Mobile Trajectory-Based Advertising: Evidence from A Large-Scale Randomized Field Experiment

Beibei Li
Carnegie Mellon University

(Joint work with Anindya Ghose and Siyuan Liu)

July 2015
Mobile technology and location-based services (LBS) are changing how we live and shop today:

- 91% own a mobile phone (PewResearch 2014);
- Consumers increasingly use mobile devices to locate and buy products, pushing revenues to $100B by 2017 (eMarketer 2013);

Mobile advertising, especially location-based, has great potential:

- 47% of users would provide their location to receive relevant offers and discounts (mBlox 2013); 76% of users agree that location sharing provides more meaningful content (Mobile Behavior Report 2014);
- Total value of real-time mobile location-based advertising will grow from $1.66B in 2013 to $14.8B in 2018 (Berg Insight 2013).
Motivation

- Mobile location-based advertising:
  - Geo-targeting, geo-fencing, geo-conquesting, …
- Location & Context $\rightarrow$ improve user digital experience:
  - Real-time geographical location (Ghose and Han 2011, Molitor et al. 2014).
  - Consumers’ current context (i.e., crowdedness, weather) (Andrews et al. 2014, Ghose et al. 2014, Luo et al. 2015).

Static Snapshot vs. Mobile Trajectory

Can we learn more about user preferences by better utilizing the information enabled by mobile technologies?
Research Questions

- Can we extract consumer preferences from large-scale, fine-grained individual mobile trajectory data?

- Can we design a new mobile advertising strategy that leverages not only static location/context information, but also consumer’s mobile trajectory?

- How can we measure the causal impact of the mobile trajectory-based advertising on consumer shopping behavior and business revenues?
Research Summary

- **Methodologies:**
  - Design a new mobile trajectory-based advertising strategy, by using statistical and machine learning methods to extract user preferences from different mobility dimensions.
  - Conduct a large-scale randomized field experiment in a large shopping mall in Beijing; 83,370 unique user responses for a 14-day period.

- **Main Findings:**
  - Higher redemption rate, satisfaction rate, and faster redemption action (compared to several benchmarks);
  - Always benefit the focal advertising store, but less effective in the overall revenues of the shopping mall during the weekend (hurt “impulse buy”);
  - Especially effective in attracting high income consumers (potential in approaching high-end customers to achieve better customer lifetime value).
Related Work

- **Mobile and Location-Based Advertising:**
  - Geo-Fencing: Molitor et al. (2014), Luo et al. (2013), Andrews et al. (2014)
  - Geo-Conquesting: Fong et al. (2014)
  - In-store Travel and Unplanned Purchase: Hui et al. (2013a, 2013b)

- **Community Detection & Trajectory-Based Clustering:**
  - Community detection based on social graph: Leskovec et al. (WWW 2010), Huang et al. (CIKM 2010), Shi et al. (CIKM 2011), Lin et al (WWW 2012)
  - Community detection based on individual trajectories: Vlachos et al. (ICDE 2002), Lee et al. (SIGMOD 2007), Ge et al. (CIKM 2010), Giannotti et al. (VLDB 2011), Tang et al. (ICDE 2012)
  - Multidimensional trajectories: (LiveLabs) Liu et al. (CIKM 2013, TKDE 2014), Sen et al. “GruMon” (Sensys 2014).
Agenda

• Modeling Consumer Similarity & Trajectory-Based Clustering
• A New Mobile Recommendation Approach
• Randomized Field Experiment
• Analyses and Results
• Conclusions
Modeling Consumer Similarity from Mobile Trajectories
“Great Minds Move alike.”

- Define a “community” as a set of similar customers with similar patterns of mobile trajectories.
  Trajectories within a community indicate similar behaviors.

- Define pairwise “similarity” as a function of different aspects of individual mobile trajectory.
  e.g., visit similar stores, visit at similar time (weekends vs. weekdays, morning vs. afternoon), similar shopping speed (explorers, raiders), etc.

- Mine communities using graph-based clustering (e.g., dense sub-graph detection).

Key: Measure similarity?

Undirected graph:
- Node – consumer;
- Weight of Edge – similarity between two consumers.
Measure Similarity (Liu et al. 2013)

- Assume two customers $i$, $i'$. We measure similarity $S(i,i')$ as a function of four different aspects of trajectories:
  - **Temporal**: Start/End time stamps, Time and day indicators;
  - **Spatial**: Spatial alignments; Crowdedness (TF-IDF);
  - **Semantic**: Stationary distribution of individual’s visit probability to each site; Time spent at sites; Transition probability from site $a$ to site $b$; Time spent to transit from $a$ to $b$;
  - **Velocity**: Speed (normalized by travel length);

- The similarity $S(i, i')$ is a weighted combination of a set of similarities calculated from the above four sources:

$$S(i,i')= a_1 K_t + a_2 K_p + a_3 K_s + a_4 K_v$$

$K$-similarity score by using various similarity functions (cosine distance, kernels).

$a$ - Weight associated with each dimension
A New Mobile Recommendation Strategy

- Recommend to a consumer stores or products that are most frequently visited by similar consumers.

- We achieve this by using a similar approach to "collaborative filtering":

\[
\hat{R}(i, j) = \frac{\sum_{i' \leq i}^N R(i', j) S(i, i')}{\sum_{i' \leq i}^N S(i, i')}
\]

*Rating of consumer $i'$ on a store $j$*

*Similarity between consumers $i$ and $i'$*

*Rating of consumer $i$ on a store $j$*
Agenda

- Modeling Consumer Similarity & Trajectory-Based Clustering
- A New Mobile Recommendation Approach
- Randomized Field Experiment
- Analyses and Results
- Conclusions
Randomized Field Experimental Setting: A Large Shopping Mall in Beijing
Experimental Setting

• A major large shopping mall in Beijing:
  – 1.3 million square feet
  – 300+ stores
  – 100,000 visitors per day; 200,000 visitors per day during holidays
  – WiFi localization system
Individual Information Validation

• At the entrance of the shopping mall, if a consumer would like to enjoy free WiFi, the consumer is required to fill a Form A with demographic information (e.g., age, gender, income range, etc.).

• At each store, upon purchase sales will fill a Form B for each consumer (e.g., demographics, whether or not a targeted coupon is used, money spent, credit card type (gold, platinum, gift card)).

• Form A and Form B cross validation.
Experimental Design

- **Control Group (G0):** Do nothing
- **Treatment Group (G1):** Send random promotion messages
- **Treatment Group (G2):** Send promotion messages based on static real-time locations
- **Treatment Group (G3):** Send promotion messages based on our trajectory-based recommendation

- On each day, randomly assign ~6000 consumers to one of the 4 groups;
- 14 consecutive days, 83,370 unique user responses;
- Promotions involve 252 participating stores;
- Different types of coupons: e.g., “50% off” and “Buy one get one free”;
- G1 uses the exact same set of mobile promotions (format & price discount) as the ones used in G2 & G3, except randomly sent;
Follow-up Survey (4 Short Questions)

1. Do you find it helpful or not to use the mobile in-store recommendation?
   [1 (Not all all) to 5(Very Helpful)]

2. Did you follow the recommendation or not?
   [Yes/No]

3. Are you willing to follow the recommendation in future or not?
   [1(Not Interested) to 5(Definitely)]

4. How do you rate your satisfaction about the overall shopping experience today?
   [1 (Disappointed) to 5(Very Satisfied)]
Group-Level Analyses
Main Results (I)
Group Mean Comparisons (14-Day Average)

Trajectory-based Mobile Recommendation leads to
• Highest promotion response rate, fastest redemption action;
• Less time spent in the focal store, but more revenue;
• Overall more time spent in the mall;
• Higher overall satisfaction.

• Location > Random (redemption rate, spending, satisfaction rate);
• Random → lowest satisfaction rate.

Demographics Breakdown:
• Age → Negative effect, (50-65+) are least likely to respond;
• Trajectory → especially effective in attracting high income group.

Individual Heterogeneity & Interactions with Treatments?
Individual-Level Analyses
Individual-Level Analyses  
- Consumer Redemption Probability

We model the probability for consumer $i$ to redeem a mobile promotion as the following:

$$Pr_i(\text{Redeem} = 1) = \frac{\exp(U_i)}{1 + \exp(U_i)}$$

$$U_i = \alpha + \beta X_i + \gamma T_i + \lambda D_i + \delta T_i \times X_i + \phi T_i \times D_i + \epsilon_i, \quad \epsilon_i \sim i.i.d., \text{Type I EV}(0,1)$$

where $X$~ individual characteristics (age, income, gender, credit card type, first time visitor, shop alone, phone type);

$T$~ treatment group indicator (Random, Location, Trajectory);

$D$~ other control variables (day in a week, time of day, coupon type);

Run the logit model over 14 days 83,370 observations.
Main Results (III)
Individual Redemption Probability

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Variable</th>
<th>Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1(Random)</td>
<td>----</td>
<td>Weekend</td>
<td>0.3612 **</td>
</tr>
<tr>
<td>T2(Location)</td>
<td>1.0068 ***</td>
<td>Afternoon</td>
<td>0.4326 **</td>
</tr>
<tr>
<td>T3(Trajectory)</td>
<td>2.3381 ***</td>
<td>Evening</td>
<td>0.3010 *</td>
</tr>
<tr>
<td>T1 × Weekend</td>
<td>1.8696 **</td>
<td>T1 × FirstTimeVisit</td>
<td>1.1648 *</td>
</tr>
<tr>
<td>T2 × Weekend</td>
<td>-0.0091</td>
<td>Coupon Type</td>
<td>Yes</td>
</tr>
<tr>
<td>T3 × Weekend</td>
<td>-2.0057 **</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*** P<0.001, ** P<0.05, * P<0.1

- On average, Trajectory-based > Location-based > Random;
- Weekend > Weekday;
- Afternoon/Evening > Morning;
- Trajectory $\rightarrow$ less effective during weekend;
- Random $\rightarrow$ more effective during weekend, and for first time visitor. Impulse buyers and explorer (random ads help explore);

Coupon Redemption $=$? Revenues
Individual-Level Analyses  
– Consumer Spending in Focal Store and in the Mall

We model consumer $i$’s spending in the focal advertising store, and the overall spending in the mall as the following:

\[
S_{\text{store}} = \alpha_{\text{store}} + \beta_{\text{store}}X_i + \gamma_{\text{store}}T_i + \lambda_{\text{store}}D_i + \delta_{\text{store}}T_i \times X_i + \varphi_{\text{store}}T_i \times D_i + \epsilon_{\text{store}}.
\]

\[
S_{\text{mall}} = \alpha_{\text{mall}} + \beta_{\text{mall}}X_i + \gamma_{\text{mall}}T_i + \lambda_{\text{mall}}D_i + \delta_{\text{mall}}T_i \times X_i + \varphi_{\text{mall}}T_i \times D_i + \epsilon_{\text{mall}}.
\]

$\epsilon_{\text{store}}$, $\epsilon_{\text{mall}} \sim i.i.d., N(0,1)$. 
Main Results (IV)
Individual Spending in the Focal Store

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Variable</th>
<th>Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1(Random)</td>
<td>2.2608 **</td>
<td>Weekend</td>
<td>1.1870 *</td>
</tr>
<tr>
<td>T2(Location)</td>
<td>4.6343 ***</td>
<td>Male</td>
<td>-0.7663 *</td>
</tr>
<tr>
<td>T3(Trajectory)</td>
<td>13.0159 ***</td>
<td>Age</td>
<td>-0.4680 *</td>
</tr>
<tr>
<td>Income</td>
<td>-0.1749</td>
<td>Age ^2</td>
<td>0.0627 *</td>
</tr>
<tr>
<td>Income ^2</td>
<td>0.0791</td>
<td>Coupon Type</td>
<td>Yes</td>
</tr>
</tbody>
</table>

*** P<0.001, ** P<0.05, * P<0.1

- On average, Trajectory-based > Location-based > Random;
- Weekend > Weekday;
- Female > Male
- Age $\rightarrow$ negative effect, diminishing;
- No significant interaction effects.

$\rightarrow$ No significant heterogeneity in the **direct effect** of mobile advertising on **focal store spending**. Always Benefit!
# Main Results (V)
## Individual Overall Spending in the Mall

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Variable</th>
<th>Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1(Random)</td>
<td>0.7527</td>
<td>Weekend</td>
<td>0.5212 ***</td>
</tr>
<tr>
<td>T2(Location)</td>
<td>1.4912 ***</td>
<td>Male</td>
<td>-0.1031 **</td>
</tr>
<tr>
<td>T3(Trajectory)</td>
<td>2.0003 ***</td>
<td>Age</td>
<td>-0.0222 ***</td>
</tr>
<tr>
<td>Income</td>
<td>-1.7117</td>
<td>Age ^2</td>
<td>0.0045 **</td>
</tr>
<tr>
<td>Income ^2</td>
<td>0.3408</td>
<td>Coupon Type</td>
<td>Yes</td>
</tr>
</tbody>
</table>

On average, consistent with focal store spending:

- Trajectory > Location;
- Weekend > Weekday;
- Female > Male;
- Age → negative effect, diminishing.
## Main Results (V)
Individual Overall Spending in the Mall

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Variable</th>
<th>Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1 × Weekend</td>
<td>0.8995 ***</td>
<td>T1 × Income</td>
<td>0.0908 *</td>
</tr>
<tr>
<td>T2 × Weekend</td>
<td>-0.2502 ***</td>
<td>T2 × Income</td>
<td>0.1328 **</td>
</tr>
<tr>
<td>T3 × Weekend</td>
<td>-0.9923 ***</td>
<td>T3 × Income</td>
<td>0.1506 ***</td>
</tr>
<tr>
<td>T1 × Male</td>
<td>0.0617 **</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>T2 × Male</td>
<td>0.0831</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>T3 × Male</td>
<td>0.1020 **</td>
<td>*** P&lt;0.001,</td>
<td>** P&lt;0.05,</td>
</tr>
<tr>
<td></td>
<td></td>
<td>* P&lt;0.1</td>
<td></td>
</tr>
</tbody>
</table>

Interestingly, significant interaction effects:
- **Trajectory** → Less effective during weekend;
- **Random** → More effective during weekends;
- **Male and higher income** → More sensitive to good mobile advertising;

→ Significant heterogeneity in the **indirect effect** of mobile advertising on overall spending in the mall.
Store-Level Heterogeneity?

- Different store may have different products (category, price, quality, brand affinity, etc.);
- e.g., Dinning, Fashion, Supermarket, ...

- Randomization of store coupons in experimental design can alleviate such concern to some extent, but potential store-level unobservable may still exist.

  e.g., Trajectory/Location-based ads are more effective for fashion stores, compared to random ads; However, effects on restaurants and supermarket are not significantly different.
Store-Level Fixed-Effect Analyses – Store Revenues

We model the overall revenues for store $j$ on day $t$ as a function of

- # different types of promotions sent by the store,
- weekend dummy,
- interaction effects,
- store-level fixed effect:

$$R_{jt} = \theta_0 + \theta_1 N_{T1}^{jt} + \theta_2 N_{T2}^{jt} + \theta_3 N_{T3}^{jt} + \theta_4 Weekend_{jt}$$
$$+ \theta_5 N_{T1}^{jt} \times Weekend_{jt} + \theta_6 N_{T2}^{jt} \times Weekend_{jt} + \theta_7 N_{T3}^{jt} \times Weekend_{jt} + \zeta_j + \epsilon_{jt},$$

We run this fixed effect model over a panel data with 14 days $\times$ 252 stores $= 3,024$ observations;

For robustness test, we also run a random effect model.
Additional Robustness Checks

- Separate analysis for each day;
- Separate analysis for weekdays vs. weekends;
- Look into in-store travel distance (Hui et al. 2013);
Summary of Main Findings

Trajectory-based Advertising:

- **On average, more effective.** Higher redemption rate, higher satisfaction rate, and faster redemption action (compared to other benchmarks);

- Especially effective in attracting **high income** consumers (potential in approaching high-end customers to achieve better customer lifetime value).

  - **Direct effect on focal advertising store** ➔ Always benefit!

  - **Indirect effect on overall mall revenues** ➔ Heterogeneity! Less effective for **weekend and first-time consumes** (may reduce exploration and impulse buy);
Major Contributions

- Extract consumer preferences from large-scale, fine-grained mobile trajectory data using statistical and machine learning methods;

- Design a new mobile trajectory-based advertising strategy;

- Evaluate using randomized field experiments, survey and econometrics to examine the causal effect.

- Digitalization of Individual Offline Behavior → user preference, business decision making
On-Going Interests

- Data Science + Social Science $\rightarrow$ Individual Decision Making
- Econometric Modeling, Bayesian Modeling, Machine Learning, Randomized Experiments
- On-going work: mobile recommendation and offline social dynamics.
Q & A
Thank you!