iVIS: Interpretable Interactive Visualization for User Behavior Clusters

Yieun Kim, Yohan Bae, Junghyun Kim, and Yeonghun Nam(✉)

Samsung Research, Samsung Electronics, Seoul, Republic of Korea
{yieun12.kim,yhan.bae,jh24.kim,yeonghun.nam}@samsung.com

Abstract. To improve Quality of Experience (QoE) and develop new features, understanding users and making a decision to target specific user groups are important to service providers. Based on the internal interviews, we find that service operators have trouble in identifying user behavior characteristics for numerous services in a short time. To address this challenge, we present iVIS, an interactive visualization system that clusters the user behaviors and visualizes the representative behavior patterns. With iVIS, service providers can interpret the user clusters (e.g., heavy/light users), and drill down to a particular cluster to get details interactively. To evaluate our system, we conduct a case study on the log data from the internal data catalog service which enables researchers to browse and use datasets. We found service operators could rapidly interpret representative user behaviors, and discover new behavior types such as *frustrated users* (i.e., users who only explored datasets for a while but not use them) or *testers* (i.e., users who used the service for testing) by refining the clustering results.

Keywords: Behavior analysis · User clustering · Interactive visualization

1 Introduction

The success of an online service is increasingly dependent on how quickly service providers (e.g., data analysts, retailers, marketers, and service managers) identify target customer characteristics and how accurately they determine their needs. Service providers can establish a market strategy, provide customized services, and develop new features aligning with consumer needs and priorities based on the behavioral analysis. However, as service functions are complicated and user characteristics are diversified, it is becoming difficult to understand user behaviors to make a decision on developing the service.

Typically, user behavioral analysis is based on collected user activity records. Request/response records, button clicks, page access paths, finger swipes and input records can all be included [1]. It takes a very long time for service providers to make data-driven decisions based on various user records. Therefore, a visual interface must be provided to help them to present the key features explaining user behaviors.

C. Stephanidis et al. (Eds.): HCII 2020, CCIS 1293, pp. 47–52, 2020.
https://doi.org/10.1007/978-3-030-60700-5_6

In addition, several visualization interfaces for user behavior clusters in online service cluster users with similar behavior patterns and provide indicators to identify the cluster characteristics. However, the existing cluster visualization interface has the following limitations. *First*, they mainly focus on identifying and drilling down the characteristics of each cluster. However, through preliminary interviews, we have found a need for service providers to compare the differences between clusters. *Second*, it is not allowed to transform the cluster results by additionally dividing or integrating clusters based on their domain knowledge.

To meet these challenges, in this paper, we present iVIS, an interactive visualization system that clusters users and visualizes their behavior patterns. iVIS provides information about cluster similarity and differences, and user distribution per cluster as a compressed view. Moreover, the usability for each service function is presented for each cluster, so that the cluster can be interpreted intuitively (e.g., heavy/light users). Furthermore, iVIS supports interaction to combine or divide user clusters inferred by the clustering model. In other words, iVIS helps service providers to answer key questions regarding their services, e.g., what is the difference between major and minor behavioral groups? What are the main user groups interested in? At what time or which regions are the groups distributed? What is the difference between the group that mainly uses the newly launched service function and the group that does not use it?

In summary, the main contributions of our work are as follows:

- We present iVIS, an interactive visualization system that clusters the user behaviors and visualizes the representative behavior patterns.
- We support not only the exploration of each group in detail, but also the comparison of several groups.
- We allow analysts to refine the clustering results by interactively merging the clusters or partitioning a cluster by user-defined rules.

2 Related Works

Several visualization interfaces based on user behavior clusters have been proposed in recent articles. In [1], a cluster visualization tool is proposed to explore hierarchically clustered online service users. The cluster hierarchy is presented as packed circles where more general patterns are presented as higher-level, and further identifying patterns as lower-level. In [2], a user clustering algorithm is proposed based on text analysis and topic modeling for online services, and an interactive visual analytics interface is designed to identify each group. It enables users recursively explore hierarchical behavior patterns in service function sequences. In [3], an interactive cluster exploration tool is designed to enable user-guided clustering. Users can cluster major clickstream patterns using their domain knowledge, and a semi-supervised automatic clustering algorithm is applied to assign complex patterns to the existing clusters.

3 iVIS Interface

3.1 Design Rationale

We interviewed service providers about how they analyze user behavior and what pain points they are facing while conducting the analysis. The result is summarized in the following three tasks they need help with:

Identify interesting cluster straightforwardly [DR-1]: The service providers confirm that it is important to view a cluster overview at the beginning of the analysis. Moreover, in order to reduce analysis time, they want to simply and clearly figure out interesting user groups (e.g., a group that has unexpected usage patterns or leaves the service) that will be considered as the starting point for analysis.

Drill down into individual groups and compare them interactively [DR-2]: For behavior analysis, they typically select multiple groups as the starting point, and understand the groups by interactively investigating as well as comparing characteristics. Therefore, they need a tool that supports interaction for both drilldown and comparison in a compressed view.

Customize the results based on domain knowledge [DR-3]: After analyzing the characteristics of all clusters, they want to use their domain-knowledge to customize the results. One said: "Other automated analytic tools have limitations to conduct domain-specific analysis. I want to apply my domain knowledge to the result, which is useful for understanding user segments and making decisions".

After the user interviews, we devise an interactive visualization for user behavior clusters that satisfying the requirements.

3.2 iVIS Interface

Cluster Overview (Fig. 1a). The widget represents each cluster as a circle. The radius of each circle is proportional to the number of users in the group, and the distance between circles represents the similarity between clusters. When a cluster is selected, other widgets reflect the selection and display the cluster detail. Moreover, service providers can customize the original result (Fig. 2a) using their domain knowledge by merging similar clusters (Fig. 2b) or splitting one cluster into multiple clusters with appropriate criteria (Fig. 2c).

Usage Segmentation (Fig. 1b). To provide explainable cluster characteristics, we categorize the user behavior into four usage patterns: *inactive*, *active only for a certain period*, *steady but less active*, and *steady and active*. We then visualize the proportion of each pattern in the widget.

Event Composition (Fig. 1c). The widget indicates distribution of event occurrences for each cluster. A cluster characteristic can be identified by comparing the event composition rate with the other clusters. We provide three switchable metrics to present event compositions: ratio, count, and average (which is the number of occurrences per user), so that service providers can perform various interpretations according to the selected metric.

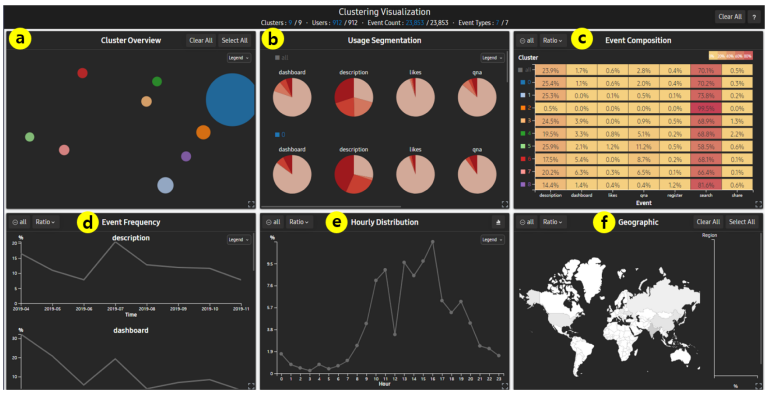

Fig. 1. Interface of iVIS. a) The number of users in each cluster and the similarities between clusters are indicated by the circle size and distances between them, respectively. b) Usage characteristics are represented by pie charts where each chart displays the usage pattern proportions for each event type. The other coordinated views show c) the event compositions, d) event frequencies over time, e) hourly event distributions, and f) geographic event distributions

(a) (b) (c)

Fig. 2. Cluster overview. a) Before customizing, b) after integrating two clusters (in the upper left), and c) after dividing the cluster (in the center of the left)

Event Frequency (Fig. 1d). The widget displays the number of event occurrences over a given time range. By selecting clusters and sub-time range, they can understand the relationship among events, clusters, and the specific period. The frequency metric can be changed from the number of events to the occurrence ratio.

Hourly Distribution (Fig. 1e). The widget shows the aggregated hourly event counts. It provides information about when users in each cluster actively use the service. The unit of time can be customized by the other units such as day, day of week, and month.

Geographic (Fig. 1f). Geographic information of each cluster is visualized in the widget. The number of users in each country is represented by the color opacity. When a cluster is selected, the regions where the cluster is mainly distributed are colored.

Cross-Filtering. All widgets in the iVIS interface can be cross-filtered, thus service providers can select any interesting cluster, event type, period, hour and region as filter conditions.

4 Case Study

4.1 Study Design

To evaluate the iVIS interface, we conducted a case study with two service administrators and a developer (referred to as P1-P3 later) using the service logs collected from the internal data catalog service which allows researchers to browse and use datasets. We asked domain experts to try our interface to freely explore the given clustering results for 30 min.

4.2 Study Results

We observed how domain experts analyze user behaviors of service using iVIS and conducted interviews to understand their analysis process. At first, P1 focused on the cluster overview to understand the user groups. Subsequently, he selected interesting clusters and events in the event composition view. On the contrary, other domain experts first looked at the usage segmentation view, selected two or three clusters of interest from the cluster overview, and explored other coordinated views. P2 stated "The usage segmentation view helped me explore the distribution of usage patterns for each user group. As a result, I was able to find the user groups of interest."

They also discovered new behavior types and rapidly interpreted representative user behaviors. Two participants found previously unknown behavior types such as *frustrated users* (i.e., users who only explored datasets for a while but not use them for a certain reason) or *testers* (i.e., users who used the service for testing). According to interviews, they discovered those types by utilizing the interaction such as switching the metrics in the event composition and event frequency widgets. In addition, they showed a great passion for discovering new behavior types using our interface.

In the early user interview stage, analysts wanted to customize the clustering results by interactively merging the clusters or partitioning a cluster by user-defined rules. In the case study, they actively used those interactions and refined the results. However, P2 mentioned "I think the partitioning function is useful, but it was difficult to determine user-defined rules. The appropriate guidance may be required when we first make them."

5 Conclusion

In this work, we present iVIS, an interactive visualization system that clusters the user behaviors and visualizes the various behavior patterns. With our interface, service providers can interpret the user clusters and drill down to a particular cluster to get details interactively. The case study conducted with two service providers and a developer showed that iVIS helps them explore each user group and identify unusual behaviors. Moreover, we gained data analysts' feedback to improve our visual representations and interactions. Though data analysts are able to customize the clustering results with iVIS, they can still be dissatisfied with the results. In the future, we plan to support the use of diverse clustering algorithms. Furthermore, we will design novel visualizations that are suitable for comparing the results based on various clustering models.

References

1. Wang, G., Zhang, X., Tang, S., Zheng, H., Zhao, B.Y.: Unsupervised clickstream clustering for user behavior analysis. In: CHI '16: Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, USA, pp. 225–236. ACM (2016)
2. Nguyen, P.H., et al.: VASABI: hierarchical user profiles for interactive visual user behaviour analytics. IEEE Trans. Vis. Comput. Graph. **26**(1), 77–86 (2020)
3. Wei, J., Shen, Z., Sundaresan, N., Ma, K.-L.: Visual cluster exploration of web clickstream data. In: 2012 IEEE Conference on Visual Analytics Science and Technology (VAST), USA, pp. 3–12, IEEE (2012)