Analyzing on User Behavior and User Experience of Social Network Services

Rong Bao, Lei Chen(✉), and Ping Cui

Jiangsu Province Key Laboratory of Intelligent Industry Control Technology, Xuzhou University of Technology, Xuzhou 221018, China
chenlei@xzit.edu.cn

Abstract. The user behavior characteristics of mobile social network services are beneficial for evaluating the user experience, and the test cases and test scenarios should be designed according to the user behavior characteristics. The current researches have been heavily addressed on the action sequence and the frequency distribution of user behavior. There is little research on the user's action triggering network flow under different scenarios. This paper analyzes the distribution character of user actions, and tests the waiting time of different user actions in different scenes. The results suggest that the complex scenarios can be consisted of some typical user behaviors.

Keywords: Quantify of user experience · Social network service · User behavior · Communication scenarios · Mobile internet

1 Introduction

With the development of 4G/5G communication technology, a large number of Social Network Services (SNS) have emerged. Mobile operators and equipment providers pay more attention to the Quality of User Experience (QoE) of such services, and improve the user experience through more intelligent scheduling strategy [1]. Some scheduling strategies adopt Deep Packet Inspection (DPI) technology to identify the specific services the packets belong to, and then use different scheduling strategies according to different services. This reflects that the communications industry has recognized that different services impact on user experience in different way. The traditional evaluation system mainly measures the service quality of large-scale business with indistinguishable general indicators. However, many test systems are still based on the data flow model [2–4], and cannot effectively trigger intelligent scheduling strategies, so it is impossible to evaluate the effectiveness of these intelligent scheduling strategies. Many communication enterprises have to use manual dial-up method to verify the new scheduling strategy, which is unable to simulate large-scale scenarios. A possible

This work is partly supported by the Natural Science Foundation of Jiangsu Province of China (No. BK20161165), the applied fundamental research Foundation of Xuzhou of China (No. KC17072), and the Open Fund of the Jiangsu Province Key Laboratory of Intelligent Industry Control Technology, Xuzhou University of Technology.

H. Song et al. (Eds.): SIMUtools 2019, LNICST 295, pp. 180–185, 2019.
https://doi.org/10.1007/978-3-030-32216-8_17

approach is to replay the real data packets captured from networks in simulation system [5]. These packets might be generated by user actions, such as login, sending message, comment, and so on. A complex test scenario should consist of these actions according to user behavior. Therefore, the new problems are how the user behavior affects the QoE, and how to reconstruct user behavior using the captured data packets.

Through the analysis of the user behavior characteristics of the main SNSs, some researchers found the concentrated distribution of common actions, action frequency and information length of users [6, 7]. Some studies suggest that the traffic model of traditional communication network can be used for reference to analyze the communication behaviors of computer network users [8]. In the model, the network user behaviors can be described by the Poisson distribution of parameter λ:

$$\boldsymbol{P_u(i,t)} = \frac{(\boldsymbol{\lambda t})^{\boldsymbol{i}}}{\boldsymbol{i}!}\boldsymbol{e}^{-\boldsymbol{\lambda t}} \tag{1}$$

where, i is the number of services during t. The amount and length of data flow of each service follow the geometric distribution of E_f and E_l respectively:

$$\boldsymbol{P_f(n)} = \frac{1}{\boldsymbol{E_f}}\left(1 - \frac{1}{\boldsymbol{E_f}}\right)^{\boldsymbol{n-1}} \tag{2}$$

$$\boldsymbol{P_l(k)} = \frac{1}{\boldsymbol{E_l}}\left(1 - \frac{1}{\boldsymbol{E_l}}\right)^{\boldsymbol{k-1}} \tag{3}$$

where, n and k are the number and length of data flows, respectively. The network traffic $n_{ij}(T)$ can thus be obtained, according to the number of users $\boldsymbol{n_{ij}}(T)$, the number of data flows $f_i(T)$ and the length of data flow $n_{ij}(T)$ during (0, T):

$$\mathbf{N(T)} = \sum\nolimits_{\mathbf{i=1}}^{\boldsymbol{u(T)}} \sum\nolimits_{\mathbf{j=1}}^{\boldsymbol{f_i(T)}} \boldsymbol{n_{ij}(T)} \tag{4}$$

And the network traffic during $(\mathrm{t}, \mathrm{t} + \tau)$ can be written by:

$$\boldsymbol{N_t(\tau)} = \boldsymbol{N(t+\tau)} - \boldsymbol{N(t)} \tag{5}$$

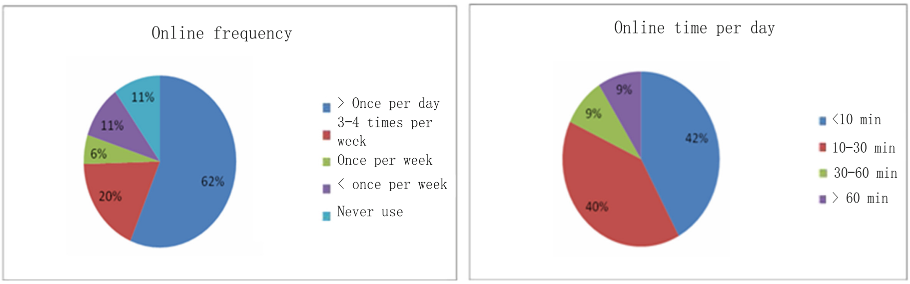
In this traffic model, the parameter λ depends on the user density and usage habits of users in the scene, $\mathrm{E_f}$ depends on the operation sequence and action frequency of the users, and $\mathrm{E_1}$ is determined by the distribution of user action. If there is centralized distribution in user behaviors, a few typical action sequences can be used to simulate the entire communication scenario and form impact scenarios similar to the real scenarios.

2 User Behavior Analysis of SNS

According to the above model, it is possible to reconstruct the real scene with a small number of typical user behaviors, if the data length and frequency of user behaviors obey a central distribution. Related studies have found the centralized distribution of user behavior in SNS. In this paper, we analyze the online time and operation frequency of SNS users. To observe the impact of user behaviors to QoE, we also test the appreciable indicators of the end users in different scenarios with different user density and behavior.

2.1 Central Distribution of Online Frequency

To analyze the distribution rule of online frequency for SNS, we survey 35 college students on their usage count per day. As shown in Fig. 1(a), the result shows that the special user group has similar usage habit.



(a) online frequency of Weibo (b) online time of Weibo

Fig. 1. Online characters of Weibo

Figure 1(b) shows the distribution of users' usage time of Weibo in campus. This implies the centralized distribution of usage habit in a special communication scenario.

2.2 Central Distribution of User Data

We collect Weibo information of 60,000 users. We compute the average number of posts these users published per day. 90% Weibo users daily post less than 10 messages.

In addition, the length of the post is also restricted by usage habit. In this paper, the length of the 2,622 Weibo posts without links is analyzed in Fig. 2, in which the vertical coordinate is the number of blog posts and the horizontal coordinate is the length of posts. As shown in Fig. 2, the lengths of most posts are between 10 and 50 bytes. To eliminate the oscillation caused by the double-byte representation of Chinese characters in the computer system. Figure 2 retains only double-byte data. This result shows centralized distribution and heavy-tailed distribution.

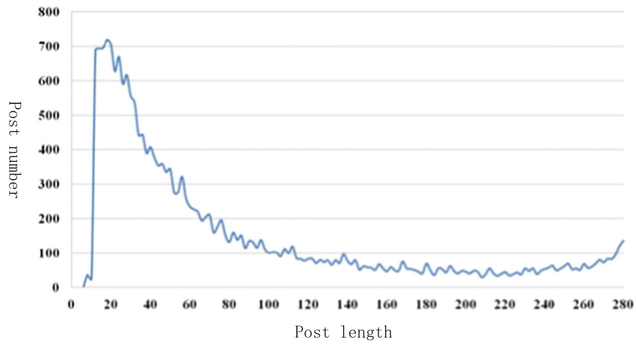


Fig. 2. Distribution of post length (only even bytes)

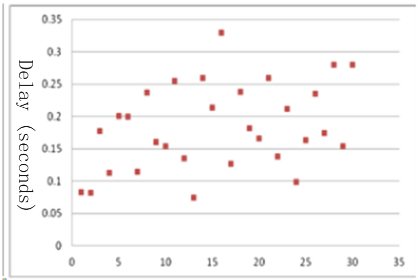
3 Impact of Scenario on QoE Indicators

We survey the degree to which the communication scenario affects the user experience. In this paper, the appreciable indicators of SNS are evaluated in the urban area. The cellular network access is provided by the same communication operators. Therefore, the infrastructure conditions of the tests are similar, and the performance fluctuation should be caused by the behavioral features of different user groups, such as user density, operation frequency, and other operation habits.

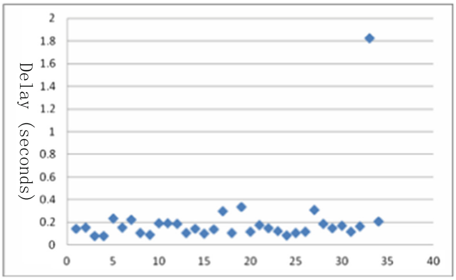
Figure 3 shows the delay test of QQ message sending in different scenarios. The test results show that the distribution range of delay is approximately the same. However, the number of delays with a significant deviation from the main distribution range is obviously different. In particular, Fig. 4 shows that the delay difference between different time periods in the same place is also obvious. The test results indicate that user behavior is key factor of communication scenarios, which has a certain impact on the QoE indicators.

In same area, communication capacity of communication operator is more similar. In different scenarios, the group behavior of users is different, and the usage habits are thus different. Therefore the degree to which the user behavior affects QoE indicators can be observed. In this paper, we chose to test the delay of posting Weibo comments through cellular network under different scenarios in a large research and development base. The scenarios include office area, experiment area, restaurant, lounge area, and so on. Except that the restaurant is significantly more densely populated than other areas, all other areas have similar user density. And the communication operators are known to identify data packets and adopt special scheduling strategies for SNS.

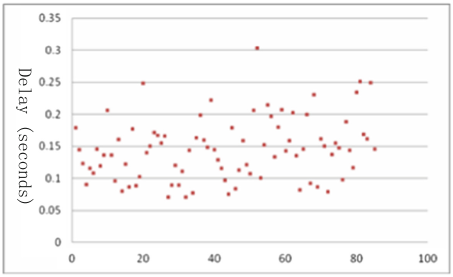
The test results in Fig. 4 also indicate that in the case of same user density and similar communication capacity user behavior causes delay fluctuation. Because the delay jitter bigger than 0.1 s is appreciable, the delay fluctuation in Fig. 4 is obvious for users.



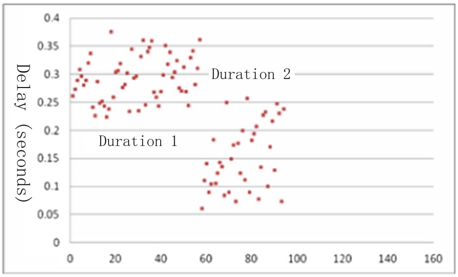
(a) shopping mall



(b) exhibition center



(c) International Finance Centre (indoor）



(d) International Finance Centre (outdoor）

Fig. 3. QQ message delay in various scenairos

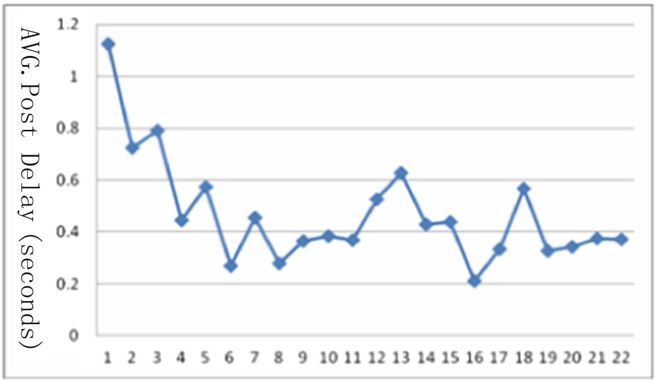


Fig. 4. Average post delay in 22 test point (5 tests for each point)

4 Conclusions

In this paper, we mainly analyze the distribution feature of user behavior for SNS, and the degree of impact of user behavior on QoE indicators. The investigation results show the centralized distribution of user behavior. The test results in real scenarios imply that the user behavior actually affects the user experience. Therefore, we propose that a QoE test scenario can be built by the typical user behaviors. The relationship between user behavior and network load should be further studied to guide the test case design.

References

1. Wang, Y., Li, P.L., Jiao, L., et al.: A data-driven architecture for personalized QoE management in 5G wireless networks. IEEE Wirel. Commun. **24**(1), 102–110 (2017)
2. Jiang, D., Wang, W., Shi, L., Song, H.: A compressive sensing-based approach to end-to- end network traffic reconstruction. IEEE Trans. Netw. Sci. Eng. (2018). https://doi.org/10.1109/tnse.2018.2877597
3. Jiang, D., Huo, L., Song, H.: Rethinking behaviors and activities of base stations in mobile cellular networks based on big data analysis. IEEE Trans. Netw. Sci. Eng. **1**(1), 1–12 (2018)
4. Jiang, D., Huo, L., Li, Y.: Fine-granularity inference and estimations to network traffic for SDN. PLoS ONE **13**(5), 1–23 (2018)
5. Chen, L., Jiang, D., Song, H., et al.: A lightweight end-side user experience data collection system for quality evaluation of multimedia communications. IEEE Access **2018**(6), 15408–15419 (2018)
6. Jin, Y., Duffield, N., Haffner, P., et al.: Can't see forest through the trees. In: The Proceedings of 9th Workshop on Mining and Learning with Graphs, San Diego, USA (2011)
7. Schneider, F., Feldmann, A., Krishnamurthy, B., et al.: Understanding online social network usage from a network perspective. In: Proceedings of the 9th ACM SIGCOMM Conference on Internet Measurement Conference, Chicago, USA (2009)
8. Zhang, S., Zhao, Z., Guan, H., et al.: A modified poisson distribution for smartphone background traffic in cellular networks. Int. J. Commun. Syst. **30**(6) (2017). https://doi.org/10.1002/dac.3117