

Where and Why Users “Check In”

Yoon-Sik Cho, Greg Ver Steeg, and Aram Galstyan

USC Information Sciences Institute

Marina del Rey, CA 90292

yoonsik@isi.edu, gregv@isi.edu, galstyan@isi.edu

Abstract

The emergence of location based social network (LBSN) services makes it possible to study individuals’ mobility patterns at a fine-grained level and to see how they are impacted by social factors. In this study we analyze the check-in patterns in LBSN and observe significant temporal clustering of check-in activities. We explore how self-reinforcing behaviors, social factors, and exogenous effects contribute to this clustering and introduce a framework to distinguish these effects at the level of individual check-ins for both users and venues. Using check-in data from three major cities, we show not only that our model can improve prediction of future check-ins, but also that disentangling of different factors allows us to infer meaningful properties of different venues.

1 Introduction

Human mobility patterns influence human behavior in both routine and profound ways. Any picture of the spread of disease, traffic congestion, or urban crime would be incomplete without understanding the movements of individuals and groups. Even though the behavior of groups of humans has many degrees of freedom, previous works (Gonzalez, Hidalgo, and Barabasi 2008; Rhee et al. 2011) have demonstrated that human mobility exhibits structural regularities.

The recent emergence of Location Based Social Network (LBSN) services such as Gowalla and Foursquare has enabled researchers to perform fine-grained analysis of users’ mobility patterns and their impact on social interactions. In LBSN services, users share their current location or the venues they have visited in the past with their friends. Most LBSNs give unique IDs to different establishments even if they share the same geographical location (i.e., Lat+Long coordinates); we emphasize this distinction by using the term “venue” rather than “location”. Typically, a user “checks in” to a specific venue by using a smartphone or tablet to choose from a list of venues near their current location as determined by Wi-Fi or GPS. This information is sent to the LBSN server and shared with their friends. A user can check-in to a venue during each visit and is often encouraged to do so through incentives.

Copyright © 2014, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

The primary LBSN data consist of check-in history of the users, where each check-in is described by a user id, venue id, and the time of the check-in. In addition, most LBSN services also provide secondary data that describe the underlying social network of the users. Prior research has studied the correlation between individual mobility patterns and social interactions, e.g., by predicting social ties based on similar mobility patterns (Crandall et al. 2010), or, conversely, by predicting the next check-in location of a user based on the recent check-in history within his local network (McGee, Caverlee, and Cheng 2013).

While most prior work has focused on user-based modeling of spatial-temporal LBSN data (Cho, Myers, and Leskovec 2011; Gao, Tang, and Liu 2012a), here we argue that a venue-centric approach is sometimes preferable. For instance, if the goal is to predict future attendance of a particular venue or to measure the impact of an ad campaign on attendance, it is more natural to focus on the check-in dynamics of venues rather than users. While recent work has studied correlations between a venue’s characteristics and its popularity (Joseph, Tan, and Carley 2012), the dynamics of venue-specific check-ins have been largely ignored.

We focus on modeling the full temporal dynamics of check-ins from a venue-centric perspective. We observe that check-ins at venues are clustered in time, sometimes exhibiting bursty behavior. We also observe that the average check-in patterns for both users and venues are not static, but change over time. We include three primary mechanisms to describe check-in dynamics: (1) Repeated behavior is captured by a self-reinforcing mechanism in which a user is strongly influenced by his recent behavior; (2) Social influence, i.e., a visit by a user triggers future visits by his friends; and (3) Exogenous effects, which include external events (such as releasing new SW for the service or a promotion campaign) that modulate the attendance rates.

Here we are especially interested in assessing social influence on visitation patterns. Toward this goal, we adopt a parametric point process model known as a Hawkes process (Hawkes 1971) to describe check-in dynamics at venues. A Hawkes process is an example of a self-exciting point process in which past events positively influence the likelihood (intensity) of future events. This model allows us to measure the likelihood that a particular (offspring) event was triggered by a past (parent) event. This allows us to dis-

tistinguish the most likely factors contributing to an individual check-in. Combining this information with the known social network structure enables us to estimate the fraction of check-ins that can be plausibly attributed to social influence.

Beyond the rich explanatory power of the model, we also demonstrate that it predicts future check-in data better than several alternatives. In particular, we consider various baseline point process models and compare them on their ability to capture temporal dynamics of check-ins. Finally, we consider each of the three mechanisms in our model separately and demonstrate their validity by distinguishing social and non-social venues and by capturing known exogenous effects like (external) promotion campaigns. This multi-faceted analysis allows a fine-grained discrimination of different types of venues. While we focus on user/venue dynamics here, the mechanisms we describe are general and could apply to other aspects of human behavior.

Related Work There is a growing body of literature on LBSN analysis. Link prediction using geo-coincidences has been studied in (Crandall et al. 2010; Wang et al. 2011; Scellato, Noulas, and Mascolo 2011). Other studies have used social network information to infer user location (Sadilek, Kautz, and Bigham 2012; Gao, Tang, and Liu 2012b) and predict next check-in (McGee, Caverlee, and Cheng 2013). Several recent studies have also attempted to cluster users (Joseph, Tan, and Carley 2012) and venues (Cranshaw et al. 2012) based on similar visitation patterns.

Most human activity patterns have bursty dynamics and cannot be adequately described by homogenous Poisson process. To describe temporal correlations in social interactions, researchers have used Non-homogeneous Poisson processes (NHPP) such as Cox-process (Lando 1998; Perry and Wolfe 2013), as well as and Hidden Markov Models (Raghavan et al. 2013). Our approach here is based on self-exciting Hawkes process (Hawkes 1971) that has been previously used for modeling urban crime (Mohler et al. 2011), inter-gang violence (Cho et al. 2013), and repeated social interactions (Blundell, Heller, and Beck 2012).

2 Dataset Description

We use the Gowalla dataset (Cho, Myers, and Leskovec 2011) in this work. Gowalla is a location-based social networking website where users share their locations by checking-in. In this dataset, the network consists of 196,591 nodes and 950,327 undirected edges. Between February 2009 and October 2010, there were 6,442,890 check-ins. We extracted all the check-ins of active users in San Francisco, New York, and Stockholm as representative samples from western U.S., eastern U.S., and Europe. Asian cities had little activity and were excluded from analysis; Check-ins from a relatively active city in Asia (Tokyo) were a quarter of those in San Francisco. We collected all activity within a rectangular box of latitude-longitude coordinates around each of the selected cities. We considered only the 20% of users who were most active to ensure sufficient statistics for parameter estimation. The 20% of most active users represented around 80% of the total number of check-ins. This 80–20 rule was universal across all the cities we examined.

Statistics from each city are presented below.

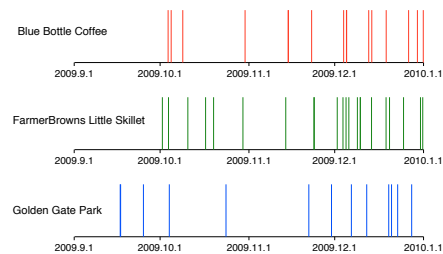
Table 1: Statistics of Check-ins from Three Cities

Dataset statistics		
San Francisco	Number of Check-ins	142,972
	Number of Venues	10,751
	Number of Users	5,989
New York	Number of Check-ins	114,777
	Number of Venues	17,062
	Number of Users	6,205
Stockholm	Number of Check-ins	184,485
	Number of Venues	15,753
	Number of Users	9,320

3 Model Description

3.1 Modeling Temporal Patterns

Figure 1: Temporal pattern of check-ins (SF)



We treat the check-ins in LBSN as a marked point process in time, where the mark represents the venue as well as the user for an event at a given time. By separating every process with respect to the venue id, each venue forms its own point process. We defer analysis of user-specific processes until Sec. 3.3. As shown in Figure 1, clustering is apparent in the three temporal point processes. Thicker lines represent the degree to which an event is explained by previous events (as opposed to background or exogenous effects). The strength of ties was mathematically computed using the self-exciting point process known as the Hawkes process (Hawkes 1971), detailed in the next section. The Hawkes process defines a (mark specific) intensity as a function of history and time. This model has been widely used in various applications that show temporal clustering of events such as shocks and after-shocks in seismology.

Hawkes process Each ‘check-in’ at a given venue is treated as an event in the given venue-specific point process. We assume that the intensity of check-in events involving the venue v at time t is given as follows:

$$\lambda_v(t) = \mu_v + \sum_{p:t_p < t} g_v(t - t_p). \quad (1)$$

This intensity function can be interpreted as a rate at which events occur (see Eq. 4). The summation in the second term is over all the events that have happened up to time t . In

Equation 1, μ_v describes the background rate of event occurrence that is time-independent, whereas the second term describes the self-excitation part, so that the events in the past increase the probability of observing another event in the (near) future. We will use a two-parameter family for the self-excitation term:

$$g_v(t - t_p) = \beta_v \omega_v \exp\{-\omega_v(t - t_p)\}. \quad (2)$$

Here β_v describes the weight of the self-excitation term (compared to the background rate), while ω_v describes the decay rate of the excitation. Intuitively, the decay term captures the notion that more recent events are more important.

3.2 Characterizing Correlations Between Events

We have seen the clustering of points in time in Fig. 1. A Hawkes process model allows us to measure the strength of ties between two events. By examining the intensity function in Equation 1 for a given event, we can further infer the likelihood that the event was triggered by a specific historical event. We use the probabilistic measure in Equation 3 as the strength of tie between i and j . For the given process (representing a specific venue, v), the probability that the j -th event is triggered by the i -th event can be expressed as below:

$$p_{i \rightarrow j}^v = \frac{g_v(t_j - t_i)}{\mu_v + \sum_{p: t_p < t_j} g_v(t_j - t_p)}. \quad (3)$$

The probability above can be inferred based on the estimated set of parameters $\{\mu_v, \beta_v, \omega_v\}$. Since we are interested in correlation of points for a given process (venue) and not the correlation across different processes (venues), we assume each process (venue) has its own parameters and we estimate them separately. We use the EM algorithm for our inference and estimation of model parameters. We follow the update equations of the parameters from (Lewis and Mohler 2011).

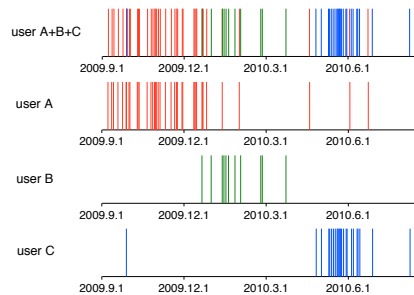
3.3 Three Factors causing Temporal-Clustering

Three factors that contribute to temporal clustering of events are considered in our studies. We describe each in turn below. We are able to disentangle these events because of the rich information in the data that include the user and venue for each event along with a social network among users. This allows us to construct a fine-grained model of the strength of the effect of one visit on another. While the ground truth cause of each visit is unknown, in the next section we consider various ways to qualitatively test the validity of our model.

Self-reinforcing Behavior Looking at the behavior of individual users already reveals strongly predictable patterns. Many users return frequently and repeatedly to the same venue. Figure 2 shows the activity of three users on Dolores Park Cafe in San Francisco. User A,B,C (bottom 3) organize their own temporal clusters, which forms a series of clusters when collected (top). A user who has recently visited a venue is much more likely to visit again soon and, conversely, a paucity of visits strongly predicts few visits in the future. This self-reinforcing tendency is measured using

Eq. 3 by summing over events i and j that were initiated by a single user. Later in our study, we see how individuals' overall activity decays over-time.

Figure 2: Three users' activity profiles at Dolores Park Cafe



Social Effects Another factor explaining temporal clustering is social influence. In this case, a user may be more likely to visit a location his friends have visited recently. LBSN allows users to see their friends' check-ins, and this in turn attracts users to visit the same venues. This effect may function by recommending venues to friends who likely share similar interests or simply by reminding users of places they have visited in the past. The increased likelihood of visits due to previous visits by friends is again captured by Eq. 3 but this time by summing only over events involving a user's friends in the social network.

Exogenous Effect If a visit is not explained by either of the effects above, we consider it to be caused by some external (exogenous) factor. Many businesses also use LBSN for marketing purposes. By reporting the physical location of their business, their venue becomes visible in the service, which allows the LBSN service users to visit in the future. Local businesses promoted check-ins to increase visibility online and to entice new customers by offering special deals. In fact, these marketing activities were not limited to local businesses. LBSN services also teamed-up with major companies for their own marketing. Gowalla attracted users in the early stages of their business by giving away presents to active Gowalla users. The influx of users during this period also forms a cluster which could not be described by the two effects above. In our studies we see how these are captured as an outside effect.

4 Experimental Evaluation

4.1 Model selection

For every popular venue, we fit the data to a Hawkes process using the EM algorithm and evaluate the goodness of fit compared against other baseline approaches (see Section 4.2 for the list of baselines). For evaluation we use the AIC score (Burnham and Anderson 2002), which has been widely used for model selection. In addition to maximizing likelihood, AIC also penalizes models with large number of parameters to discourage overfitting. The model with

the smallest score is chosen from the candidates. In our experiments, we found that the baselines generally compared poorly with the homogenous Poisson process (HPP), so here we focus on comparison of Hawkes process with HPP.

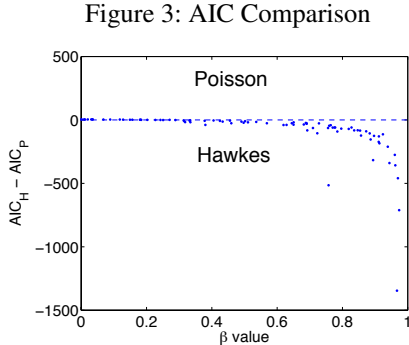


Fig. 3 shows the difference in AIC scores plotted against the inferred value of the parameter β . Positive difference (above the dashed line) suggests a better fit for HPP, while negative difference suggests that the Hawkes process model should be selected. We observe that Hawkes process is the better choice overall, and especially so for the venues with large β . When β is large, our model predicts significant temporal clustering, so it is natural to expect that the Hawkes process should do a better job explaining the data. We also note that when β decreases, the gap in AIC between the two models becomes small, and actually reverses sign when $\beta \rightarrow 0$. This is also understandable, because if there is no temporal clustering, there is no advantage to include a self-exciting term in the model. We would like to note, however, that in the latter scenario the difference between the AIC scores is minuscule, compared to significant differences observed for large β .

4.2 Predicting Venue attendance

In this experiment, we predict the number of daily visitors in the future. For all the venues, we compute the mid-time, which is the mid-point between initial check-in time and final check-in time on each venue. To have enough temporal data for the training set, we sample the time that appears after mid-time. The check-ins made before the mid-point are collected as a training set. Using the training set, we fit the data to a Hawkes process and estimate the parameters. With the estimated parameters, the rate function at time t can be computed based on the history up to time t and the parameters estimated from the training set. The number of events between time interval t and $t + \Delta t$ can be computed using the counting process as below ($\Delta t > 0$):

$$N(t + \Delta t) - N(t) = \int_t^{t+\Delta t} \lambda(\tau) d\tau. \quad (4)$$

In our experiment, we focus on predicting daily check-ins, so we set $\Delta t = 24$ hrs. The time, t is randomly sampled from some random time that includes at least half of the data so that we have enough history for parameter estimation. For

each venue, we repeat the experiment 1,000 times for different random t 's and compare our prediction to the actual number of events. The prediction error is computed using the gap between the actual number of events and our prediction: $abs(\text{true count} - \text{predicted count})$. The number of predicted events is estimated using Equation 4. We compare the Hawkes process to several other baselines including non-homogenous Poisson processes (NHPP) and Cox processes.

Baseline 1: piecewise-constant NHPP

Check-in data from three cities shows strong activity during the weekend compared to weekdays. We separate weekend check-ins and weekday check-ins to estimate the rate parameter respectively. Each parameter is λ_{weekend} and λ_{weekday} is constant and can be easily estimated. For predicting the number of visits on a given day, we simply use the appropriate rate parameter for weekends or weekdays. As for the Hawkes process, we repeat the experiment 1,000 times for each venue.

Baseline 2: NHPP with drifting

$$\lambda(t) = at + b \quad (5)$$

We define the rate function as a linear function of time. On many venues, the check-ins became more frequent as time elapsed from the first check-in. This is because more and more people joined the Gowalla service after its introduction. This intensity function well captures the birth of users, and we see that this simple intensity function predicts the number of visitors relatively well.

Baseline 3: Cox proportional hazard model

A Cox process is a generalization of Poisson process where the random intensity is a stochastic process. Cox proportional hazard model (Cox 1972) associates covariates which modulate up or down to the baseline rate $\lambda_0(t)$. Often the baseline is a function of time with non-parametric form while the covariates involve coefficient β which is estimated using the partial likelihood. For our experiment, we define $x(t)$ as a number of unique visitors assuming the size of unique visitors in the past affects the intensity function.

$$\lambda(t) = \lambda_0(t) \exp(\beta x(t)) \quad (6)$$

In our experiment, we assume the baseline rate as constant, and repeat the same experiment for baseline 3 with the same sampled time/ training set.

Baseline 4: Sigmoidal Gaussian Cox process

$$\lambda(s) = \lambda^* \sigma(g(s)) \quad (7)$$

This is another variant of the Cox process for which the intensity function is a transformation of a random realization from a Gaussian process. Adams et. al. (Adams, Murray, and MacKay 2009) suggested this model which achieves tractable inference on unknown intensity function. The random intensity function $\lambda(s)$ has an upper-bound λ^* and a sigmoid function which projects $g(s)$ to the intensity function where the $g(s)$ is sampled from Gaussian process. In (Adams, Murray, and MacKay 2009), sigmoidal Gaussian Cox process (SGCP) inferred the intensity functions defined in simple form in their synthetic experiment. We also compare the Hawkes process to SGCP on prediction of future events.

Error Comparison We use 360 venues (120 each) from three cities, and repeat 1,000 predictions for each sampled time range between t and $t + \Delta t$. The venues with few check-ins (less than 100 during 400 days) or with a short history (less than 200 days from the first check-in to the final check-in) were excluded in this experiment. Some venues have more frequent check-ins, hence we evenly divide the 360 venues into three groups based on the total number of actual check-in counts from 1,000 test samples, and name them inactive/moderate/active venues reflecting fewer to more check-ins, respectively. Splitting venues by activity was done on a per city basis to avoid city-specific bias due to higher average usage. As for the 1,000 test samples, we divide them into two groups, ones which had no check-ins (zero) and others which had more than zero check-ins. On average, $\sim 70\%$ of test samples from inactive venues had zero check-ins, $\sim 50\%$ of test samples from moderate venues fell into the zero group, and $\sim 35\%$ of test samples from active venues fell into the zero group.

Table 2: Performance of Predictions

Process	Obs. (each sample)	Avg. Prediction Error		
		inactive venues	moderate venues	active venues
Hawkes	zero	0.3202	0.3710	0.7210
	non-zero	0.7238	0.5361	0.4712
Baseline 1	zero	0.3273	0.4455	1.0306
	non-zero	0.7305	0.7029	0.5937
Baseline 2	zero	0.5318	0.6795	1.6901
	non-zero	0.6011	0.5707	0.5086
Baseline 3	zero	0.7289	0.9347	2.2185
	non-zero	0.6040	0.6361	0.5660
Baseline 4	zero	0.2477	0.4037	0.9989
	non-zero	0.7927	0.7331	0.5990

Th results are presented in Table 2. The average prediction error for zero group has been averaged over the total number of zero occurrence in the test sample, while the average prediction error for non-zero group has been averaged over the total number of the counts in the test sample. We observe that Hawkes process clearly outperforms the other baselines for venues with moderate and high activity levels. In particular, for those venues the Hawkes process produces more accurate prediction for both the rate of events when they occurred, and the absence of events. For the inactive venues, we find that Baseline 4 makes more accurate predictions for non-events, while Baseline 2 (and also Baseline 3) make better prediction of the rate of events when they occur. The former observation can be attributed to the fact that Baseline 4 tends to under-predict, which results in low prediction error for non-events and the higher prediction error (among all methods) for events. Out of all the processes under consideration, Hawkes process is the only process which captures influence between check-ins while the other processes only capture fluctuation of rates over time.

4.3 Evaluating the Three Factors

We now focus on understanding the relative importance of the three main factors put forward in Section 3.3 that are responsible for temporal clustering. Toward this goal, recall that the (directional) correlation between two check-in events can be measured using Equation 3. Furthermore, by using the existing social network information, we can estimate the relative contribution of each factor by analyzing the identity of users in those check-ins. Namely, if both check-ins are by the same user, then the event pair contributes to the *self-reinforcing behavior*. Similarly, when the check-ins are by two users that are connected in the social network, then the events contribute to the *social effect*. Finally, event pairs that belong to neither of these groups are attributed to *exogenous effects*. Since we are interested in differentiating the strength of effects, we separately add all the pairs of $p_{i \rightarrow j}^v$ in Equation 3 for the cases when event i and j involve the same person, or two people who are friends with each other, or neither of these two. To understand which factor contributes more to the temporal patterns, we define the following scores corresponding to each of three factors, respectively:

$$\begin{aligned}
 S_{\text{self}} &= \frac{\sum_{t_i < t_j} p_{i \rightarrow j}^v \mathbf{1}_{\langle u_i = u_j \rangle}}{\sum_{t_i < t_j} p_{i \rightarrow j}^v} \\
 S_{\text{social}} &= \frac{\sum_{t_i < t_j} p_{i \rightarrow j}^v \mathbf{1}_{\langle u_j \in F(u_i) \rangle}}{\sum_{t_i < t_j} p_{i \rightarrow j}^v} \\
 S_{\text{exgn}} &= 1 - S_{\text{self}} - S_{\text{social}} \quad (8)
 \end{aligned}$$

where $\mathbf{1}$ is an indicator function, u_i is the user corresponding to event i , and $F(u_i)$ is the set of friends of u_i . We measured the above scores for 120 venues from the three cities.¹ In Table 3 we show the top 5 venues with estimated high self-reinforcing behavior score.

Table 3: Top 5 Venues with Self-Reinforcing Behavior

San Francisco	New York
Laguna Honda station	G. Washington bridge
Bernie’s (local coffee)	Manhattan bridge
San Francisco Caltrain station	Port authority bus terminal
Mail Access	Lincoln tunnel
Research Institute	Grand central terminal

Interestingly, high S_{self} score seems to capture venues that reflect repeated behaviors such as commuting from work to home or regularly visiting favorite local places.

We next turn to venues that are characterized by high social effect score S_{social} . Intuitively, we expect that such venues will consist of bars or local restaurants in a community that have higher chances of attracting users who are friends of each other compared to other type of venues such as popular tourist attractions or stations where large numbers of random

¹We excluded Stockholm from the analysis below because most of the venues there were hard to identify using geo-coordinates only.

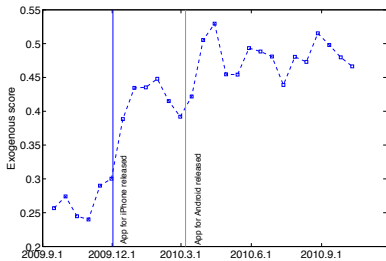
users visit. We list the venues with high social effect scores in Table 4 below. Indeed, based on the name and type of these venues, these venues seem to intuitively reflect what might be expected for highly social venues.

Table 4: Top 5 Venues with High Social Effect

San Francisco	New York
303 second st. Plaza Chinatown (restaurant)	Tasti D lite (ice cream) Cafe 28
Golf smith (shop) or others Restaurant (name unknown)	Moschino Botique Ace Hotel NY
Western athletic clubs	Radio City Music Hall

Finally, we analyze the relative importance of, and temporal variations in the *exogenous effects* as predicted by our model. Toward this goal, we average S_{exgn} over all the popular venues in San Francisco for all the check-ins during a given week, and then track the variation of the averaged score over time. The resulting dynamics is shown in Figure 4. We observe that the exogenous score increases starting from $S_{\text{exgn}} \approx 0.25$ in Sep, 2009 and reaches over 0.5 in June, 2010. Remarkably, the onsets of two growth periods as identified in the figure correspond to important external events. Namely, according to Gowalla blog, the company released its software for iPhone users on Dec. 2nd 2009, and for Android users on March 7th 2010. We clearly see steep growth of S_{exgn} after the release of the software, thus vindicating our intuition. More generally, we believe that by tracking the dynamics of S_{exgn} (perhaps with only local averaging), it might be possible to detect the impact of even smaller promotional events.

Figure 4: Exogenous Effect Score S_{exgn} plotted against time (San Francisco)

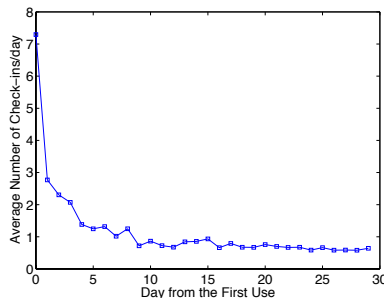


Comparison between Cities Between three cities, San Francisco showed the highest score in *social effect*, meaning many of the events within a cluster had social relationships. For the experiment, we used the top ~ 120 venues from each city and compared the average social effect. To learn the Hawkes process, enough samples are needed, which led us to use ~ 120 venues. The average *social effect* scores from San Francisco, New York and Stockholm were **0.0895**, **0.0220**, and **0.0188**, respectively. The average social density, which we defined as the fraction of true edges over all the

possible pairs, for all collected venues was measured. The social densities are **0.0374** (San Francisco), **0.0552** (New York), **0.0406** (Stockholm). Interestingly, while San Francisco had the lowest social density, it showed the highest *social effect* score. This is because the users who were friends in San Francisco visited venues during same time period while friends in other cities did not cluster as much of their activity at similar times and venues.

Long-term Activity Trends We conclude this section by commenting on activity patterns of individual users in the long term. As indicated in Figure 2, the temporal patterns of individuals' behavior seem to decay over time. One possibility is that a user stops visiting a venue. Another possible explanation is that users tend to use the Gowalla service less often as time passes. Indeed, this trend is indicated in Figure 5, where we plot the average number of check-ins (averaged over all the active users in all three cities) against the time passed after the first check-in. We observe that user activity rapidly decays after the first week of use. Remarkably, this decay seems to be fairly universal across the cities. A more detailed analysis of the user-turnover dynamics is an interesting open problem.

Figure 5: Average Number of Check-ins with Respect to the Days Since the First Check-in (San Francisco)



5 Conclusion and Future Work

In conclusion, we have introduced a point process model describing check-in activity for users and venues participating in LBSNs. Our model, which is based on self-exciting Hawkes process, outperforms benchmarks in both data explanation and prediction tasks. More importantly, the proposed approach allows to construct a fine-grained view of events that enables us to distinguish relevant factors like self-reinforcing behavior, social effects, and exogenous effects. Qualitative results suggest that we are able to meaningfully distinguish these factors. Future work will provide a more in-depth analysis of these important effects and their repercussions on human mobility patterns.

Acknowledgments

This research was supported in part by DARPA grant No. W911NF-12-1-0034.

References

- Adams, R. P.; Murray, I.; and MacKay, D. J. C. 2009. Tractable nonparametric bayesian inference in poisson processes with gaussian process intensities. In Danyluk, A. P.; Bottou, L.; and Littman, M. L., eds., *ICML*, volume 382 of *ACM International Conference Proceeding Series*, 2. ACM.
- Blundell, C.; Heller, K. A.; and Beck, J. M. 2012. Modelling reciprocating relationships with hawkes processes. In Bartlett, P. L.; Pereira, F. C. N.; Burges, C. J. C.; Bottou, L.; and Weinberger, K. Q., eds., *NIPS*, 2609–2617.
- Burnham, K. P., and Anderson, D. R. 2002. *Model selection and multimodel inference: a practical information-theoretic approach*. Springer, 2 edition.
- Cho, Y. S.; Galstyan, A.; Brantingham, J.; and Tita, G. 2013. Latent point process models for spatial-temporal networks. *arXiv:1302.2671*.
- Cho, E.; Myers, S. A.; and Leskovec, J. 2011. Friendship and mobility: user movement in location-based social networks. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, KDD '11, 1082–1090. New York, NY, USA: ACM.
- Cox, D. R. 1972. Regression models and life-tables (with discussion). *Journal of the Royal Statistical Society B*(34):187–220.
- Crandall, D. J.; 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):22436–22441.
- Cranshaw, J.; Schwartz, R.; Hong, J. I.; and Sadeh, N. 2012. Utilizing social media to understand the dynamics of a city. In *ICWSM'12*.
- Gao, H.; Tang, J.; and Liu, H. 2012a. Exploring social-historical ties on location-based social networks. In Breslin, J. G.; Ellison, N. B.; Shanahan, J. G.; and Tufekci, Z., eds., *ICWSM*. The AAAI Press.
- Gao, H.; Tang, J.; and Liu, H. 2012b. gscorr: modeling geo-social correlations for new check-ins on location-based social networks. In *Proceedings of the 21st ACM international conference on Information and knowledge management*, CIKM '12, 1582–1586. New York, NY, USA: ACM.
- Gonzalez, M. C.; Hidalgo, C. A.; and Barabasi, A.-L. 2008. Understanding individual human mobility patterns. *Nature* 453(7196):779–782.
- Hawkes, A. G. 1971. Spectra of Some Self-Exciting and Mutually Exciting Point Processes. *Biometrika* 58(1):83–90.
- Joseph, K.; Tan, C. H.; and Carley, K. M. 2012. Beyond "local", "categories" and "friends": clustering foursquare users with latent "topics". In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*, UbiComp '12, 919–926. New York, NY, USA: ACM.
- Lando, D. 1998. On cox processes and credit risky securities. *Review of Derivatives Research* 2(2-3):99–120.
- Lewis, E., and Mohler, G. 2011. A Nonparametric EM algorithm for Multiscale Hawkes Processes. *preprint*.
- McGee, J.; Caverlee, J.; and Cheng, Z. 2013. Location prediction in social media based on tie strength. In *CIKM*, 459–468. ACM.
- Mohler, G. O.; Short, M. B.; Brantingham, P. J.; Schoenberg, F. P.; and Tita, G. E. 2011. Self-Exciting Point Process Modeling of Crime. *Journal of the American Statistical Association* 106(493):100–108.
- Perry, P. O., and Wolfe, P. J. 2013. Point process modelling for directed interaction networks. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 75(5):821–849.
- Raghavan, V.; Steeg, G. V.; Galstyan, A.; and Tartakovsky, A. G. 2013. Modeling temporal activity patterns in dynamic social networks. *CoRR* abs/1305.1980.
- Rhee, I.; Shin, M.; Hong, S.; Lee, K.; Kim, S. J.; and Chong, S. 2011. On the levy-walk nature of human mobility. *IEEE/ACM Trans. Netw.* 19(3):630–643.
- Sadilek, A.; Kautz, H.; and Bigham, J. P. 2012. Finding your friends and following them to where you are. In *Proceedings of the fifth ACM international conference on Web search and data mining*, WSDM '12, 723–732. New York, NY, USA: ACM.
- Scellato, S.; Noulas, A.; and Mascolo, C. 2011. Exploiting place features in link prediction on location-based social networks. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, KDD '11, 1046–1054. New York, NY, USA: ACM.
- Wang, D.; Pedreschi, D.; Song, C.; Giannotti, F.; and Barabasi, A.-L. 2011. Human mobility, social ties, and link prediction. In *Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining*, KDD '11, 1100–1108. New York, NY, USA: ACM.