Mapping Emerging Forms of Tourism in Central Tokyo Through SNS Data

Andrew Burgess[1], Xinyue Song[2], Kaori Ito[1]
[1] Department of Architecture, Faculty of Science and Technology, Tokyo University of Science [2]Aeon Mall
Email: < a.burgess@rs.tus.ac.jp>

(1) Purpose: The dramatic increase of foreign visitors to Japan in recent years coupled both social changes and changes in technology has led to profound changes in tourist behavior. Greater access to the internet via free wi-fi spots, availability of rental sim cards, increasing use of services such as Airbnb, online maps and translation tools has led to a wider variety of tourist activities in previously non-tourist locations. This research presents an update and expansion of research previously presented at CSIS DAYS by analyzing the spatial distribution and spatial relationships between geotagged posts by non-Japanese users from Twitter and Sina Weibo in Tokyo. The ongoing goal of this research is to understand the spatial characteristics of emerging forms of tourism and the role that user-generated information plays in informing and influencing tourist behavior.

(2) Approach: This research compares the activity of Chinese language users and English language users not only by geolocation but also text content. Words are weighted by tf-idf scores are attributed to places before spectral coclustering is carried out to simultaneously cluster words and places. The use of tf-idf scores rather than frequency counts extracts spatially significant terms and phrases rather than simply common terms.

(3) Originality: While research utilizing Twitter data abounds, the comparison of Weibo data is rare. Further, the simultaneous analysis of both sets of data at the urban scale provides a unique insight into tourist behavior.

(4) Findings: Figure 1 shows the variation in location of Twitter and Weibo Users by comparing zscores. Figures 2 & 3 show the comparison of two communities detected in the cluster analysis. The cluster analysis found significant variations in the clustering of areas due to variations in the types of activities and impressions of posters. In general, the prevalence of fan-based tourist activities, for example concert-going, of Weibo users influenced this variation.

(5) Data and Software: Data was collected through the Twitter API and Weibo API (v.1). Data collection, preprocessing and analysis was carried out in python 3.6 (sciPy, Natural Language Toolkit (NLTK), tweepy) and visualized in QGIS 2.14.

(6) References:

Dhillon, I. (2001) Co-clustering Documents and Words Using Bipartite Spectral Graph Partitioning. UT CS Technical Report #TR 2001-05.

Lansley. G, Longley, P. (2016) The Geography of Twitter Topics in London. *Computers, Environment and Urban Systems*, 58, 85–96.

宋歆月, バージェス・アンドリュー,伊藤香織（2016）「SNS 投稿による訪都外国人の行動特性分析に関する研究」, CSIS DAYS 2016.

Figure.2: Overlapping Areas Between Clusters

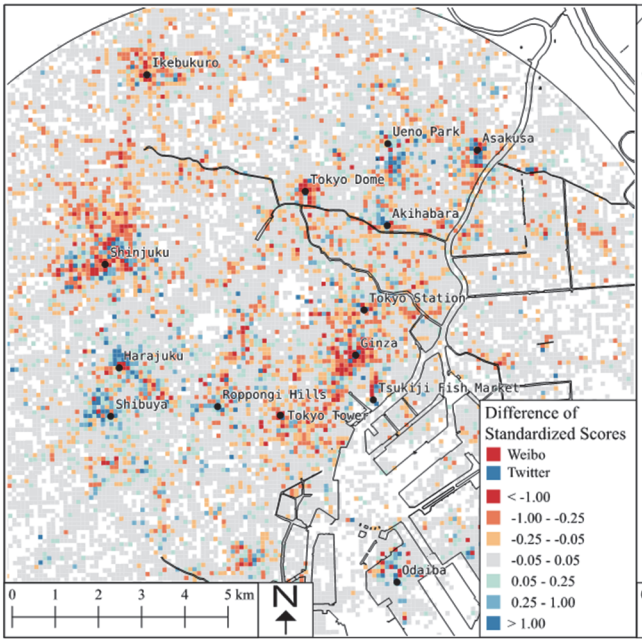

Figure.1: Comparison of Weibo and Twitter Posts by Location.

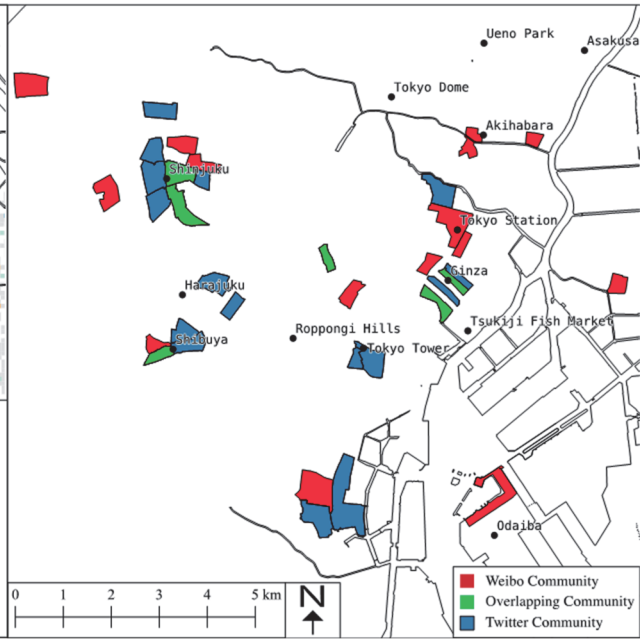

Figure.3: Spatial Comparison of Overlapping Communities.