

Chaoming Song is an Assistant Professor in Physics at University of Miami. He received the B.S. degree from the Fudan University in 2001 and the Ph.D. degree in Physics from the City University of New York (CUNY) in 2008. Between 2008 and 2013, he worked at the Northeastern University and Harvard Medical School as a Postdoctoral Fellow. Since August 2013 he joined in the Physics Department of University of Miami. Song's research interests include statistical physics, network science, and computational social science.