Revealing Universities' Atmosphere from Visitor Interests Using Search Queries and GPS Logs

Kaoru Miyanaga
Institute of Science Tokyo
Tokyo, Japan
miyanaga@miubiq.cs.titech.ac.jp

Soto Anno
Institute of Science Tokyo
Tokyo, Japan
anno@miubiq.cs.titech.ac.jp

Kota Tsubouchi
LY Corporation
Tokyo, Japan
ktsubouc@lycorp.co.jp

Masamichi Shimosaka
Institute of Science Tokyo
Tokyo, Japan
simosaka@miubiq.cs.titech.ac.jp

I. INTRODUCTION

When visiting universities, you might notice the distinctive atmospheres of each university, such as a calm and serious environment or a lively enthusiasm for sports. Capturing these atmospheres could help in promoting universities and fostering development in the communities around the universities. To explore the atmospheres of universities, we analyze the thoughts and interests of university community members, such as students and faculty members. Specifically, we use a large-scale dataset derived from search queries and GPS logs to quantify visitors' interests. Additionally, to extract the meaningful atmospheres of universities, we apply topic modeling to the dataset.

II. METHOD

In recent years, to understand the characteristics of urban areas from visitors' interests, the combination of search queries and GPS logs has gained attention. For instance, Sakamoto *et al.*'s method represented each urban area with numerical vectors based on embedded words searched by users visiting each urban area [1].

Originally developed for extracting topics from documents, topic models have also been applied to urban analysis [2]. To apply topic models to our task, we treat the atmospheres within the university as topics in the document.

III. EXPERIMENT

In this experiment, we used the City-Atmosphere Dataset provided by LY Corporation. This dataset, generated using Sakamoto *et al.*'s method, does not contain any personally identifiable information. The dataset covers 9 prefectures, including Tokyo and Osaka, and spans the period from March 1, 2022, to August 31, 2022. The data consists of 12 dimensions based on search categories. To increase the data volume per urban area, we prepared data based on visitor attributes, such as gender and age group. We filtered the data based on the number of search queries and grouped it by urban areas.

Firstly, by considering both visibility and model performance, we extracted atmospheres and the proportions of atmospheres by each university. Figure 1 shows features of some extracted atmospheres characterizing each university. Interestingly, Atmosphere 10 suggests an interest not only in knowledge but also in venues and shopping.

Based on these atmospheres, Figure 2 shows a comparison of the proportions of atmospheres across the universities in Japan. For instance, the high proportion of Atmosphere 6 at The University of Tokyo and Waseda University indicates that these universities may focus on academics and club activities. These results also suggest the potential demand for entertainment facilities for students in the areas surrounding those two universities.
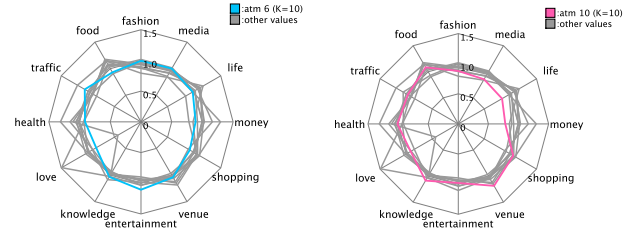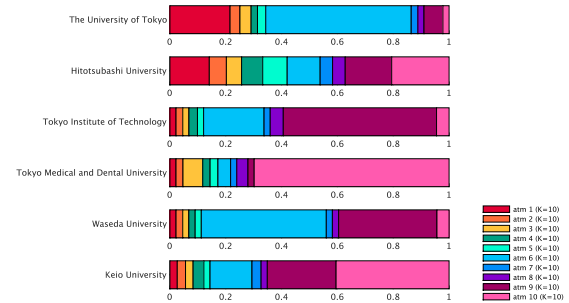

Fig. 1: features of some extracted atmosphere

Fig. 2: Comparison of the proportions of atmospheres

IV. CONCLUSION

We captured the different atmospheres in each university from visitors' interests with a topic model. Future work will address analyzing urban areas not limited to universities, such as stations and commercial facilities.

REFERENCES

[1] T. Sakamoto, H. Yasuda, A. Tsujimoto, M. Iwamoto, Y. Hayakawa, T. Usami, K. Tsubouchi, and M. Shimosaka. Cityatmosphere: Vr image to glimpse wishes in the air. In *In Proc. of UbiComp*, 2019.
[2] J. Yuan, Y. Zheng, and X. Xie. Discovering regions of different functions in a city using human mobility and pois. In *In Proc. of SIGKDD*, 2012.