Does Emotion Influence the Use of Auto-suggest during Smartphone Typing?

Surjya Ghosh, Kaustubh Hiware, Niloy Ganguly, Bivas Mitra
Indian Institute of Technology Kharagpur

Pradipta De
Georgia Southern University
Background

- Text-entry in small touch-based devices — relatively inconvenient due to limited space

- To facilitate, different techniques like *auto-complete, auto-suggest* are used
Background

• However,
  – reserves fixed space in keyboard layout
  – increase cognitive load
    • to read, parse and select correct word [CHI 2016]

Opportunity to improve the layout by making these techniques adaptive
Objective

• Multiple factors can guide auto-suggest usage
  – typing volume
  – application type
  – user emotion
    • as it influences typing behaviour [MobileHCI 2017, ACII 2017]

• Can auto-suggestion usage be determined based on human emotion?
• Auto-suggest usage
  – tracing user typing and providing suggestions
  – labelling auto-suggest usage
    • **accepted** \(\rightarrow\) if at least one suggestion is used in a session
    • **skipped** \(\rightarrow\) if no suggestion is used in a session
  – tracking emotion during typing session
Experiment Apparatus

Suggestions

- Custom keyboard
  - tracing user typing and providing suggestions (based on English dictionary)
- Emotion self-report collection UI
  - collects emotions at end of a session
  - four emotions - happy, sad, stressed, relaxed (based on Circumplex model)

(a) App keyboard       (b) Emotion collection UI       (c) Circumplex model

Auto-suggest Usage Scenario ➔ Methodology ➔ User Study ➔ Evaluation ➔ Take-home Points
Auto-suggest Usage Prediction Model

- Auto-suggest usage prediction model
  - Personalized
    - Random Forest
  - Two classes
    - accepted
    - skipped
  - Features
    - Emotion-related features

<table>
<thead>
<tr>
<th>Feature name</th>
<th>Feature description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Emotion_curr</td>
<td>Emotion associated with current session</td>
</tr>
<tr>
<td>Emotion_prev</td>
<td>Emotion associated with previous session</td>
</tr>
<tr>
<td>Time_elapsed</td>
<td>Elapsed time between previous and current session emotion recording timestamp</td>
</tr>
</tbody>
</table>

Table 2: Features used for auto-suggest usage prediction
# User Study

<table>
<thead>
<tr>
<th>Study parameters</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study duration</td>
<td>3-weeks (in-the-wild)</td>
</tr>
<tr>
<td>Total participants</td>
<td>20 (15 M, 5 F) → university students</td>
</tr>
<tr>
<td>Age range</td>
<td>20 – 35 years</td>
</tr>
<tr>
<td>Installed the app in the participants mobile phones</td>
<td></td>
</tr>
<tr>
<td>Excluded participants</td>
<td>7 (as recorded less than 20 suggestions)</td>
</tr>
<tr>
<td>Final participants</td>
<td>13 (10 M, 3 F)</td>
</tr>
</tbody>
</table>
# Dataset

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total typing sessions</td>
<td>3,284</td>
</tr>
<tr>
<td>No Response sessions</td>
<td>330 (~ 10%)</td>
</tr>
<tr>
<td>Effective sessions</td>
<td>2954</td>
</tr>
<tr>
<td>Avg. session per user</td>
<td>227 (std dev. 151.7)</td>
</tr>
<tr>
<td>Auto-suggest accepted</td>
<td>841 (~28%)</td>
</tr>
<tr>
<td>Auto-suggest skipped</td>
<td>2,113 (~ 72%)</td>
</tr>
</tbody>
</table>

User-wise auto-suggest usage reveals that most of the users have more skipped session.

Auto-suggest Usage Scenario ➔ Methodology ➔ **User Study** ➔ Evaluation ➔ Take-home Points
Role of Emotion on Auto-suggest Usage

Does auto-suggest usage vary across emotions?

Auto-suggest usage across different emotions for all users

- Relaxed, sad \(\rightarrow\) **Low activity level**
- May influence more auto-suggest usage

Users more likely to use auto-suggest when sad or relaxed
Role of Emotion on Auto-suggest Usage

Does auto-suggest usage vary across individual user emotions?

- Majority of the users use auto-suggest in relaxed
- Distribution of emotions in accepted and skipped is statistically significant \((p < 0.05)\) using chi-square test

Auto-suggest usage comparison across different emotions for individual user
Evaluation

• Experiment setup
  – 10-fold cross validation
  – Metrics: AUCROC, F-score

• Baselines
  – Most-represented Emotion (MRE)
    • Auto-suggest usage → highest in one emotion
    • Personalized model, which determines auto-suggest usage in these emotions
  – Generalized (GEN)
    • Aggregating data from all other users
    • Leave-one-out-cross-validation
Evaluation

• How accurate is the auto-suggest usage prediction model?

(a) Classification performance

(b) Comparison with baselines

Avg. accuracy (AUCROC) – 73% (std dev 9%)
Outperforms the baselines
Evaluation

• How effective are the features to predict auto-suggest usage?

<table>
<thead>
<tr>
<th>Feature</th>
<th>Rank</th>
<th>Avg. IG</th>
</tr>
</thead>
<tbody>
<tr>
<td>Emotion$_{curr}$</td>
<td>1</td>
<td>0.1194</td>
</tr>
<tr>
<td>Emotion$_{prev}$</td>
<td>2</td>
<td>0.1098</td>
</tr>
<tr>
<td>Time$_{elapsed}$</td>
<td>3</td>
<td>0.0794</td>
</tr>
</tbody>
</table>

Ranking different features based on InformationGain

✅ Current session emotion is the most discriminating one
Evaluation

- How to improve auto-suggest usage prediction performance?
  - Balancing the dataset (accepted and skipped class)

Classification performance after balancing the dataset

- AUCROC – 82%
- F-score (Accepted) – 75%
- F-score (Skipped) – 78%

Classification performance improves after balancing dataset
Conclusion

• Auto-suggest usage is related with human emotion
  • more likely to use in sad or relaxed state
• Based on emotion, auto-suggest usage can be detected with an accuracy of 82%
Thank You!!

Acknowledgement

- Microsoft Research India
- LRN Foundation India
- IUI student travel grant

http://cse.iitkgp.ac.in/~surjya.ghosh/

Complex Network Research Group (CNeRG)

@iitkgpcnerg @cnerg http://www.cnergres.iitkgp.ac.in/
Dataset

- How close to the typing session, the emotion is collected?

Distribution of elapsed time between typing and emotion recording for all sessions across all users.

- Median elapsed time is less than 5 minutes; 75th and 90th percentile elapsed time is less than 30 minutes and 1 hour respectively.
Users more like to use auto-suggest when sad or relaxed

- Majority of the users use auto-suggest in relaxed
- Distribution of emotions in accepted and skipped is statistically significant using chi-square test
Outline

• Auto-suggest Usage Scenario

• Methodology
  – Experiment Apparatus
  – Model Construction

• User Study
  – Field Study
  – Dataset

• Evaluation

• Take-home Points
Background

• Text-entry in small touch-based devices
  – relatively inconvenient due to limited space

• To facilitate, different techniques like auto-complete, auto-suggest are used
  – However,
    • need additional space
    • increase cognitive load to read, parse and select correct word
    [CHI 2016]

• Making auto-suggestions adaptive, when the users are more likely to use can overcome these
Thank You!!

Acknowledgement
• Microsoft Research India
• LRN Foundation India
• IUI student travel grant

http://cse.iitkgp.ac.in/~surjya.ghosh/

Complex Network Research Group (CNeRG)

@iitkgpcnerg @cnerg http://www.cnergres.iitkgp.ac.in/
Background

- Text-entry in small touch-based devices—relatively inconvenient due to limited space

- To facilitate, different techniques like *auto-complete, auto-suggest* are used

  — However,
  - reserves fixed space in keyboard layout
  - increase cognitive load
    — to read, parse and select correct word [CHI 2016]

**Opportunity to improve the layout by making these techniques adaptive**