

# Incorporating User Grouping into Retweeting Behavior Modeling

## Supplementary Material

Jinhai Zhu\*   Shuai Ma\*   Jing Wang\*   Hui Zhang\*   Chunming Hu\*   Xiong Li†

### Appendix: Extra Experiments

**Exp-5: Case Study for User Clustering** In this section, we show the results of our system GruBa for User Clustering. We carefully analyzed the statistical information of four groups obtained from user clustering.

(1) Most users of group one are male (as shown in Sub-Fig. 1(a)), and Sub-Fig. 1(b) shows that they are mainly in Beijing (北京). The amount of their followees is mostly similar as their followers (as shown in Sub-Fig. 1(d)). The most active time of this group is mainly between 10 am and 12 am (as shown in Sub-Fig. 1(e) and 1(f)).

(2) Mostly users of group two are female (as shown in Sub-Fig. 2(a)) and are also mainly in Beijing (北京) (shown as Sub-Fig. 2(b)). Similar to group one, most of them have followees close to followers (as shown in Sub-Fig. 2(d)). The most active time of this group is mainly between 10 pm and 12 pm (as shown in Sub-Fig. 2(e) and 2(f)).

(3) Most users of group three are male (as shown in Sub-Fig. 3(a)). They come from a wide range of provinces (as shown in Sub-Fig. 3(b)). Most of them have number of followers far more than followees (as shown in Sub-Fig. 3(d)). The most active time of this group is mainly in 10-12 am and 10-12 pm (as shown in Sub-Fig. 3(e) and 3(f)).

(4) Most users of group four are female (as shown in Sub-Fig. 4(a)), and they come from a wide range of provinces just as group three (as shown in Sub-Fig. 4(b)). Most of these users have the amount of followers far more than followees (as shown in Sub-Fig. 4(d)). And the most active time of these users is mainly between 10 pm and 12 pm (as shown in Sub-Fig. 4(e) and 4(f)).

According to Sub-Figs. 1(c), 2(c), 3(c) and 4(c),

the number distribution of microblogs for four groups is similar with each other. We visualize each group's long-term interest expressed as words cloud (shown in Sub-Figs. 1(i), 2(i), 3(i) and 4(i)). As we can see, group one and group three have similar interests.

---

\*SKLSDE Lab, Beihang University, China. Email: {zhujh@act., mashuai, zhanghui@act., hucm@}buaa.edu.cn, wjseawind@gmail.com

†National Computer Network Emergency Response Technical Team/Coordination Center of China. Email: li.xiong@foxmail.com



Figure 1: The Statistics of User Group One

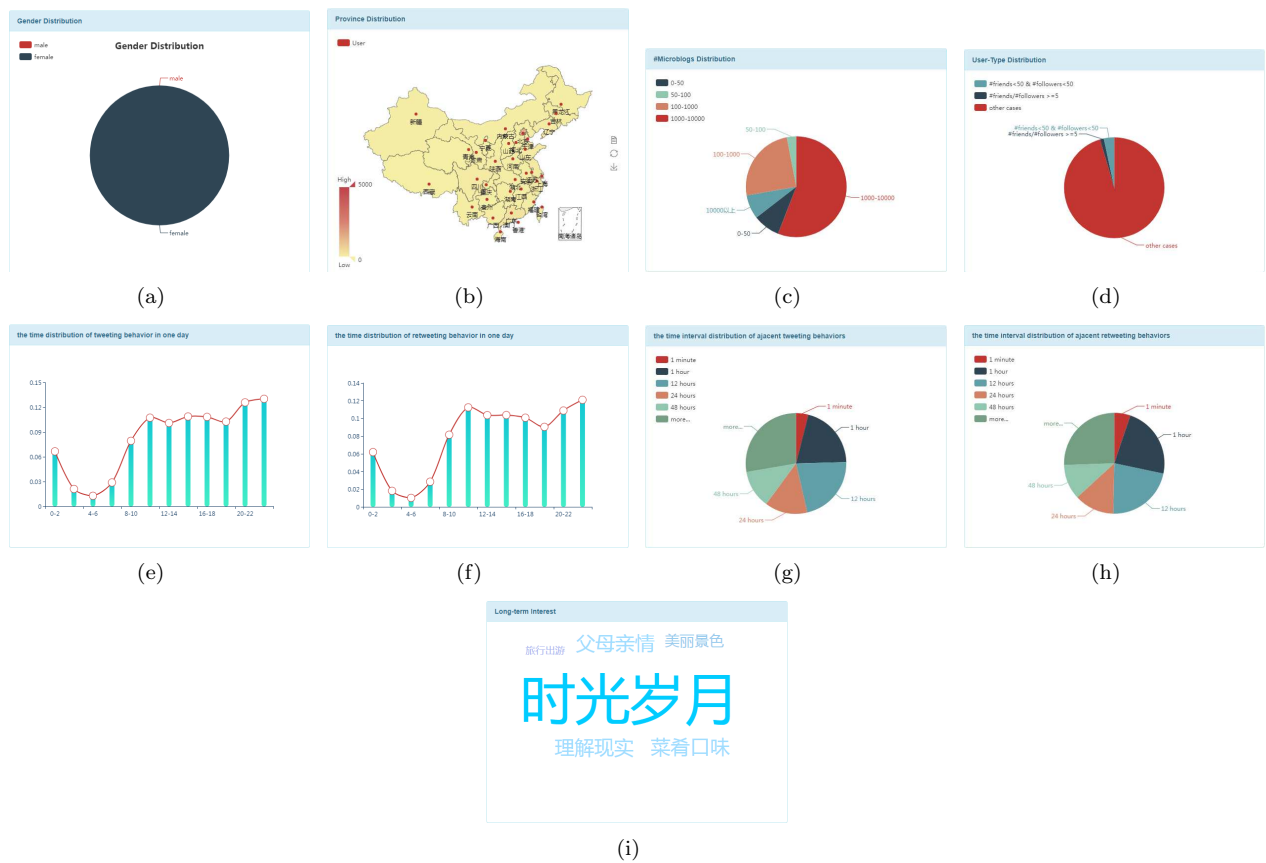


Figure 2: The Statistics of User Group Two



Figure 3: The Statistics of User Group Three



Figure 4: The Statistics of User Group Four