Social Information Improves Location Prediction in the Wild

Jia Li,[1*] **Ivan Brugere,**[1*] **Brian Ziebart,**[1]
Tanya Berger-Wolf,[1] **Margaret Crofoot,**[2] **Damien Farine**[2]
1 Department of Computer Science, University of Illinois at Chicago, Chicago, IL
2 Department of Anthropology, University of California, Davis, Davis, CA
{jli213, ibruge2, bziebart, tanyabw}@uic.edu; {mccrofoot, drfarine}@ucdavis.edu

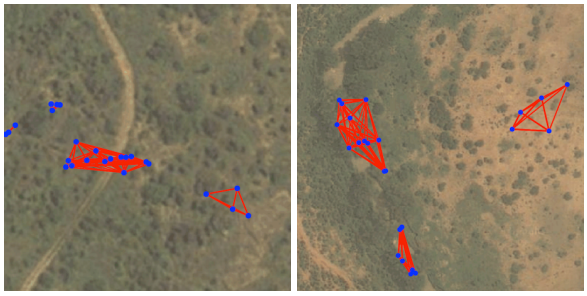
Abstract

How can knowing the location of my friends be used to more accurately predict my location? This paper explores socially-aware location prediction under a particularly challenging setting where the underlying interactions and social network are unknown and must be inferred over continuous spatiotemporal data. Our method samples inferred network topology using a linear regression model to predict future individual locations. We present an in-depth empirical study comparing different network models and network sampling regimes under a bootstrapped sampling baseline. Furthermore, our qualitative analysis demonstrates the value of social information in population mobility modeling under our application's challenges.

Introduction

When modeling movement of social individuals, does the knowledge of social information improve the accuracy of the model? Research in movement modeling using mobile social sensors (GPS, mobile phone) has become a burgeoning area of research with applications in multi-user coordination (Ashbrook and Starner 2003; Davidoff et al. 2011), personalized place recommendation (Zhang and Chow 2013; Bao, Zheng, and Mokbel 2012), location-aware information ranking (Chow, Bao, and Mokbel 2010) and place discovery (Do and Gatica-Perez 2014). However, much of this work has focused on individual dynamics (González, Hidalgo, and Barabási 2008), or incorporating personal points of interest (e.g. home, work) (Cho, Myers, and Leskovec 2011) but rarely a social context. Alternatively, research has focused on aggregate estimation of large urban population dynamics (Hasan, Zhan, and Ukkusuri 2013) for applications such as destination prediction (Xue et al. 2013), improved driving directions (Yuan et al. 2010), and large-scale spatiotemporal event detection (Zheng et al. 2014), with scarce individual-level knowledge. Past work scarcely incorporates detailed social information in individual mobility modeling. Our paper investigates the question: "can knowing the location of my friends be used to more accurately predict my location?"

(a) A troop of baboons in the wild wearing GPS collars



(b) Example GPS data from baboons (blue nodes) and inferred co-location network structure (red edges). Note that interacting communities change between two time samples (left, and right). (see: Data Description).

Figure 1: Baboon GPS data overview

We develop methods for answering this question and establish an initial positive result in this paper.

The capability to infer location information from that of "friends" is important for the study of movement and behavior of social animals (including humans), and has important privacy implications. An increased understanding of the social role of individual movement allows anticipatory location-based services, and a more detailed user context. However, this capability may also yield unwanted surveillance of the users of these services, as well as their social connections who may not even be users of the service. For example, even if a user forgoes a mobile device entirely due to concerns of surveillance, an attacker with the knowledge of some small number of her friends may be able to predict her location with access to the friends' location-based social network check-ins (often published on other social networks), by having privileged access to their mobile call-

[*]Both authors contributed equally.

record data, or other geo-location data.

This work focuses on a location prediction problem on a tracked population of baboons wearing GPS collars. This instrumentation allows researchers to observe the traces and interactions of animals in unprecedented detail (Krause et al. 2013; Morales et al. 2010). Baboons are highly social animals and these devices provide high-resolution GPS data that would not generally be appropriate to collect on human subjects due to privacy concerns, while still providing a 'surrogate' for several human movement dynamics. Such data provides a testbed for modeling the movement of any social animal (including humans), and can be down-sampled to simulate many different sensor environments.

Figure 1 (a) shows the troop in the wild, wearing these devices (see: Data Description). Figure 1 (b) visualizes the GPS data against landscape imaging. Between individuals (blue nodes), we infer interaction networks (red edges) based on several network inference models (see: Methods). In the left and right images we see a different number of interacting subgroups of varying densities. Analogous to urban spaces, this population is characterized by intentional, periodic interactions under differing (hidden) contexts, as well as a mixing of individuals in chance encounters. This inferred time-evolving 'interaction network' can be used as a representation of social information. Our prediction task directly tests the predictability of individuals simulated to be 'outside' of our dataset: given the social information of *this* individual baboon and hiding his location information, how accurately can we track this individual? We use several network models (see: Experimental Design) to demonstrate improved predictive performance over baseline sampling strategies, including population demographic-based networks and randomly constructed networks (see: Experimental Results).

Related Work

Improved localization technologies (GPS and cellular phone triangulation) have enabled significant amounts of research for reasoning about human-generated location trajectories from individuals and groups. Our work extends upon previous work demonstrating that interactions and social relations are structured by geography, and that social-spatiotemporal modeling can be predictive of behavior or future interactions.

Individual mobility modeling

This area of work focuses on developing descriptive models of human movement (González, Hidalgo, and Barabási 2008) or predictive models demonstrating periodic and other regular aspects of individual location (Cho, Myers, and Leskovec 2011). Many of these models incorporate points of interest (e.g. home and work) or operate on constrained spaces such as urban road networks. Early work focuses on extracting places of interest from individuals daily routines and predicting which place of interest the individual would visit next (Marmasse and Schmandt 2000; Ashbrook and Starner 2003). Extensions also infer an individual's mode of transportation at each point of a location trajectory (Patterson et al. 2004) and the motivations for selecting specific trajectories (Ziebart et al. 2008).

Geographic constraints on social networks and applications

This area of research examines the aggregate relationships between social network topology and geography. (Onnela et al. 2011) model social network communities derived from call record data using the standard Louvain (Blondel et al. 2008) method of community detection. (Backstrom, Sun, and Marlow 2010) examine social-spatial statistics across the Facebook network for location prediction, and uses friend locations to infer the location of a user, assuming the spatial autocorrelation of friendship to make this prediction. Several studies use Location Based Social Networks (LBSNs), analyzing both aggregate individual-place (Noulas et al. 2011), and social-place (Scellato et al. 2011) statistics. The main difference in our setting is that we predict future locations *and* infer the social network topology.

Geo-spatial network prediction

This work is most related to spatial-social prediction problems. Methods for link prediction incorporating spatial features such as co-location have been developed and evaluated on mobile call record data (Wang et al. 2011), and location-based social networks (Scellato, Noulas, and Mascolo 2011). Other work on LBSNs has incorporated social and spatial influence for place recommendation (Ye et al. 2011). A key difference in our setting is that these studies have discrete interaction events (calls, check-ins) while we must find a predictive definition of interactions—and therefore, network—from continuous spatial positions rather than discrete points (e.g. places, cell towers). (Noulas et al. 2012) propose a method with spatiotemporal features for next-place prediction, also on Location Based Social Networks, but this does not incorporate social features. Previous work has also inferred networks from spatiotemporal co-occurrence of mobile phone data (Eagle, Pentland, and Lazer 2009), and multimedia events in Flickr (Crandall et al. 2010). However, both studies use a "ground-truth" social network for evaluation, where none exists in our setting.

Data Description

Our dataset is collected from high-resolution GPS collars worn by a group of 26 baboons in the wild at Mpala Research Centre, Kenya (Figure 1). These devices capture geolocation of each baboon at one observation per second. The spatial error of this data is low enough to identify several types of interactions including coordinated movement, subgroupings and communities, and group splits over time. Data is collected for 12 hours a day ($\approx 43,000$ observations a day) for 35 days for a maximum of $1.5M$ latitude-longitude coordinates per individual. Due to battery and transmission failures in collars over the duration of the study, we examine a subset of 20 individuals with complete data over a four day period.

Methods

We aim to demonstrate that linear regression with social features is an effective starting point in location prediction using social information. In this section we give an overview of our regression model, and the network models used to infer social interactions between individuals.

Linear Regression

Linear regression is a well-known and widely used prediction model. It predicts a vector of real-valued outputs y (n-length) given an input dataset X composed of n examples (e.g. individuals), each with p features (an $n \times p$ sized matrix). $\vec{w}$ is a p-length learned weight vector for X, and w_0 is a constant offset term. Linear regression is equivalent to minimizing the residual sum of squares (Equation 2):

$$y = X\vec{w} + w_0. \tag{1}$$

$$\min_w \frac{1}{2n} \sum_{i=1}^{n} (y_i - w_0 - \vec{x}_i^{\mathrm{T}} \vec{w})^2 \tag{2}$$

This weight vector is then used to make predictions on test data:

$$\hat{y} = X_{test}\vec{w} + w_0. \tag{3}$$

In our experiments, the input X contains $\langle Latitude, Longitude \rangle$ locations of a set of *predictor* individuals, excluding the predicted individual. The output y is the $\langle Latitude, Longitude \rangle$ location of our predicted individual. In the next subsection, we describe how the set of individuals is selected. We also add the normalization constraint that $\mathbf{1}^T\vec{w} = 1$ to make the predicted individual's location a combination of other individuals' locations.

Network construction models

We choose predictor individuals in our above regression from an inferred social network between individuals. We assume that frequent co-location approximates *undirected interactions*, although we can incorporate positive, negative, and asymmetric relationships for location prediction. We define a graph $G = \langle V, E \rangle$ and construct unweighted edges $\langle n_i, n_j \rangle \rightarrow \{0, 1\}, i \neq j, n_i \in V \wedge n_j \in V$ between pairs of nodes (i.e. individuals) based on two network inference procedures:

Linear threshold kernel (PD) For each time step, this model calculates the pairwise Euclidean distance between nodes n_i and n_j, $i \neq j$. A linear threshold kernel constructs an edge if and only if a distance threshold is met: $\langle n_i, n_j \rangle \in E \Leftrightarrow d(n_i, n_j) \leq \gamma$. We apply TWIN (Sulo, Berger-Wolf, and Grossman 2010) to discover the distance threshold $\gamma = 10$ meters, which provides maximum stability over time (optimizing network measure variance vs. compression ratio). We then construct a time-evolving network (at 1Hz) under threshold γ. For each time window w ($|w| = 3600 =$ one hour), we apply a second "persistence" threshold λ, which we later vary for sensitivity analysis. We aggregate the time-varying graph defined on this time window into a static network for prediction. To do this, for each edge we count the number of time steps the edge exists on the time window, and build a distribution over all $\binom{|V|}{2}$ edge counts. An edge exists in the aggregated E_w edge set if and only if its count is above the λ-th percentile in the frequency distribution. Note that this threshold is adaptive to the edge occurrences over w. $|(\langle n_i, n_j \rangle)_w| \geq \lambda_{k,w}$ denotes the occurrences of $\langle n_i, n_j \rangle$ greater than the kth percentile measured on the window w. For simplicity, we use λ_k to threshold each w according to k; We do our analyses on: $\{\lambda_{90}, \lambda_{80}, \lambda_{70}, \lambda_{60}\}$. For clarity, this set of thresholds loosens from left to right.

Gaussian Mixture Model (GMM) This method is an extension to 2D geographic space from previous work for inferring time-varying graphs from time series event data over a set of discrete points (i.e. disjoint sensor sites) (Psorakis et al. 2012). This method is parameter-free, and learns a minimal set of $(\mu, \sigma)_s$, a variable-number of Gaussian distributions which best cluster the data. For each time step on the window w, this method assigns a cluster label to each node. We then construct a clique among all nodes with the same label (i.e. at each time step, this graph contains exactly s connected components). We follow the same procedure as (PD) to construct hourly networks on all windows using the percentile threshold set $\{\lambda_{90}, \lambda_{80}, \lambda_{70}, \lambda_{60}\}$.

Experimental Design

In order to test our initial question, "can knowing the location of my friends be used to accurately predict my location?", we first formulate two precise definitions of "friends" on our networks. We then formulate a baseline null model using bootstrapped samples of nodes. Bootstrapping (Efron and Tibshirani 1993) is a common statistical technique informally equivalent to "random sampling with replacement." In this case, the baseline model measures the predictability of the population in aggregate, *without* any detailed social information. Our sampling strategies then directly measure the predictive power of our inferred networks.

Network sampling

We construct a regression model for each individual $n_i \in V$, based on the locations of a subset of other individuals, $X = \langle lat_{w,C}, lon_{w,C} \rangle, C \subseteq V \setminus n_i$. We define the subset C according to three sampling strategies:

1. *Neighborhood*, $C = N(n_i) := \{v | \langle v, n_i \rangle\} \in E$, the nodes v adjacent to n_i,
2. *Community*, $C = \{L(v) = L(n_i), v \neq n_i | Louvain(G)\}$, nodes sharing the same label output assignment from the Louvain community detection method (Blondel et al. 2008).
3. *Null*, $C_{rand} = \{Rand(V \setminus n_i)\}_k$ is defined relative to the *Neighborhood* or *Community* set C, where C_{rand} contains $k = |C|$ random nodes, sampled without replacement.

We will now describe the procedure to build the bootstrapped sample and calculate the percentile error statistic. For each individual, for each training/test window pair

$\langle w_i, w_j \rangle, i \neq j$ we build a distribution of bootstrapped Euclidean distance errors by sampling C_{rand} from V, then running the regression model with location data of individuals in C_{rand} as input, and reporting the mean error over all predictions on w_j. C_{rand} is sampled with replacement for $K(= 150)$ trials. Let $\{e(C_{rand}, w_i, w_j)\}_K$ denote this null distribution.

We calculate location prediction error of each network sampling method in a similar procedure as above. For each individual, for each training/test window $\langle w_i, w_j \rangle, i \neq j$ we sample C according to the network sampling method, using location data of all individuals in C as input to the regression, and report the mean error over all predictions on w_j, denoted $e(C, w_i, w_j)$. Figure 2 illustrates querying the percentile of the network sampling error $e(C, w_i, w_j)$ (red) within the $\{e(C_{rand}, w_i, w_j)\}_K$ (blue) null distribution. A well-performing predictor falls within a low percentile of the null error distribution (i.e. the area within the distribution *to the left* of the red line is small). Specifically, in Figure 2, $e(C, w_i, w_j)$ performs better than random, in the 26th percentile of the null distribution. This percentile error statistic is adaptive to arbitrary null distribution models. Because it is a relative measure, it is adaptive to the relative hardness of prediction without social information. To abstract one step further, we can summarize this error percentile statistic as $E(i, \langle w_i, w_j \rangle, S)$ for individual i, on training/test windows w_i, w_j, and where S is a 3-tuple encoding the network setting: (e.g.: $S = \langle$ PD, Neighborhood, $\lambda_{60} \rangle$). These error percentile values serve as the basic units of our entire analysis.

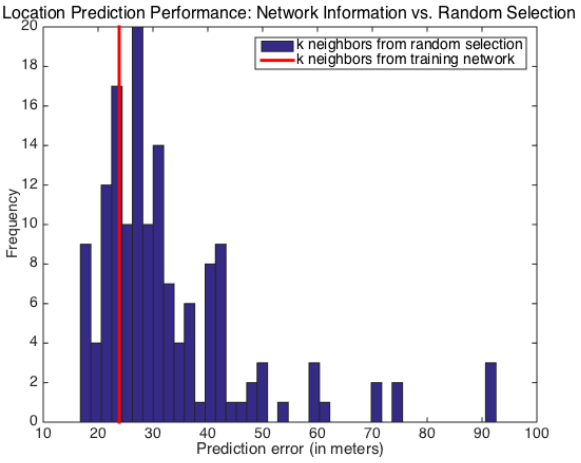


Figure 2: A comparison of the location prediction error using network *Neighborhood* information (red) vs. the null error distribution of 150 sets of randomly selected individuals of equal size to each neighborhood. The area of the distribution to the left of the red line are random trials which outperform prediction with *Neighborhood* information. A well-performing location prediction method relative to random sampling minimizes this area.

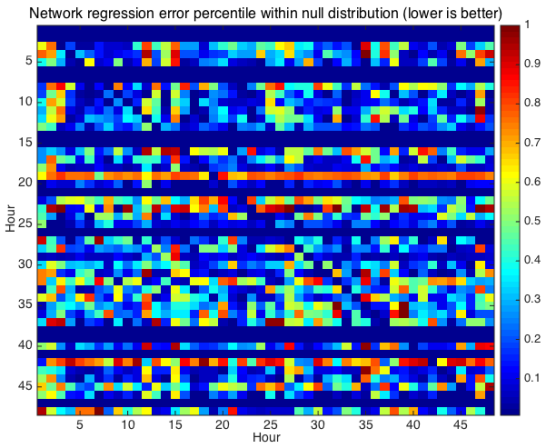


Figure 3: Summary of hours vs. hours prediction error, reporting the network regression error *percentile* relative to the null distribution (lower is better).

Experimental Results

Hour-pairs performance distribution

We now have all the tools we need to formulate the full experiment comparing *Neighborhood* and *Community* sampling against *Null*. For each of M $(= 20)$ individuals, and all $\langle w_i, w_j \rangle$ ordered training/test hour window pairs: $i \times j, i \neq j$, we compare $e(C, w_i, w_j)$ against $\{e(C_{rand}, w_i, w_j)\}_K$ if and only if $C \neq \emptyset$. Figure 3 visualizes these errors over a particular network setting: $\langle$PD, Neighborhood, $\lambda_{90} \rangle$ for a particular baboon. We call this baboon Karl (= UID 1). Since we will follow him through the rest of this experiment, we should be on familiar terms with him. Note that some rows in Figure 3 are empty. This occurs when Karl has no neighbors ($C = \emptyset$) in the inferred network over training window w_i. Also note that some rows have consistently poor performance against random (e.g. rows 19 and 42). This suggests Karl's inferred neighborhood on these hours is typically *more* distant than random. This analysis may be extended for network anomaly detection under a node label or location prediction task.

Cumulative distribution of hour-pairs

Figure 4 directly summarizes Figure 3, reporting the cumulative distribution of error percentiles over all hour-pairs for the same network settings, $\langle$PD, Neighborhood, $\lambda_{90} \rangle$. For example, for the best 10% of null distributions ($x = 10$), 23% ($y = .2324$) of hour-pair errors fall within this distribution for the network setting $\langle$PD, Neighborhood, $\lambda_{90} \rangle$. Alternatively, one-fourth of the time knowing only a few of Karl's "friends" is more predictive than even very lucky guesses. This is in a moderately large combinatorial space for 20 individuals. Figure 4 can be interpreted similar to an ROC curve, where a line above the diagonal indicates prediction performance better than the bootstrapped null distribution.

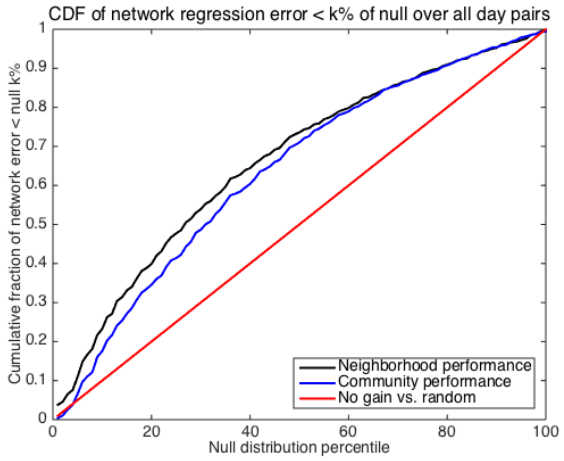


Figure 4: Cumulative distribution function comparing the distribution of all pairwise hour performance vs. a bootstrapped sample. Similar to an ROC curve, area above the diagonal is the improvement over random.

Average predictability of the whole population

We also conduct a sensitivity analysis on the parameter λ to see how the "quality" of edges—measured in frequency over the window—affects prediction under *Neighborhood* and *Community* sampling. We compare the population mean of individual median network performance vs. random. For clarity, this mean given by: $Mean_i(Median_{j\times k}(E(i, \langle w_j, w_k\rangle, S))\ i = 1...20$ (all individuals), and $j \times k = \langle 1...48 \times 1...48\rangle, j \neq k$ (all ordered hour-pairs), where functional subscripts define the indices which the function aggregates on. To measure change in performance under different λ thresholds, we subtract these medians and report the mean difference across the population under different λ (more negative is better). This can be interpreted as how much the individual hour-pairs error distribution changes on average. Table 1 summarizes a subset of changes in λ. We see that tightening the threshold monotonically increases performance on average; looking at the underlying performance per individual, there are very few that have poorer performance after any threshold tightening across all increasing combinations. Note that this mean difference is a signed measure, so unbiased noise across the population cancels to ≈ 0.

S $\rightarrow \lambda(top)$	λ_{70}	λ_{80}	λ_{90}
$\langle$ PD, Neighborhood, $\lambda_{60}\rangle$	-.014	-.021	-.051
$\langle$ PD, Community, $\lambda_{60}\rangle$	-.028	-.045	-.063
$\langle$ GMM, Neighborhood, $\lambda_{60}\rangle$	-.004	-.042	-.066
$\langle$ GMM, Community, $\lambda_{60}\rangle$	-.035	-.068	-.076

Table 1: Each cell reports the change in population mean of individual median network percentile vs. random, when moving from λ_{60} (left) to a higher threshold (top), (more negative is better).

Figure 4 raises an important question about the population: is the greater performance of *Neighborhood* relative to *Community* consistent across other individuals and network settings? Table 2 follows a similar evaluation as Table 1, reporting the change in mean error across the population, this time changing *Community* to *Neighborhood* sampling. The results show that at all λ thresholds, performance is better for *Neighborhood* sampling. This result follows our intuition because non-adjacent neighbors of the same community have strictly lower persistence than adjacent neighbors (by the persistence threshold definition λ), and we see these *Community* edges are "weaker" with respect to location prediction as well. However, Figure 4 illustrates that *Community* sampling still improves predictive performance over *Null*.

S $\rightarrow$ *Neighborhood*	λ_{60}	λ_{70}	λ_{80}	λ_{90}
$\langle PD, Community, :\rangle$	-.036	-.050	-.060	-.048
$\langle GMM, Community, :\rangle$	-.027	-.058	-.054	-.037

Table 2: Each cell in this table reports the change in population mean of individual median network percentile vs. random, when moving from a *Community* sampling (left) to *Neighborhood* sampling, reported over all thresholds (top), (more negative is better).

Comparing predictability of individuals

Figure 5 summarizes performance for each network settings for each individual, at λ_{90}. Figure 5 (a) summarizes the mean of the hour-pair performance distribution visualized in Figure 3, for each network setting (lower is better). For a pair of bars on the same network model (e.g. yellow, red for GMM), *Neighborhood* outperforms *Community* sampling across nearly all individuals. Furthermore, we see individual differences in performance both in absolute terms across all models (e.g. Individual 4 is less predictable than the population), and in relative terms where either the differences between GMM and PD or *Neighborhood* and *Community* are high. Figure 5 (b) reports the fraction of the hour-pair error distribution which falls under the bottom half of its bootstrapped sample (higher is better). For example, the blue bar for Karl (UID = 1) summarizes his CDF (Figure 4) at the 50th percentile of random ($x = 50$). The range is between 57.45% and 79.30%, again showing individual differences in predictability using "friends."

Predictability of population demographics

We now test the predictability of an explicit biological network defined on the population compared our network models. For the network *Demographics*, two nodes n_i, n_j are adjacent if and only if they share labels for both Sex and Age (on the categorical vocabulary {'adult', 'subadult', 'juvenile'}). However, any biological hypothesis testing v.s. co-location network models could be incorporated in this type of analysis. Because the *Demographics* network is a disjoint set of cliques, both *Neighborhood* and *Community* sampling are equivalent, and because it is static, all λ thresholds yield the same network. We do the same bootstrapped analysis

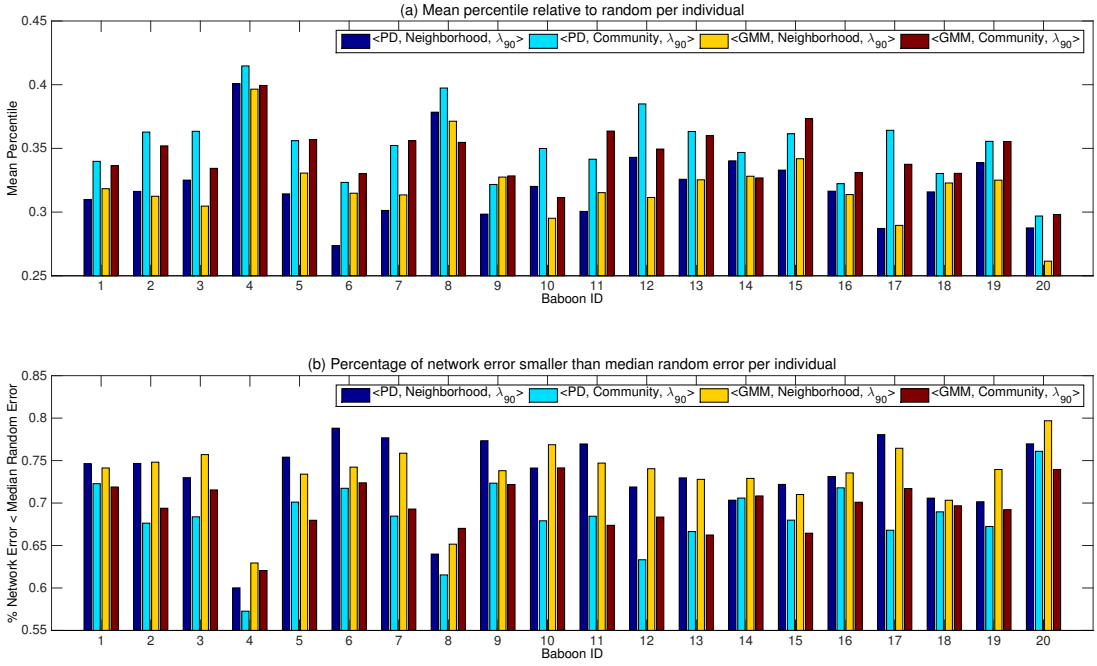


Figure 5: (a) the mean of the hour-pair performance distribution, for each network setting, for each individual (lower is better). (b) the fraction of the hour-pair performance distribution which falls under the bottom half of its bootstrapped sample (higher is better).

as above. Table 3 reports the change in population mean of individual median percentile when substituting the *Demographics* network with PD or GMM networks, on each sampling strategy constructed at varying λ. Both *Community* and *Neighborhood* sampling outperform *Demographics* suggesting that inferred social structure from repeated interaction is not merely a surrogate for this particular demographic structure. A closer look shows that of the five demographic pairs in our population, only the adult female class has a mean error better than random ($.459$). Comparing with Figure 5 (a), this subgrouping mean is worse than any of our reported individual errors for our network models, suggesting that the demographic information is much weaker relative to even the most predictive demographic. The mean error for *Demographics* across the entire population is $.511$ (random).

Demographics $\rightarrow$ S (left)	λ_{60}	λ_{70}	λ_{80}	λ_{90}
$\langle PD, Neighborhood, : \rangle$	-.142	-.171	-.189	-.206
$\langle PD, Community, : \rangle$	-.107	-.122	-.128	-.158
$\langle GMM, Neighborhood, : \rangle$	-.134	-.168	-.202	-.210
$\langle GMM, Community, : \rangle$	-.107	-.109	-.149	-.173

Table 3: Each cell reports the change in population mean of individual median network error vs. random, when moving from the *Demographics* network to the network setting S (left), at varying thresholds for S (top), (lower is better).

Conclusions and Future Work

In this paper, we have analyzed how social information can improve location prediction performance. Two important extensions are to discover rules for "when" and "where" a particular network is more predictive. In an intuitive sense, we can think of co-workers and family as predictive of location in different times and places. Empirically, in Figure 3 we want to identify predictors for well-performing hour-pairs based on network topology, location, or periodic patterns (Lahiri and Berger-Wolf 2010).

In this analysis, we construct two different network models (PD, GMM) on four different thresholds $\{\lambda_{90}, \lambda_{80}, \lambda_{70}, \lambda_{60}\}$. But, which of these is *the correct* network? These and many other networks could be constructed from location data. Choosing the most appropriate one is an open problem for statistically inferred networks where no explicit, unambiguous edge relation exists. Our future work aims to tune the network definition according to predictive performance, under the assumption that the most *predictive* network best-represents one dimension of social interaction. This is a difficult problem which balances predictive accuracy against the semantic interpretability of edge definitions. Across both network models there is a consistent trade-off between predictability and sparsity according to the quality of edges (see: Table 1). Investigating different quality criteria may help formulate a principled tuning solution.

Acknowledgements

This work was supported in part by the following grants: NSF CNS-1248080 (Berger-Wolf), NSF IGERT Grant Number 1069311 (Brugere), and NSF IOS-1250895 (Crofoot, Berger-Wolf). We thank the reviewers for their valuable comments.

References

Ashbrook, D., and Starner, T. 2003. Using gps to learn significant locations and predict movement across multiple users. *Personal and Ubiquitous Computing* 7(5):275–286.

Backstrom, L.; Sun, E.; and Marlow, C. 2010. Find me if you can: improving geographical prediction with social and spatial proximity. In *Proceedings of the 19th international conference on World wide web*, 61–70. ACM.

Bao, J.; Zheng, Y.; and Mokbel, M. F. 2012. Location-based and Preference-Aware Recommendation Using Sparse Geo-Social Networking Data. In *ACM SIGSPATIAL GIS 2012*. ACM.

Blondel, V. D.; Guillaume, J.-L.; Lambiotte, R.; and Lefebvre, E. 2008. Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment* 10.

Cho, E.; Myers, S.; and Leskovec, J. 2011. Friendship and Mobility: User Movement In Location-Based Social Networks. *Proceedings of KDD 11*.

Chow, C.; Bao, J.; and Mokbel, M. 2010. Towards location-based social networking services. In *Proceedings of the 2nd ACM SIGSPATIAL International Workshop on Location Based Social Networks*, 31–38. ACM.

Crandall, D.; Backstrom, L.; Cosley, D.; Suri, S.; Huttenlocher, D.; and Kleinberg, J. 2010. Inferring social ties from geographic coincidences. *Proceedings of the National Academy of Sciences* 107(52):22436.

Davidoff, S.; Ziebart, B. D.; Zimmerman, J.; and Dey, A. K. 2011. Learning patterns of pick-ups and drop-offs to support busy family coordination. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1175–1184. ACM.

Do, T. M. T., and Gatica-Perez, D. 2014. The Places of Our Lives: Visiting Patterns and Automatic Labeling from Longitudinal Smartphone Data. *Mobile Computing, IEEE Transactions on* 13(3):638–648.

Eagle, N.; Pentland, A.; and Lazer, D. 2009. Inferring friendship network structure by using mobile phone data. *Proceedings of the National Academy of Sciences* 106(36):15274–15278.

Efron, B., and Tibshirani, R. J. 1993. *An Introduction to the Bootstrap*. New York: Chapman & Hall.

González, M. C.; Hidalgo, C. A.; and Barabási, A.-L. 2008. Understanding individual human mobility patterns. *Nature* 453(7196):779–82.

Hasan, S.; Zhan, X.; and Ukkusuri, S. V. 2013. Understanding Urban Human Activity and Mobility Patterns Using Large-scale Location-based Data from Online Social Media. UrbComp '13, 6:1—-6:8. New York, NY, USA: ACM.

Krause, J.; Krause, S.; Arlinghaus, R.; Psorakis, I.; Roberts, S.; and Rutz, C. 2013. Reality mining of animal social systems. *Trends in Ecology & Evolution* 28(9):541–551.

Lahiri, M., and Berger-Wolf, T. 2010. Periodic subgraph mining in dynamic networks. *Knowledge and Information Systems* 24(3):467–497.

Marmasse, N., and Schmandt, C. 2000. Location-aware information delivery with commotion. In *Handheld and Ubiquitous Computing*, 157–171. Springer.

Morales, J. M.; Moorcroft, P. R.; Matthiopoulos, J.; Frair, J. L.; Kie, J. G.; Powell, R. A.; Merrill, E. H.; and Haydon, D. T. 2010. Building the bridge between animal movement and population dynamics. *Philosophical Transactions of the Royal Society B: Biological Sciences* 365(1550):2289–2301.

Noulas, A.; Scellato, S.; Mascolo, C.; and Pontil, M. 2011. An Empirical Study of Geographic User Activity Patterns in Foursquare. *Artificial Intelligence* 570–573.

Noulas, A.; Scellato, S.; Lathia, N.; and Mascolo, C. 2012. Mining User Mobility Features for Next Place Prediction in Location-Based Services. In *Data Mining (ICDM), 2012 IEEE 12th International Conference on*, 1038–1043.

Onnela, J.; Arbesman, S.; González, M.; Barabási, A.-L.; and Christakis, N. 2011. Geographic constraints on social network groups. *PloS One* 6(4).

Patterson, D. J.; Liao, L.; Gajos, K.; Collier, M.; Livic, N.; Olson, K.; Wang, S.; Fox, D.; and Kautz, H. 2004. Opportunity knocks: A system to provide cognitive assistance with transportation services. In *UbiComp 2004: Ubiquitous Computing*. Springer. 433–450.

Psorakis, I.; Roberts, S. J.; Rezek, I.; and Sheldon, B. C. 2012. Inferring social network structure in ecological systems from spatio-temporal data streams. *Journal of The Royal Society Interface* 9(76):3055–3066.

Scellato, S.; Noulas, A.; Lambiotte, R.; and Mascolo, C. 2011. Socio-spatial Properties of Online Location-based Social Networks. *Proceedings of ICWSM'11* 11:329–336.

Scellato, S.; Noulas, A.; and Mascolo, C. 2011. Exploiting Place Features in Link Prediction on Location-based Social Networks. *Proceedings of KDD 11*.

Sulo, R.; Berger-Wolf, T.; and Grossman, R. 2010. Meaningful Selection of Temporal Resolution for Dynamic Networks. In *Proceedings of the Eighth Workshop on Mining and Learning with Graphs*, MLG '10, 127–136. New York, NY, USA: ACM.

Wang, D.; Pedreschi, D.; Song, C.; Giannotti, F.; and Barabási, A.-l. 2011. Human Mobility, Social Ties, and Link Prediction. *Proceedings of KDD 11*.

Xue, A. Y.; Zhang, R.; Zheng, Y.; Xie, X.; Huang, J.; and Xu, Z. 2013. Destination Prediction by Sub-Trajectory Synthesis and Privacy Protection Against Such Prediction. In *IEEE International Conference on Data Engineering (ICDE 2013)*. IEEE.

Ye, M.; Yin, P.; Lee, W.-C.; and Lee, D.-L. 2011. Exploiting Geographical Influence for Collaborative Point-of-interest

Recommendation. SIGIR '11, 325–334. New York, NY, USA: ACM.

Yuan, J.; Zheng, Y.; Zhang, C.; Xie, W.; Xie, X.; and Huang, Y. 2010. T-Drive: Driving Directions Based on Taxi Trajectories. In *ACM SIGSPATIAL GIS 2010*. Association for Computing Machinery, Inc.

Zhang, J.-D., and Chow, C.-Y. 2013. iGSLR: Personalized Geo-social Location Recommendation: A Kernel Density Estimation Approach. SIGSPATIAL'13, 334–343. ACM.

Zheng, K.; Zheng, Y.; Yuan, N. J.; Shang, S.; and Zhou, X. 2014. Online Discovery of Gathering Patterns over Trajectories. *IEEE Transaction on Knowledge Discovery and Data Engineering*.

Ziebart, B. D.; Maas, A. L.; Bagnell, J. A.; and Dey, A. K. 2008. Maximum entropy inverse reinforcement learning. In *AAAI*, 1433–1438.