

# Digitizing Offline Shopping Behavior Towards Mobile Marketing

*Completed Research Paper*

**Anindya Ghose**

Leonard N. Stern School of Business

New York University

44 West 4<sup>th</sup> St., New York, NY 10012

[aghose@stern.nyu.edu](mailto:aghose@stern.nyu.edu)

**Beibei Li**

Heinz College

Carnegie Mellon University

5000 Forbes Ave, Pittsburgh, PA 15213

[beibeili@andrew.cmu.edu](mailto:beibeili@andrew.cmu.edu)

**Siyuan Liu**

Smeal College of Business

Penn State University

University Park, PA 16802

[sxl68@psu.edu](mailto:sxl68@psu.edu)

## Abstract

*The proliferation of mobile technologies makes it possible for mobile advertisers to go beyond the real-time snapshot of the static location and contextual information about consumers. In this study, we propose a novel mobile advertising strategy that leverages full information on consumers' offline moving trajectories. To evaluate the effectiveness of this strategy, we design a large-scale randomized field experiment in a large shopping mall in Asia based on 83,370 unique user responses for two weeks in 2014. We found the new mobile trajectory-based advertising is significantly more effective for focal advertising store compared to several existing baselines. It is especially effective in attracting high-income consumers. Interestingly, it becomes less effective during the weekend. This indicates closely targeted mobile ads may constrict consumer focus and significantly reduce the impulsive purchase behavior. Our finding suggests marketers should carefully design mobile advertising strategy, depending on different business contexts.*

## Introduction

Smartphone usage is expected to exceed 6.1 billion users worldwide by 2020 (Ericsson 2014). The proliferation of mobile and sensor technologies has contributed to the rise of mobile location-based advertising. Such advertising can enable businesses to deliver information to mobile users in real time about offers in geographical proximity to them. Recent studies using randomized field experiments have causally shown that mobile advertisements based on static location and time information can significantly increase consumers' likelihood of redeeming a geo-targeted mobile coupon (Molitor et al. 2014, Luo et al. 2014, Fong et al. 2014), that mobile ads have a synergistic relationship with PC ads (Ghose et al. 2013), that expiry length of mobile coupons influences their redemption rates (Danaher et al. 2015), and that understanding consumers' context, in particular the crowdedness of their environment, is important for marketers to increase mobile marketing effectiveness (Andrews et al. 2015).

Beyond the real-time snapshot about the static geographical location and the consumer contextual information, the overall mobile trajectory of each individual consumer can provide even richer information about consumer preferences. In particular, "trajectory" hereby refers to the physical behavioral trace of an individual's offline movement. For example, it can include information on the locations where the individual has been to in the past, at what time, for how long, as well as the associated contexts. Considering the significant search costs of consumers in the offline world, such physical behavioral trace of individuals can be highly informative in revealing consumer preferences for real-time decision making. This information is analogous to the search-click stream data that we have been studying in the online environment. Mobile and sensor technologies allow us to digitalize such individual behavioral trajectory in the offline environment. Hence, extracting and measuring the digital trace of consumers' offline behavior has become increasingly critical for businesses today in order to understand consumers' inherent preferences to improve user digital experiences and business marketing strategies.

In this study, we propose a new mobile advertising strategy that infers consumer preferences by leveraging full information on consumers' offline moving trajectories from four different mobility dimensions: temporal duration, spatial dispersion, semantic information, and movement velocity. We extract these multidimensional trajectory features from large-scale, fine-grained user-level behavioral data. We learn individual preferences from the trajectory information, using statistical and machine learning methods such as kernel-based similarity functions, dense sub-graph detection algorithms for graph-based clustering, and collaborative filtering. Finally, to examine the effectiveness of this new mobile trajectory-based advertising strategy, we conduct a randomized field experiment in one major large shopping mall in Asia in June 2014. Our experiment results are validated based on 83,370 unique user responses for a 14-day period. Our follow-up econometric analyses from both group level and individual user level demonstrate high consistency in our final results.

Our main findings are the following. First, we find that mobile trajectory-based advertising can significantly increase the likelihood of a consumer redeeming a mobile promotion at the focal advertising store. On average, the coupon-redemption rate is 34.78% higher when compared to mobile advertising based on real-time location information, and it is 93.75% higher when compared to mobile advertising based on purely random recommendations. Second, the mobile trajectory-based advertising leads to the fastest redemption behavior from customers. On average, the time elapsed between receiving the mobile promotion and redeeming it under mobile trajectory-based advertising is approximately one-third the time elapsed under the static location-based mobile advertising (4.55 vs. 12.83 minutes), and one-fourth the time under the purely random-based mobile advertising (4.55 vs. 16.43 minutes). Third, our follow-up survey shows that mobile trajectory-based ads lead to the highest overall satisfaction rates and the highest future-willingness-to-redeem rates from the customers. Fourth, interestingly, our findings indicate that mobile trajectory-based advertising becomes much less effective during the weekend, compared to pure location-based advertising or random advertising strategies. This finding suggests customers are likely to be in an "exploratory" shopping stage during the weekend and are likely to incur "impulse buys." However, closely targeted mobile advertising tends to constrict consumer focus and significantly reduce the impulse purchase behavior. This finding suggests that businesses and marketers need to be careful when implementing mobile advertising strategies, depending on different contexts. Finally, we also find that mobile trajectory-based advertising is especially effective in attracting high-income consumers, suggesting the potential of mobile advertising in approaching high-end customers to achieve better customer lifetime value.

Our major contributions can be summarized as follows. First, we demonstrate the value of mining large-scale, fine-grained offline mobile trajectory information on understanding individual preferences, and the importance of leveraging such information to improve the effectiveness of digital marketing. Second, we are able to establish a link between individuals' offline behavioral trajectories and their digital behaviors. To the best of our knowledge, our work is the first to bridge the understanding of individual behavioral path and preference from both physical and digital worlds. Third, our analyses based on a combination of field experiments and surveys allow us to quantify the economic effects of the new mobile trajectory-based advertising approach from a causal perspective. Advertisers can learn from our results in order to improve the design and effectiveness of their mobile marketing strategies. Finally, our interdisciplinary approach incorporates methodologies from large-scale statistical and machine learning, econometric modeling, field experiments, and surveys. It provides an application opportunity in data-driven decision-making processes. It also paves a path on which future studies can build to analyze problems that lie at the intersection of technology and social science.

## Related Literature

Our study builds on the following three streams of research.

### ***Mobile Marketing and Location-based Advertising.***

Our paper is highly related to mobile and location-based advertising. Recent studies have shown that mobile ads have a synergistic relationship with PC ads, and geographical proximity matters more for mobile ads than for PC ads (Ghose et al. 2013). Using randomized field experiments, researchers have causally shown that mobile ads based on static location and time information can significantly increase users' likelihood of redeeming a geo-targeted mobile coupon (Molitor et al. 2014, Luo et al. 2014, Fong et al. 2015). Molitor et al. (2014) have shown that the higher the discount from mobile coupons and the closer the consumers are to the physical store offering the coupon, the more likely they are to download the mobile coupons. Luo et al. (2014) have found that temporal targeting and geographical targeting individually increase sales. However, the sales effects of employing these two strategies simultaneously are not straightforward, and advertisers need to carefully choose both temporal and spatial dimensions when designing mobile strategies. Furthermore, Fong et al. (2015) have focused on the effectiveness of competitive locational targeting, and found that competitive locational targeting can produce increasing returns to promotional discount depth. More recently, studies have shown that understanding consumers' hyper-context, such as the crowdedness of their immediate environment, is critical for marketers to measure mobile marketing effectiveness (Andrews et al. 2015). In particular, the authors have found that the more crowded the customer's current location environment, the more likely the customer will respond to a mobile ad. Danaher et al. (2015) show that besides location and time of delivery, how long m-coupons are valid (expiry length) can influence redemption rates, because redemption times for m-coupons are much shorter than for traditional coupons. Our paper distinguishes itself from all the existing studies by leveraging the full historical information on consumers' digitalized offline trajectories from four different mobility dimensions, including temporal duration, spatial dispersion, semantic information, and movement velocity, to infer preferences and improve mobile advertising.

Moreover, previous studies have examined consumer perceptions and attitudes toward mobile location-based ads (e.g., Brunner and Kumar 2007; Xu et al. 2009). Gu (2012) examines both the short- and long-term sales effects of location-based advertising. Bart et al. (2014) study mobile advertising campaigns and find they are effective at increasing favorable attitudes and purchase intentions for higher- (versus lower-) involvement products, and for products that are seen as more utilitarian (vs. more hedonic).

### ***Spatial-Temporal Mining and Trajectory Clustering.***

Second, our study builds on the spatial-temporal mining and trajectory-clustering literature from machine learning. Researchers have studied trajectory, using a variety of measures, ranging from mining frequent trajectory patterns for activity monitoring (Liu et al. 2012), probability function of time (Gaffney et al. 1999), behavior correlation representation (Xiang et al. 2010), density-based distance function (Nanni et al. 2006) and uncertainty measurement of trajectories (Pelekis et al. 2011). Different similarity measures (e.g., time and location distances) and clustering methodologies have strengths and weaknesses. In contrast to most prior work, our method is able to handle multiple information sources (not just movement trajectories, but also the semantics of the underlying space) and apply a general metric-based

learning framework to the clustering problem. Studies have used trajectory-based clustering for different broad objectives, such as discovering common sub-trajectories (Lee et al. 2007) and identifying spatial structures (Ng et al. 2002). But such work is based purely on spatial locations, making extending it to incorporate semantic, velocity, or other information that may contain distinctive markers of real community interaction difficult. It is also related to the community-detection literature from machine learning and computer science. Communities in networks/graphs are groups of vertices within which connections are dense, but between which connections are sparser. Four types of methods primarily exist: hierarchical clustering (Huang et al. 2010), similarity in edge-betweenness scores (Leskovec et al. 2010), counts of short loops (Newman et al. 2004), and voltage differences in resistor networks (Shi et al. 2011). However, these existing methods focus on detection given a network structure and social-link distance between nodes, which are difficult to capture from physical mobile trajectories. Instead, in our study, we focus on detecting communities of similar users based purely on their movement trajectory patterns.

### ***Behavior-Based Recommendation.***

Finally, our work is related to the stream of literature on recommendation systems, especially behavior-based recommendation. Link, content, and location can be viewed as the results of users' different behaviors, but little previous work builds trajectory community models to provide the online recommendation. In recommender systems, behavior models are proposed for different purposes, such as effects of behavior monitoring and perceived system benefits (Nowak et al. 2012), navigational patterns to model relationships between users (Esslimani et al. 2011), effect of context-aware recommendations on customer purchasing behavior and trust (Adomavicius et al. 2011, Gorgoglione et al. 2011), and utility query recommendation by mining users' search behaviors (Ghose et al. 2012, 2014; Zhu et al. 2011). Compared with the previous studies, one unique feature in this paper is that we aim to model individual preferences based on the large-scale and granular information extracted from individuals' heterogeneous offline behavior using mobile trajectories and offline contexts.

## **Extraction of Multidimensional Mobility Features from Individual Trajectories**

In this section, we discuss how we characterize individual mobility by extracting unique movement features from various dimensions of individuals' mobile trajectories. Building upon prior literature (Liu et al. 2013), we focus on four different dimensions of mobility features: temporal duration, spatial dispersion, semantic information, and movement velocity. Through these four mobility dimensions, we aim to capture similar patterns in individuals' physical movement from different perspectives. Note that this step allows us to learn consumer behavior not only through static locational or contextual proximity information, but also through dynamic movement similarity from the underlying mutual interaction or shared relationship.

### ***Temporal Duration***

We define temporal duration to contain information on the starting and ending time of the mobile trajectory, as well as the day-of-the-week index. More specifically, for each consumer, we extract a vector with three different temporal features: the starting time of a consumer's trajectory, the ending time of this trajectory, and the day index. These temporal features aim to capture the temporal activity pattern for real-life communities. To measure the similarity between two user trajectories on their temporal dimension, we adopt a similar approach as in Liu et al. (2013) using a temporal kernel.

### ***Spatial Dispersion***

Spatial dispersion measures the spatial alignment of different user trajectories. The close alignment of two trajectories might indicate high behavioral similarity between the two users. Note that to account for the popularity of the location, we inversely weigh the spatial similarity in proportion to the crowdedness of a specific location. Intuitively, this approach is similar to TF-IDF in text mining (e.g., Manning et al. 2008).

More specifically, our method builds on the Global Alignment Kernel (GAK) to measure the spatial similarity between two trajectories (Cuturi 2011, Liu et al. 2013). The intuition is to capture the spatial closeness between two individuals over time. However, the popularity of a location can potentially bias GAK. For example, if customers A, B, and 100 other customers are waiting in a concourse area, the spatial closeness between A and B becomes less informative of the similarity between them, because this

concourse is clearly a popular location for almost everyone. However, if  $A$  and  $B$  are the only two customers in the concourse, this spatial closeness can instead reveal significant information about the similarity between them. Based on this intuition, we apply the Global Alignment Kernel with Inverse Proportion method (GAK-IP) (Liu et al. 2013), which weighs the spatial similarity in inverse proportion to how many other people are co-located within the nearby area.

### **Semantic Information**

Semantic information aims to capture the contextual information related to the mobile trajectory. For example, it contains the stationary probabilistic distribution of individuals' visits to different stores in the mall, time spent at each store, time spent to transit from one store to another, and the transition probability between two stores. More specifically, let  $X$  denote the set of all trajectories. Our goal is to measure the traverse statistic on the sites, and use this statistic to measure the semantic correlation of user trajectories. Let  $L$  denote the total number of spatially distinct sites. We can extract the following features of the sites visited by an individual.

**Markov State Transition.** We construct the Markov state transition matrix  $A \in R^{L \times L}$ , where  $A(s_a, s_b)$  represents the transition probability from site  $s_a$  to site  $s_b$ . To calculate  $A$ , we first collect all the pairs from the overall trajectory set  $X$ . Then we count the number of occurrence of each transition pair. Finally, we conduct the column normalization of  $A$ , satisfying  $\sum_{s_b} A(s_a, s_b) = 1$ .

**Temporal Intervals.** We measure the time spent at each site and the time taken to transit from site  $s_a$  to site  $s_b$  to capture the "level of interest" shown by the users (e.g., when a shop is very "interesting," the shoppers may choose to stay longer), and the convenience of moving from site  $s_a$  to site  $s_b$ , which indicates the semantic relation of two sites.

### **Movement Velocity**

Finally, movement velocity contains information about the speed of moving subjects (users). The information encoded in the velocity pattern of moving subjects is critical. However, we face two challenges when modeling the velocity pattern. The first challenge is that the overall length of each individual trajectory is different, which causes difficulty in directly measuring their pairwise similarity from a velocity aspect. The second challenge is that even within the same individual mobile trajectory, velocity can vary largely at different times and locations; therefore, constructing the direct measurement is hard. To account for these challenges and to make velocity comparable across heterogeneous individual trajectories, we normalize the velocity by applying a temporal pyramid matching method (Liu et al. 2013). This method is inspired by the normalization method in calculating the image similarity in image classification while accounting for the different scales of resolution (Lazebnik et al. 2006).

More specifically, each trajectory has a velocity vector  $v_k$  with unequal lengths. We uniformly quantize the velocity into  $L$  levels. Given  $v_k$  with length  $l_k$ , we calculate the normalized histogram  $h_k(0)$  on  $v_k$ . Then we equally divide  $v_k$  into two parts  $v_k \rightarrow [v_k(1), v_k(2)]$ , where both  $v_k(1)$  and  $v_k(2)$  are also velocity vectors with length  $l_k/2$ . We then calculate the normalized histogram  $h_k(1)$  and  $h_k(2)$  on  $v_k(1)$  and  $v_k(2)$ , respectively, and normalize them so that  $\sum h_k(1) + \sum h_k(2) = 1$ . Consequently, we further equally divide  $v_k(1)$  or  $v_k(2)$  into two parts again and calculate the histograms in the same way. We continue this process until we achieve a predefined level. We concatenate all the histograms with predefined weights. Then, we can extract a velocity histogram  $h_k$  of equal length with coarse-to-fine temporal resolution. The pair-wise similarity between user trajectories  $k$  and  $k'$  can be calculated using different similarity functions, such as histogram intersection or chi-square kernel (Liu et al. 2013).

## **A New Mobile-Recommendation Approach**

In this section, we discuss how we design a new mobile-recommendation strategy based on the multidimensional mobility features we extract as described above.

### **Measuring Consumer Similarity from Multiple Trajectory Dimensions**

We first discover consumers with similar behavioral paths based on their mobile trajectories (Liu et al. 2013). This step aims to identify groups of similar consumers based on the behavior-driven mobility features of individuals as they move together in a shopping mall. It allows us to learn consumer

preferences not only through static spatial or contextual proximity information, but also dynamic movement similarity from the underlying mutual interaction or shared relationship.

We focus on inferring consumer similarity based on a combination of the four dimensions of mobility features we extract from the previous stage. Specifically, based on the above four dimensions of mobility features, we calculate the consumer similarity by combining the features as follows:

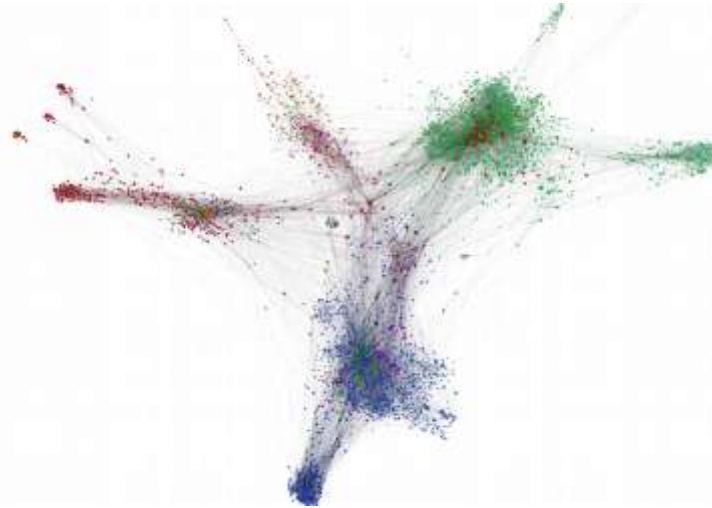
$$S(i, i') = \sum_{m=1}^M \alpha_m S_m(i, i'), \alpha_m \geq 0, \sum_{m=1}^M \alpha_m = 1, \quad [1]$$

where  $S(i, i')$  denotes the similarity of consumer  $i$  and consumer  $i'$ ,  $M$  denotes the number of dimensions of mobility features (here  $M = 4$ ),  $S_m(i, i')$  denotes the similarity on the  $m$ -th dimension of mobility features, and  $\alpha_m$  denotes the pre-assigned weights reflecting the specific interests of the problem domain. In this study, we obtain the weight  $\alpha_m$  using two different approaches. First, we assume an equal weight of 0.25 for each dimension. Alternatively, we are able to learn the weight using the machine learning method. In particular, we construct a training data set by manually rating the overall pairwise similarity between two trajectories on a scale from 0 to 1. Then, we use logistic regression to learn the corresponding weights based on the training set. For model evaluation, we use 10-fold cross-validation to avoid overfitting. We find the two approaches give us highly consistent results. We applied equal weights to the four mobility dimensions in our main experiments.

### Using Graph-based Clustering to Identify Similar Consumers

Based on the pairwise similarity of consumers derived from the previous steps, we can cluster similar individuals from their pairwise similarities. The main goal of this step is to identify clusters of consumers where the consumers within a cluster are similar to each other with regard to their mobile trajectories but dissimilar to consumers not in the cluster.

To do so, we use a graph-based clustering method. In particular, we apply the Markov Clustering Algorithm (MCL) for dense sub-graph detection (Van Dongen 2012). It is an unsupervised learning method that allows us to leverage the network structure to extract groups of similar items. MCL has several advantages (Satuluri et al. 2010) over distance-based clustering algorithms such as k-means (MacQueen 1967) and hierarchical clustering (Eisen et al. 1998). First, unlike the k-means-based algorithm, MCL is less sensitive to the initial starting conditions. Second, MCL does not take any default number of clusters as an input. Instead, it allows the internal structure of the network to determine the granularity of the cluster. Third, compared with many state-of-the-art network clustering algorithms, MCL is more noise tolerant and more effective at discovering the cluster structure (Brohee and Helden 2006, Liu et al. 2013).



**Figure 1. An example of Graph-Based Trajectory Clustering**

More specifically, we construct an undirected probabilistic graph of individual trajectories (an example is shown in Figure 1), where each node in the graph represents a consumer's trajectory, and the weight on each edge between two nodes represents the pairwise similarity between two consumers. Therefore, if two

consumers are very similar to each other in their trajectory patterns, the weight on the edge between the two corresponding consumer nodes would be very high. Our goal is to detect a set of highly connected sub-graphs (cliques) from the graph where the weight on the edge between each pair of two nodes in the sub-graphs is relatively high (i.e., dense sub-graph). The basic intuition of the MCL algorithm is based on the idea of a random walk. The probability of visiting a connected node is proportional to the weight on the edge. In other words, the random walk will stabilize inside the dense regions of the network after many steps. The stabilized regions shape the clustered sub-graph and reflect the intrinsic structure of the network. The sub-graphs hence represent the identified clusters of similar consumers.

### **Mobile Recommendation**

With the detected clusters of similar consumers from the previous steps, we then target advertising promotions by offering recommendations to a consumer from stores that are most frequently visited or products that are most frequently purchased by similar consumers. This approach is similar to the collaborative filtering approach widely used in traditional recommender systems.

In practice, recommendations are achieved by calculating the ratings of the consumers for the stores. More specifically, the rating of a consumer for a store is a measurement of one's interest in that store. It is defined as a weighted sum of time and money the consumer spent in that store. Given consumer  $i$  and store  $j$ , one common approach to predict the rating is to average the ratings of similar consumers on store  $j$  weighted by their similarity information. Thus, the average rating can be calculated by

$$\hat{R}(i, j) = \frac{\sum_{i'=1}^{N_i} R(i', j) S(i, i')}{\sum_{i'=1}^{N_i} S(i, i')}, \quad [2]$$

where  $N_i$  denotes the number of similar consumers to consumer  $i$ , and  $R(i', j)$  denotes the observed rating of consumer  $i'$  on store  $j$ . More specifically, our recommendation consists of four distinct phases:

1. Model trajectory similarities from four distinct mobility feature dimensions: semantic properties of the locations, spatial proximity to other objects, temporal duration of the trajectory, and movement velocity at different timescales;
2. Compute a weighted similarity measure that linearly combines the similarity measures along each dimension;
3. Cluster similar consumers using graph-clustering techniques (e.g., dense sub-graph detection algorithms) over a graph with edges corresponding to the computed pairwise similarity score;
4. Make recommendations to a consumer regarding most frequently visited stores or most frequently purchased products by similar consumers.

## **A Field Experiment on Mobile Trajectory-based Advertising**

To examine the effectiveness of the mobile trajectory-based advertising strategy, we designed and executed a large-scale randomized field experiment in collaboration with one of the largest shopping malls in Asia in June 2014.

### **Experimental Setting**

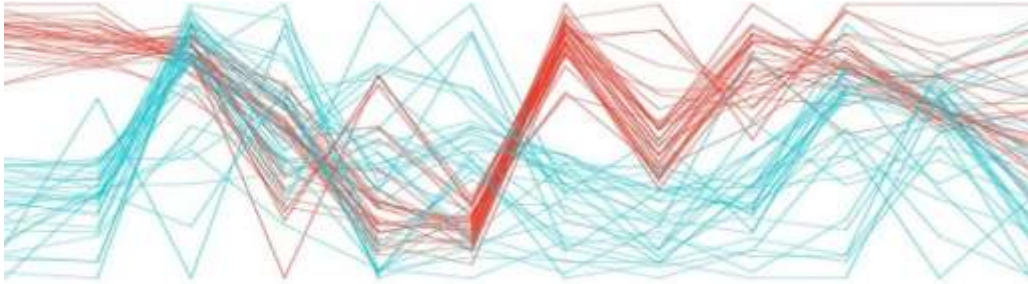
The shopping mall contains over 300 stores spanning 1.3 million square feet. On average, it attracts over 100,000 visitors daily. At the entrance of the shopping mall, if a consumer wanted to enjoy free WiFi, she was required to complete a Form A with information on age, gender, income range, credit card type (gold, platinum, gift card, others), and phone type (iPhone, Android, others). At each store, when the consumer purchased a product, she was required to complete a Form B, which involved similar socio-demographic information plus the amount spent and whether the purchase was related to a mobile coupon. We cross-validated Form A and Form B to check for the accuracy of the individual-level information.

Once the consumer connected to the WiFi, we were able to track the detailed mobile trajectory information during her visit in the shopping mall with precise time stamps. Finally, when the consumer left the mall, we conducted a short follow-up survey via mobile asking whether she followed the mobile



recommendation, whether she wanted to follow such recommendations in the future, overall satisfaction with the shopping experience, and additional personal information (first-time visitor or not, WiFi user or not, shop alone or with others, money spent in the focal advertising store, total money spent in the mall).

Figure 2 provides an example of movement trajectories of individual customers traveling upstairs and downstairs in the shopping mall. More specifically, the trajectories of customers contain information such as what kind of stores the customers visited, how long they stayed in each store, the transition probability between two stores, how fast they were walking, and so on. We are then able to generate mobile recommendations based on the four dimensions of mobility features extracted from the trajectory information as described in the previous section.



**Figure 2. Example Mobile Trajectories of Consumers in a Large Shopping Mall**

### ***Profiling Stores Based on Crowd-sourced Data***

To better contextualize the mobile trajectories and to capture detailed store-profile information in the shopping mall, we had to resort to crowdsourcing techniques. We applied a similar approach as in (Guo et al. 2013). In particular, we deployed ShopProfiler, a shop-profiling system on crowdsourcing data. ShopProfiler only relies on sensor readings from mobile devices. The process of data collection is automatic and running in the background. Inertial sensor readings reflect human movement information such as acceleration, heading, and speed. Microphone and WiFi modules provide additional information on surrounding conditions. From the customer's point of view, mobile-sensing data in a mall contain information about what shops that customer visited, how long she stayed, and how fast she was walking. From the shop's point of view, mobile-sensing data reveal information about the shops' inside layout and how many people visit the shops in a particular time period. Through mining and learning mobile-sensing data, we capture unique features of different shops and categorize them into different types (e.g., restaurant, books, supermarket, electronics, fashion, etc.). Furthermore, a complete profile of shops contains specific location and precise brand names. We extract the brand-name information from the Service Set Identifier (SSID) from each store using the text-mining method.

### ***Randomized Experiment Design***

We designed our randomized experiment to contain the following four groups:

- Control Group (C): Do not send any mobile promotion;
- Treatment Group 0 (T<sub>0</sub>): Send mobile promotion from randomly selected store;
- Treatment Group 1 (T<sub>1</sub>): Send mobile promotion based on real-time location information;
- Treatment Group 2 (T<sub>2</sub>): Send mobile promotion based on consumers' trajectory information.

We sent mobile coupons by using short message service (SMS) texts. Note that to control for the potential bias introduced by the stores and products, we randomized the participation among 252 stores from various categories including fashion, dining, supermarket, and so on. To control for the potential bias introduced by the coupon type, we considered different coupon designs with regard to both format and price discount, and randomized these coupon designs among the four experimental groups. For example, for the same store, we randomized the level of price discount (e.g., 10% off, 20% off, 30% off, or 50% off). For the same level of price discount, we also randomized the coupon format (e.g., "price 50% off" vs. "buy one get one free") to minimize the potential bias the coupon format might introduce. Moreover, to make sure our results were comparable across groups, we considered the same set of mobile promotions (in terms of both format and price discount) in T<sub>0</sub> as the ones used in T<sub>1</sub> and T<sub>2</sub>. The only difference is that the mobile promotions were sent randomly in T<sub>0</sub>, whereas they were tailored in T<sub>1</sub> and T<sub>2</sub>. Note that to



design real-time location-based mobile promotions (T1), we used a similar approach as the one used in the previous studies (e.g., Spiekermann et al. (2011), Ghose et al. (2013), Luo et al. 2014). In particular, we define “distance to a store” as the mobile user’s physical distance from the center of the store. We sent the real-time location-based mobile promotion to a consumer based on the store that had the shortest distance to the consumer at the moment of sending the coupon.

Finally, to control for the potential bias introduced by the timing of the coupon, we randomized the timing of when the mobile coupon was sent. Note that for efficiency and effectiveness of the recommendation, we conducted trajectory mining based on a large pool of historic individual consumer trajectories collected by the shopping in the past one year. This process allowed us to quickly identify trajectory similarity when a new customer walked into the shopping mall. Moreover, to avoid “cold start,” we waited for a random time period ( $\geq 10$  mins) after the customer walked into the mall before sending mobile coupons.

| Variable              | Definition   | Mean                         | Std. Dev. | Min    | Max |
|-----------------------|--|------------------------------|-----------|--------|-----|
| <i>C</i>              | Control Group, Do nothing  | .2472                        | .4436     | 0      | 1   |
| <i>To</i>             | Treatment Group 0, Random promotions                               | .2434                        | .4291     | 0      | 1   |
| <i>T1</i>             | Treatment Group 1, Location-based ads                              | .2517                        | .4546     | 0      | 1   |
| <i>T2</i>             | Treatment Group 2, Trajectory-based ads                            | .2577                        | .4656     | 0      | 1   |
| <i>Sunday</i>         | Whether the visit was on Sunday                                    | .1416                        | .3681     | 0      | 1   |
| <i>Monday</i>         | Whether the visit was on Monday                                    | .1406                        | .3413     | 0      | 1   |
| <i>Tuesday</i>        | Whether the visit was on Tuesday                                   | .1433                        | .3399     | 0      | 1   |
| <i>Wednesday</i>      | Whether the visit was on Wednesday                                 | .1418                        | .3405     | 0      | 1   |
| <i>Thursday</i>       | Whether the visit was on Thursday                                  | .1418                        | .3404     | 0      | 1   |
| <i>Friday</i>         | Whether the visit was on Friday                                    | .1431                        | .3397     | 0      | 1   |
| <i>Saturday</i>       | Whether the visit was on Saturday                                  | .1477                        | .3754     | 0      | 1   |
| <i>Morning</i>        | Whether the visit was in the morning                               | .3217                        | .4013     | 0      | 1   |
| <i>Afternoon</i>      | Whether the visit was in the afternoon                             | .3407                        | .4900     | 0      | 1   |
| <i>Evening</i>        | Whether the visit was in the evening                               | .3376                        | .4762     | 0      | 1   |
| <i>TimeElapse</i>     | Time elapse between receiving a coupon and redeeming it            | 11.0285                      | 9.3735    | 0      | 36  |
| <i>Male</i>           | Whether the customer is male customer                              | .3475                        | .4762     | 0      | 1   |
| <i>Age</i>            | Age of the customer  | 37.9407                      | 10.8149   | 18     | 64  |
| <i>Income</i>         | Monthly Income (1000RMB)   | 16.9538                      | 8.0364    | 3      | 33  |
| <i>FirstTime</i>      | Whether the customer is first-time visitor                         | .0227                        | .1488     | 0      | 1   |
| <i>Redeem</i>         | Whether the customer redeemed the coupon                           | 0.2484                       | 0.4973    | 0      | 1   |
| <i>FutureRedeem</i>   | Whether the customer is willing to redeem the coupon in the future | 0.3451                       | 0.4754    | 0      | 1   |
| <i>TimeSpentStore</i> | Total time spent in focal advertising store (min)                  | 16.2987                      | 19.0283   | 0      | 51  |
| <i>TimeSpentMall</i>  | Total time spent in the mall (min)                                 | 60.7137                      | 36.2377   | 9      | 273 |
| <i>Satisfaction</i>   | Satisfaction rate of customer                                      | 2.8813                       | 1.7551    | 0      | 5   |
| Total # Observations: |  |                              |           | 83,370 |     |
| Time Period:          |  | 6/9/2014-6/22/2014 (14 days) |           |        |     |

**Table 1. Definitions and Summary Statistics of Variables**

On each day, we randomly assigned approximately 6,000 unique consumers who visited the shopping mall to one of the above four groups. To account for potential daily variation in a week, we conducted the same experiment for 14 consecutive days over two weeks from June 9, 2014, to June 22, 2014. Our experiment results are based on 83,370 unique user responses for a 14-day period. For a better understanding of our data, we present the definitions and summary statistics of all variables in Table 1.

## Main Results

In this section, we discuss our experimental results based on different levels of econometric analyses. We demonstrate first our results from group-level analyses on the mean treatment effect. Then, we discuss our findings from individual-level analyses on the distribution of the treatment effect.

### Group-level Analyses and Findings

First, we conduct group-level analyses. We compare daily group means (based on 14-day average) on consumer coupon-redemption rate, time elapsed until redemption, money and time spent in store, total money and time spent in the mall, satisfaction rate, and future willingness-to-redeem rate. To examine the statistical significance in the difference in-group means, we conduct a one-way ANOVA test. The results from group-level analyses are shown in Table 2a.

| Group                    | Redeem Rate | Future Redeem Rate | Spending in Focal Store (\$) | Spending in Mall (\$) | Time Elapse | Satisfaction Rate | Time in Store (min) | Time in Mall (min) |
|--------------------------|-------------|--------------------|------------------------------|-----------------------|-------------|-------------------|---------------------|--------------------|
| C                        | --          | --                 | --                           | 84.98                 | --          | 2.6               | --                  | 46.75              |
| To                       | 16%         | 21%                | 23.50                        | 88.19                 | 16.43       | 2.1               | 28.19               | 56.72              |
| T1                       | 23%         | 34%                | 41.25                        | 166.87                | 12.83       | 3.2               | 13.24               | 63.85              |
| T2                       | 31%         | 56%                | 56.78                        | 193.06                | 4.55        | 4.3               | 9.82                | 71.98              |
| <b>P&lt;0.05 (ANOVA)</b> |             |                    |                              |                       |             |                   |                     |                    |

**Table 2a. Group-Level Comparisons (Daily Mean, 14-Day Period)**

We find that on average, the new trajectory-based mobile advertising can lead to a statistically significant increase in the coupon-redemption rate, higher overall spending in the shopping mall, and higher overall satisfaction rate from the customers, compared with all the baseline strategies. In particular, we find that on average, the mobile trajectory-based advertising strategy leads to a 34.78% increase in the coupon-redemption rate when compared with the static location-based advertising strategy, and a 93.75% increase when compared with the random advertising strategy.

Interestingly, we also find that the new advertising strategy can lead to significantly lower amount of time customers spend in the focal advertising store (9.82m vs. 13.24m/28.19m), but higher revenues (\$56.78 vs. \$41.25/\$23.50). This finding indicates that mobile trajectory-based advertising can help make customers' shopping experience more efficient. We also note that based on group mean-level comparison, random advertising strategies on average perform the worst. Such strategies can lead to lower customer satisfaction rates due to the potential annoying effect from the improper ads.

| Group / Redeem Rate      | Age (20-30) | Age (30-40) | Age (40-50+) | Income <sup>1</sup> (2k-5k) | Income (6k-10k) | Income (11k-50k) |
|--------------------------|-------------|-------------|--------------|-----------------------------|-----------------|------------------|
| C                        | --          | --          | --           | --                          | --              | --               |
| To                       | 25%         | 17%         | 6%           | 34%                         | 15%             | 2%               |
| T1                       | 24%         | 20%         | 10%          | 33%                         | 23%             | 9%               |
| T2                       | 21%         | 23%         | 15%          | 32%                         | 27%             | 36%              |
| <b>P&lt;0.05 (ANOVA)</b> |             |             |              |                             |                 |                  |

**Table 2b. Demographic Subgroup-Level Redemption Rate Comparisons (Daily Mean, 14-Day Period)**

Furthermore, to understand how the treatment effect may vary across different demographic subgroups, we compare the mean treatment effects by breaking down the overall subject population into different demographic subgroups. The results from different demographic subgroups are shown in Table 2b. First, we find that on average, the younger age group (i.e., 20-30) is more responsive to mobile advertising, whereas the older age group (i.e., 40-50+) is less responsive, regardless of the type of mobile ads. Second,

<sup>1</sup> Income is measured in RMB at monthly level.

our results show that on average, customers with lower monthly income (i.e., 2k-5k) are more active in responding to mobile promotions. However, they are not sensitive to the type of mobile ads. This finding is reasonable because low-income customers are often highly price sensitive; hence, any mobile ads that offer price promotions would attract them. On the contrary, interestingly, customers with higher monthly income (i.e., 11k-50k) on average are not as responsive to random ads (2%) or regular static location-based mobile ads (9%). However, they are highly attracted by mobile trajectory-based ads (36%). Our findings indicate the potential of mobile trajectory-based advertising in attracting high-end customers to achieve better customer lifetime value. In particular, these high-income customers are usually the “challenging type” for mobile advertising. They are likely to be extremely sensitive to the quality of mobile targeting. They will not respond to a mobile ad just because it offers a lower price, unless it is carefully designed and is a good fit for their personal preferences.

### Individual-level Analyses and Findings

Beyond the group-level analyses, our unique data set acquired from the field experiment also allows us to conduct individual-level analyses on the effect of mobile advertising on consumer mobile coupon redemption and purchase behavior. In particular, we observe individual-level consumer characteristics, mobile advertising response, and purchase behavior. Such data help us further examine the distribution of the treatment effect of mobile advertising, as well as its interaction with consumer heterogeneity.

We aim to examine the effect of different mobile advertising strategies (i.e., random, static location-based, and trajectory-based) on the likelihood of consumer mobile coupon-redemption behavior. To do so, we apply a Logit model at the individual consumer level and model the consumer coupon-redemption rate as a function of consumer characteristics and different mobile advertising strategies. To account for the potential variation in the mobile advertising effects induced by the consumer heterogeneity, we include interaction effects between consumer characteristics and different mobile advertising strategies.

More specifically, we model the utility for consumer  $i$  to redeem a mobile coupon as follows:

$$U_i = \bar{U}_i + \varepsilon_i = \alpha + \beta X_i + \gamma T_i + \lambda D_i + \delta T_i \times X_i + \varphi T_i \times D_i + \varepsilon_i, \quad \varepsilon_i \sim i.i.d., EV(0,1), \quad [3]$$

where  $X_i$  is an individual-specific vector representing characteristics of consumer  $i$  (e.g., age, gender, income level, credit card type, first-time visitor, shop alone, phone type, etc.),  $T_i$  is an individual-specific vector containing three binary indicators for three treatment groups ( $T_0$  – Random,  $T_1$  – Location,  $T_2$  – Trajectory), and  $D_i$  represents other individual-specific control variables for consumer  $i$  (e.g., day of the week, time of day, coupon type, advertising store category).  $\varepsilon_i$  is an individual stochastic error term that captures any randomness during consumer  $i$ 's decision process. We assume the error term follows the type I extreme value distribution. Hence, the probability of consumer  $i$  redeeming a mobile coupon is

$$Pr_i(Redeem = 1) = \frac{\exp(\bar{U}_i)}{1 + \exp(\bar{U}_i)}. \quad [4]$$

We provide the estimation results from the Logit model in Table 3. First, we find that on average, mobile trajectory-based advertising outperforms all the baseline advertising strategies at the individual consumer level. In particular, mobile trajectory-based ads show the highest significant and positive effect on customer coupon-redemption rates ( $\gamma_2=2.3381$ ), compared to the corresponding effects by location-based ads ( $\gamma_1=1.0068$ ) and by random ads (i.e., baseline). Note that when studying the effect of mobile advertising on coupon redemption rate, we used the “random coupon” group ( $T_1$ ) as the baseline because participants in control group (C) did not receive any mobile coupon by experimental design. Second, we find significant differences in coupon-redemption behavior at different times. On average, customers are more likely to redeem a mobile coupon during weekends than during weekdays, and they are more likely to respond to a mobile coupon in the afternoon and evening than in the morning.

Interestingly, our model with interaction effects demonstrates significant heterogeneity in the treatment effect. In particular, we find that although on average, mobile trajectory-based ads perform the best in increasing coupon-redemption responses, they become 85.78% less effective during the weekends (i.e., interaction effect between  $T_2$  and Weekend is -2.0057, whereas the average  $T_2$  effect is 2.3381). On the other hand, a random advertising strategy becomes much more effective during the weekends and for first-time visitors. In particular, random advertising outperforms all the advertising strategies during the

weekends, showing a significant and positive interaction effect (i.e., 1.8696) on coupon-redemption probability. Meanwhile, we also find a similar trend for first-time visitors. These findings are intriguing. They seem to suggest that weekend customers and first-time visitors tend to react to the random ads more frequently compared to the targeted ads.

| Variables                             | Coefficient (Std. Err.) |
|---------------------------------------|-------------------------|
| <i>To-Random</i>                      | --                      |
| <i>T1-Location</i>                    | 1.0068 (.2851) ***      |
| <i>T2-Trajectory</i>                  | 2.3381 (.2747) ***      |
| <i>Weekend</i>                        | 0.3612 (.0129) ***      |
| <i>Afternoon</i>                      | 0.4326 (.0165) ***      |
| <i>Evening</i>                        | 0.3010 (.1069) **       |
| <i>Male</i>                           | -0.0027 (.0264)         |
| <i>Age</i>                            | -1.1987 (1.0025)        |
| <i>Age<sup>2</sup></i>                | 0.3968 (2.4472)         |
| <i>Income</i>                         | 0.5758 (.4521)          |
| <i>Income<sup>2</sup></i>             | -0.1882 (.2168)         |
| <i>To-Random × Weekend</i>            | 1.8696 (.0305) ***      |
| <i>T1-Location × Weekend</i>          | -0.0091 (.0259)         |
| <i>T2-Trajectory × Weekend</i>        | -2.0057 (.0286) ***     |
| <i>To-Random × Male</i>               | 0.0113 (.0301)          |
| <i>T1-Location × Male</i>             | -0.0237 (.0254)         |
| <i>T2-Trajectory × Male</i>           | 0.0092 (.0357)          |
| <i>To-Random × Income</i>             | -0.0356 (.0284)         |
| <i>T1-Location × Income</i>           | 0.0026 (.0242)          |
| <i>T2-Trajectory × Income</i>         | 0.0170 (.4693)          |
| <i>To-Random × FirstTimeVisit</i>     | 1.1648 (.0940) *        |
| <i>T1-Location × FirstTimeVisit</i>   | -0.0702 (.0841)         |
| <i>T2-Trajectory × FirstTimeVisit</i> | -0.0168 (.0894)         |
| <i>First Time Visitor</i>             | Yes                     |
| <i>Credit Type</i>                    | Yes                     |
| <i>Phone Type</i>                     | Yes                     |
| <i>Shopping Context</i>               | Yes                     |
| <i>Advertising Store Category</i>     | Yes                     |
| <i>Coupon Type</i>                    | Yes                     |
| *** P<0.001, ** P<0.05, * P<0.1       |                         |

**Table 3. Model Estimation Results**

One potential reason is that customers who visit the shopping mall during the weekends tend to be “explorers,” who often walk around the mall randomly and may not have a concrete purchase plan beforehand. As a result, they may be more open to random ads, which can help them with their explorative and variety-seeking habits in the shopping mall. Moreover, previous marketing and psychology literature has shown that when customers are in the unplanned purchasing stage, they are more likely to engage in impulse purchases (e.g., Stern 1962, Engel and Blackwell 1982), that a positive relationship exists between variety-seeking behavior and impulse buying (e.g., Sharma et al. 2010), and

that a consumer's propensity to purchase on impulse receives a further impetus when she sees a random item on sale (Ramaswamy and Namakumar, 2009). Therefore, being exposed to a random promotion can significantly increase the likelihood of impulse purchases for customers who are in an unplanned shopping and explorative stage. On the other hand, targeted ads tend to directly lead the customers to the focal advertising stores. As a result, they may constrict the scope and physical range of exploration, and reduce the probability of impulse purchases, especially for customers who are in an unplanned shopping stage. Previous literature has found that shorter in-store travel length has a negative effect on consumers' in-store impulse buying behavior (Hui 2013a, 2013b), and that historical-behavior-based targeting may lead to less variety-seeking behavior from consumers (e.g., Fleder and Honsanagar 2009). Our findings are in line with these previous studies indicating that marketers need to carefully design the targeted campaign based on the shopping context and mental stage of customers.

## Conclusion and Future Work

The proliferation of mobile and sensor technologies make it possible to leap beyond the real-time snapshot of the static location and context information about consumers. In this study, we propose a novel mobile advertising strategy that infers consumer preferences by leveraging full information on consumers' offline moving trajectories from four different mobility dimensions. To measure its effectiveness, we conduct a large-scale randomized field experiment in a major shopping mall in Asia based on 83,370 unique user responses for a 14-day period.

We find that by extracting and incorporating the overall offline behavioral trajectory of each individual consumer, we are able to significantly improve the performance of mobile advertising. In particular, our results show that on average, mobile trajectory-based advertising can help businesses achieve higher coupon-response rate compared to the existing baseline location-based advertising strategies. Meanwhile, our study also reveals significant heterogeneity in the mobile advertising effect. Targeted mobile trajectory-based ads may not always perform the best. They may reduce the amount of impulse purchases from customers, especially during the weekends. Therefore, businesses must understand the heterogeneity in the effect of different mobile ads. Marketers should carefully design mobile advertising strategies based on the business contexts.

On a broader note, our paper is the first to analyze the digital trace of individual offline behavior and how it can be linked to better understand individual preferences and decision making. Our work can be viewed as a first step to study the digitalization of human offline behavior at a large-scale and granular level. We demonstrate the value of leveraging mobile and sensor technologies to digitalize, measure, understand, and predict individual behavioral trajectory in the physical environment to improve user digital experiences and business marketing strategies.

Our paper has some limitations, which can serve as fruitful areas for future research. First, because of technical limitations of our GPS tracking system, we could recruit only customers who were interested in accessing Wi-Fi, which resulted in approximately 80% of the customers in the shopping mall. However, this number could potentially improve in the future with a tracking system based on more advanced sensor technologies (e.g., wearable devices). Second, in the current analysis, we were able to control for various observed individual characteristics, such as age, income, gender, and so on. However, individual-level unobserved heterogeneity might still exist. In the future, we can incorporate random coefficient models to better account for such individual-specific unobservables. Third, currently our recommendations are based on similarity between customers. In the future, we could potentially experiment with alternative recommendation strategies, for example, recommendation based on dissimilarity between customers.

Another interesting future direction is to explore whether and how consumers' trajectories may diverge from their original predicted patterns after receiving the intervention. In particular, we would like to compare the predicted trajectory patterns with the observed resulting patterns after the intervention to study the changes in individual behavior. Interestingly, our experiments show that the potential divergence in consumer trajectory after intervention actually results in better outcomes in general. Note that in this study we are able to predict which sites will likely be on the future predicted trajectory path, but it would be even more interesting if we can predict the full future trajectory path/sequence. This is important because trajectory divergence after intervention can be extremely informative in explaining why this new mobile strategy works and why it does not work under certain circumstance. Our current

experimental data allow us to provide some theoretical explanations and empirical evidence for support (e.g., weekend impulse buyers). However, we cannot perfectly verify them using our current data unless we survey people in details why they behave in such a way after receiving the mobile coupon. Individual-level information on trajectory divergence after intervention can help us better understand the distribution and heterogeneity in the treatment effect.

Finally, due to privacy policy of the shopping mall, we could not identify repeated customers if they visit the shopping mall multiple times during the 14-day experimental period. In this study, we treated each individual trajectory as a unique customer. In the future, it would be interesting if we can identify return customers or the same customers who visit different shopping malls, to study the individual long-term learning behavior facilitated by the mobile advertising interventions.

## References

- Adomavicius, G., B. Mobasher, F. Ricci, A. Tuzhilin. 2011. Context-Aware Recommender Systems. *AI Magazine*, 35(4).
- Andrews, M., X. Luo, Z. Fang, and A. Ghose. 2015. Mobile Ad Effectiveness: Hyper-Contextual Targeting with Crowdedness forthcoming, *Marketing Science*. Forthcoming.
- Bart Y, Stephen A, Sarvary M (2014) Which products are best suited to mobile advertising? A field study of mobile display advertising effects on consumer attitudes and intentions. *J. Marketing Res.* 51(3):270-285.
- Brohee, Sylvain and Jacques Van Helden. 2006. Evaluation of Clustering Algorithms for Proteinprotein Interaction Networks. *BMC Bioinformatics*, 7 (1), 488.
- Cuturi, M.. Fast global alignment kernels. In *ICML'11*.
- Danaher, P., M. Smith, K. Ranasinghe, and T. Danaher. 2015 Where, When and How Long: Factors that Influence the Redemption of Mobile Phone Coupons. *Journal of Marketing Research*.
- Eisen, Michael B., Paul T. Spellman, Patrick O. Brown, and David Botstein. 1998. Cluster Analysis and Display of Genome-wide Expression Patterns. *Proceedings of the National Academy of Sciences*, 95 (25), 14863–14868.
- Ericsson 2014. Ericsson Mobility Report. Nov. 2014.
- eMarketer. 2014. Worldwide Mobile Phone Users: H1 2014 Forecast and Comparative Estimates. <http://www.emarketer.com/Article/Smartphone-Users-Worldwide-Will-Total-175-Billion-2014/1010536>
- Engel, J. and Blackwell, R. 1982. *Consumer Behavior*. Chicago: Dryden Press.
- Esslimani, I., A. Brun, and A. Boyer. From social networks to behavioral networks in recommender systems. In *ASONAM'09*, pages 143–148, 2009.
- Fleder, D. and K Hosanagar. 2009. Blockbuster culture's next rise or fall: The impact of recommender systems on sales diversity. *Management science*, 55(5), 697-712.
- Fong, N. M., Z. Fang, X. Luo. 2015. Real-Time Mobile Geo-Conquesting Promotions. *Journal of Marketing Research*. Forthcoming.
- Gaffney, S. and P. Smyth. Trajectory clustering with mixtures of regression models. In *SIGKDD'99*.
- Ghose, A., Ipeirotis, P. and Li, B. 2012. Designing Ranking Systems for Hotels on Travel Search Engines by Mining User-Generated and Crowdsourced Content. *Marketing Science*. 31(3), 493-520.
- Ghose, A., Ipeirotis, P. and Li, B. 2014. Examining the Impact of Ranking on Consumer Behavior and Search Engine Revenue. *Management Science*. 60(7).
- Ghose, A., S. Han and S. Park. 2013. Analyzing the Interdependence Between Web and Mobile Advertising: A Randomized Field Experiment. Working paper, NYU.
- Gorgoglione, M., U. Panniello, and A. Tuzhilin. The effect of contextaware recommendations on customer purchasing behavior and trust. In *RecSys'11*, pages 85–92, 2011.
- Guo, X., Eddie C. L. Chan, Ce Liu, Kaishun Wu, Siyuan Liu, Lionel M. Ni. ShopProfiler: Profiling Shops with Crowdsourcing Data. The 33rd Annual IEEE INFOCOM International Conference on Computer Communications. 2014. Toronto, Canada.
- Gu, Bin. 2012. Quantifying the Dynamic Sales Impact of Location-based Mobile Advertising Technologies. Working Paper, Arizona State University.
- Huang, J., H. Sun, J. Han, H. Deng, Y. Sun, and Y. Liu. 2010. Shrink: a structural clustering algorithm for detecting hierarchical communities in networks. In *CIKM'10*.

- Hui, Sam, Jeffrey Inman, Yanliu Huang, and Jacob Suher. 2013a. Estimating the Effect of Travel Distance on Unplanned Spending: Applications to Mobile Promotion Strategies. *Journal of Marketing*, 77 (March), 1-16.
- Hui, Sam, Yanliu Huang, Jacob Suher, and Jeffrey Inman. 2013b. Deconstructing the 'First Moment of Truth': Understanding Unplanned Consideration and Purchase Conversion Using In-Store Video Tracking. *Journal of Marketing Research*, 50 (4), 445-462.
- Lee, J.-G., J. Han, and K.-Y. Whang. Trajectory clustering: a partition-and-group framework. In SIGMOD'07.
- Leonard, H. 2013. There Will Soon Be One Smartphone for Every Five People in the World. <http://www.businessinsider.com/15-billion-smartphones-in-the-world-2013-2>.
- Leskovec, J., K. J. Lang, and M. W. Mahoney. Empirical comparison of algorithms for network community detection. In WWW'10.
- Liu S, Wang S., Jayarajah K., Misra A., Krishnan R.. 2013. TODMIS: Mining Communities from Trajectories. ACM International Conference on Information and Knowledge Management, CIKM'13.
- Liu, Y., Y. Zhao, L. Chen, J. Pei, and J. Han. Mining frequent trajectory patterns for activity monitoring using radio frequency tag arrays. *IEEE Trans. Parallel Distrib. Syst.*, pages 2138–2149, 2012.
- Luo, X., M. Andrews, Z. Fang, and Z. Phang. 2014. Mobile Targeting. *Management Science*. 60 (7), July, 1738-56.
- MacQueen, James. 1967. Some Methods for Classification and Analysis of Multivariate Observations. *Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability*, 1 (14), 281- 297
- Manning, C. D.; Raghavan, P.; Schütze, H. 2008. Scoring, term weighting, and the vector space model. *Introduction to Information Retrieval*. p. 100
- Molitor D, Reichhart P, Spann M, and A. Ghose 2014. Measuring the effectiveness of location-based advertising: A randomized field experiment. Working paper, NYU.
- Nanni, M. and D. Pedreschi. Time-focused clustering of trajectories of moving objects. *J. Intell. Inf. Syst.*, 2006.
- Newman, M.. Detecting community structure in networks. *EPJ B*, 2004.
- Ng, R. T. and J. Han. Clarans: A method for clustering objects for spatial data mining. *IEEE TKDE*, 2002.
- Nowak, M. and C. Nass. Effects of behavior monitoring and perceived system benefit in online recommender systems. In CHI'12, pages 2243–2246, 2012.
- Pelekis, N., I. Kopanakis, E. E. Kotsifakos, E. Frentzos, and Y. Theodoridis. Clustering uncertain trajectories. *Knowl. Inf. Syst.*, 2011.
- Ramaswamy & Namakumar. 2009. *Marketing Management* (4th ed.). New Delhi: McMillan Publishing.
- Satuluri, Venu, Srinivasan Parthasarathy, and Duygu Ucar. 2010. Markov Clustering of Protein Interaction Networks with Improved Balance and Scalability. *Proceedings of the First ACM International Conference on Bioinformatics and Computational Biology*, 247-256.
- Shankar, Venkatesh and Sridhar Balasubramanian. 2009. Mobile Marketing: A Synthesis and Prognosis. *Journal of Interactive Marketing*, 23 (2), 118-129.
- Shankar, Venkatesh, Alladi Venkatesh, Charles Hofacker, and Prasad Naik. 2010. Mobile Marketing in the Retailing Environment: Current Insights and Future Research Avenues. *Journal of Interactive Marketing*, 24 (2), 111-120.
- Sharma, P., Sivakumaran, B., & Marshall, R. 2010. Impulse buying and Variety seeking: A trait-correlates perspective. *Journal of Business Research*, 63, 276-83.
- Shi, C, P. S. Yu, Y. Cai, Z. Yan, and B. Wu. 2011. On selection of objective functions in multi-objective community detection. In CIKM'11.
- Spiekermann S, Rothensee M, Klafft M. 2011. Street marketing: How proximity and context drive coupon redemption. *J. Consumer Marketing* 28(4):280–289.
- Stern, H. 1962. The Significance of Impulse Buying Today. *Journal of Marketing*, April, 59-62.
- Paco Underhill. 1999. *Why We Buy: The Science of Shopping*.
- van Dongen, Stijn, Ceil Abreu-Goodger. 2012. Using MCL to extract clusters from networks. *Bacterial Molecular Networks*, 281-295.
- Xiang, T. and S. Gong. Beyond tracking: Modelling activity and understanding behaviour. *IJCV*, 2006.
- Zhu, X., J. Guo, X. Cheng, and Y. Lan. More than relevance: high utility query recommendation by mining users' search behaviors. In CIKM'12, pages 1814–1818, 2012.