Understanding and Recommending Green Spaces: Tapping the Power of Social Media Analytics and Artificial Intelligence

Text by Kwan Hui Lim

Green spaces are thought to improve the overall mental and physical well-being of residents in urban areas. For example, research have shown that green spaces provide the benefits of reducing stress levels [Tyrvainen et al., 2014] and myopia rates [Dadvand et al., 2017], while improving life expectancy [Takano et al., 2002] and thermal comfort [Wang et al., 2017]. Due to the numerous benefits of green spaces, the inclusion of green spaces is now a key consideration in many of today's urban development plans. The accessibility to green spaces is especially important today where more than half of the world are living in urban areas and rapidly increasing[1]. While we aim to improve accessibility to green spaces, how can we better measure the various effects that green spaces have on residents and better understand how green spaces are used? Traditional methods involve the use of surveys, questionnaires and case studies to answer these questions. The methods typically involve the participation of a small group of green space users, who are interviewed or asked questions relating to their use of green spaces.

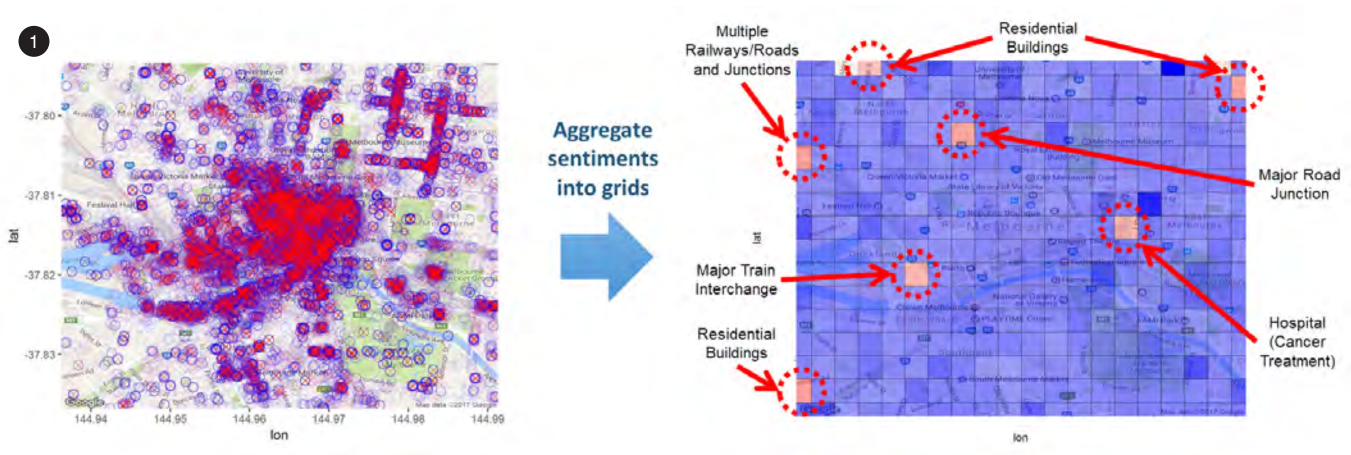
Apart from traditional surveys, questionnaires and case studies, the popularity and ubiquity of social media now provides another effective medium for conducting similar studies at a larger-scale and lower cost. Social media is now a prevalent part of our everyday life and oftentimes, the digital footprints left behind by users serve as a good reflection of their real-life activities. Moreover, the popularity of GPS-enabled smart phones have further enhanced the ubiquitous nature of social media. These digital footprints thus provide numerous opportunities for us to understand various real-life phenomenon at a large scale and in fine-grained details. Some of these phenomenon include understanding the effects of green spaces on its visitors and the different types of activities in green spaces.

This article aims to summarize some of our earlier works that utilize data mining and artificial intelligence techniques on geo-tagged social media data to: (i) understand the effects that green spaces have on the emotional well-being of residents [Lim et al., 2018]; (ii) discover popular activities and events that are undertaken in different types of parks [Lim et al., 2019]; and (iii) developing a recommender system for green spaces based on a user's activity preferences and sentiments [Wang et al., 2018].

Deriving Digital Footprints from Social Media

Twitter is a popular social networking site used by millions across the world. Twitter users are able to post short messages or tweets of up to 280 characters, which are shared with their followers. People routinely share numerous aspects of their daily life, ranging from their current activities to discussions on various topics. More importantly, users using GPS-enabled mobile devices can tag their tweets with their current location. In addition to these location information, each tweet is also tagged with the person who posted it and the time it was posted.

Our research involves the collection of 2.2 million public tweets that are tagged with such location information, which are collectively posted



1. An example showing individual tweet sentiments aggregated into small grids and the different facilities associated with each grid. Red grids refer to areas with negative sentiments and blue grids represent positive sentiments. Retrieved from Lim et al., 2018, 2019.

by more than 10k users in Melbourne, Australia. By mapping these tweets to locations of known green spaces and urban areas, we are able to determine the location that has been visited, the time it was visited and what was mentioned about that location.

Are people happier in green spaces or urban areas?

People convey various types of sentiments by their choice of different words, and this characteristic applies to tweets as well. Sentiment analysis techniques enable us to automatically determine the polarity of a sentence based on the words used. By applying these sentiment analysis techniques on tweet text, we can categorise each tweet into broad categories of positive, negative and neutral. Subsequently, we aggregate the polarity of all tweets posted from a specific location to obtain the overall polarity associated with that location, such as a green space or urban area. Our big-data analysis of 2.2 million tweets show that tweets in green spaces were more positive than those in urban areas.

Beyond simple polarity, each tweet can be further separated into specialized categories of anger, fear, sadness, disgust, surprise, anticipation, trust, and joy. Specifically, tweets in green spaces exhibit higher levels of anticipation, trust, and joy, and lower levels of anger and fear compared to those in urban areas. The benefits are not limited to visiting green spaces. Being in close proximity to a green space also has the effect of reducing negative sentiments such as anger and fear, although there are no clear trends in improving positive sentiments.

Furthermore, we identified various trends that demonstrate the positive effects of green space. For example, we observed that people decreased in positivity from mid-day before recovering at the end of the day, a trend that reflects daily work/school life. On a weekly basis, people are generally more positive on weekends than weekdays, with green spaces being consistently more positive than urban areas. People also exhibit more positivity during warmer months and less positivity during colder months in green spaces, but no clear trends was observed in urban areas.

Using a finer resolution of 250m grids, we can adapt our earlier analysis to study different sentiments associated with small pockets of areas. Figure 1 shows an example of this study where tweets sentiments are aggregated into individual grids, thus serving as a sentiment heatmap where red grids represent negative sentiments and blue positive sentiments. This heatmap shows that most negative sentiments are associated with residential areas or major transport infrastructure such as train stations and busy road junctions.

What do people do in green spaces?

To understand the popular activities in green spaces, topic modelling techniques are applied on tweet text to identify the main topics. These techniques examine the text used in each tweet in order to identify a set of representative keywords for a specific topic. In turn, these detected topics serve as a proxy for popular activities in the respective green spaces. For instance, in Carlton Gardens, this approach detected the events of GABS Beer, Cider & Food Fest (a food and drinks festival) and Brickvention (a convention and exhibit for Lego fans).

[1] https://esa.un.org/unpd/wup/publications/files/wup2014highlights.pdf

"

The study uses a big data approach, utilizing geo-tagged social media to understand the effects of green spaces on urban residents.

Combining this activity detection with our earlier sentiment analysis, we can now determine the sentiments associated with different activities in the specific green spaces. The positivity associated with green spaces also applies to most activities, much in line with our earlier analysis on green space in general. The only exception was for the Education activity in Flagstaff Gardens, a green space that is in close proximity to various educational institutes. We also observe that different green spaces are popular for different types of activities, e.g., some for social gatherings, others for sporting activities.

Which green space should I visit?

Extending the study of activities and sentiments across the year, we are able to understand how different green spaces are popular for different activities and how the popularity of these activities change throughout the year. Using this finding, we developed a recommender system that aims to make personalized recommendation of green spaces to users based their preferred activities, e.g. workout, relaxing or socializing, and the popularity of these activities at different times.

Figure 2 shows a screenshot of our recommendation system and an example of the recommended green spaces. To use this system, a user can provide his/her current location, preferred activity, time and travel distance. Using this information, the recommender system will examine its database of green spaces and recommend the top three choices based on the popularity and sentiment towards the selected activity and current distance from the user. The three recommended green spaces are listed based on their popularity and sentiment, as well as illustrated on the map. Users who do not want to enter their preferences can also get a default recommendation based on their current location.

Conclusion

In this article, we utilized a big data approach utilizing geo-tagged social media to understand the effects of green spaces on urban residents and their usage in terms of popular activities. Based on these findings, we developed a recommender system that offers personalized suggestions of suitable green spaces for specific activities based on the current context of the user. While our study focuses on green spaces in Melbourne, Australia, we believe that this approach is generalizable to other land use types as well as other cities with a sufficient social media usage rate.

This study contributes to the existing body of work on the benefits of green spaces. In particular, these findings have implications for urban planning authorities that aim to improve accessibility to green spaces and improve the overall well-being of urban residents. Future research can examine if certain negative effects in specific areas can be better mitigated. For example, could the decline in sentiments during colder months be due to falling leaves, and thus potentially be resolved by more maintenance and cleaning? Similarly, could negative sentiments associated with specific transport hubs be mitigated by encouraging the use of alternative routes or transport services?

References

[1] Kwan Hui Lim, Kate E. Lee, Dave Kendal, Lida Rashidi, Elham Naghizade, Stephan Winter and Maria Vasardani. The Grass is Greener on the Other Side: Understanding the Effects of Green Spaces on Twitter User Sentiments. Proceedings of the 2018 Web Conference Companion (WWW'18). Pg 275-282. 2018.

[2] Kwan Hui Lim, Kate E. Lee, Dave Kendal, Lida Rashidi, Elham Naghizade, Yungang Feng and Jia Wang. Understanding Sentiments and Activities in Green Spaces using a Social Data-driven Approach. Transitioning to smart cities: Mapping the Political, Economic, and Social Risks and Threats. Elsevier. 2019

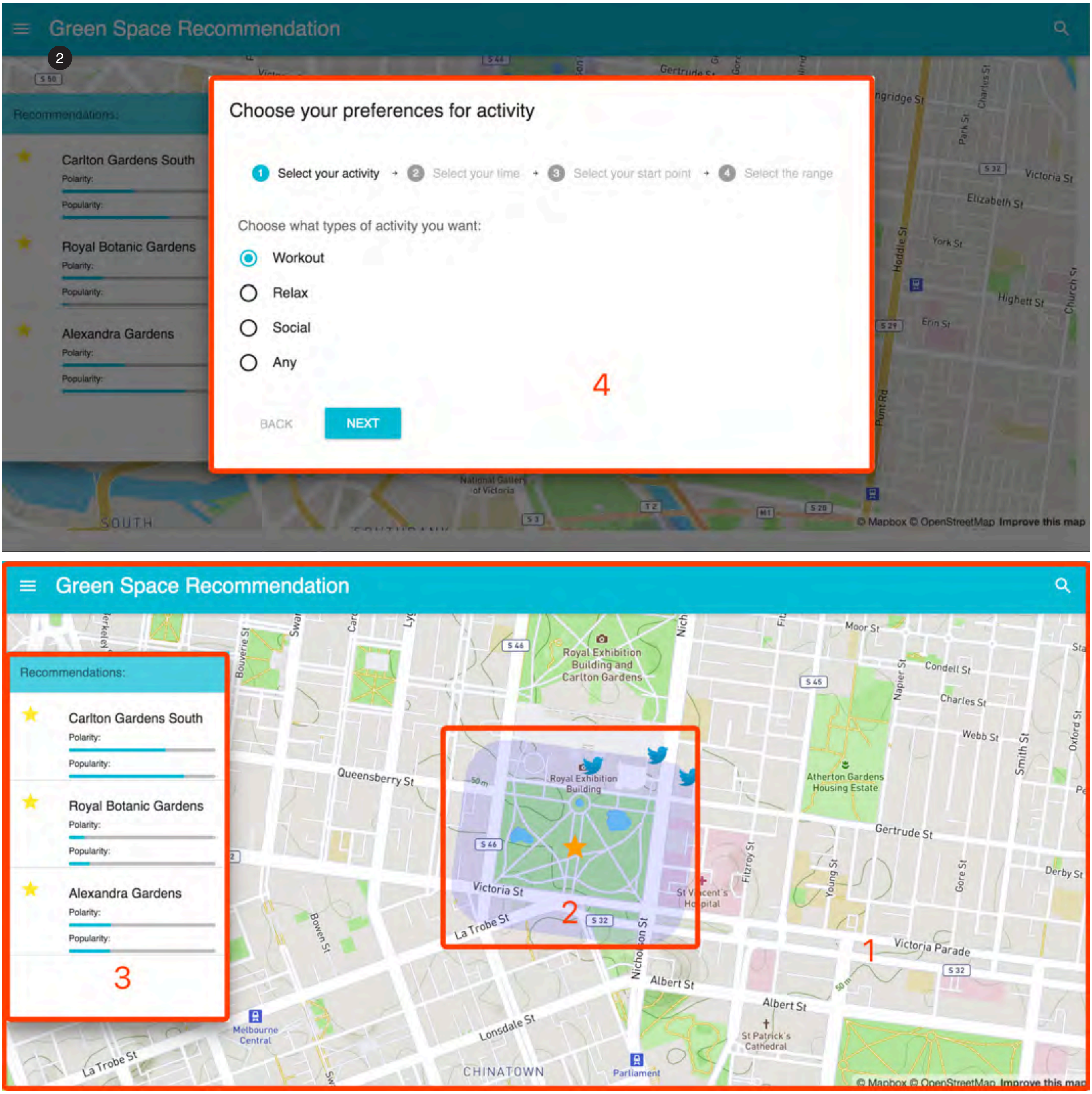
[3] Jia Wang, Yungang Feng, Elham Naghizade, Lida Rashidi, Kwan Hui Lim and Kate E. Lee. Happiness is a Choice: Sentiment and Activity-Aware Location Recommendation. Proceedings of the 2018 Web Conference Companion (WWW'18). Pg 1401-1405. 2018.

[4] Payam Dadvand, Jordi Sunyer, Mar Alvarez-Pedrerol, Albert Dalmau-Bueno, Mikel Esnaola, Mireia Gascon, Montserrat De Castro Pascual, Xavier Basagana, Ian G. Morgan, and Mark J. Nieuwenhuijsen. "Green spaces and spectacles use in schoolchildren in Barcelona." Environmental research 152 (2017): 256-262.

[5] Takehito Takano, Keiko Nakamura, and Masafumi Watanabe. "Urban residential environments and senior citizens' longevity in megacity areas: the importance of walkable green spaces." Journal of Epidemiology & Community Health 56, no. 12 (2002): 913-918.

[6] Liisa Tyrvainen, Ann Ojala, Kalevi Korpela, Timo Lanki, Yuko Tsunetsugu, and Takahide Kagawa. "The influence of urban green environments on stress relief measures: A field experiment." Journal of Environmental Psychology 38 (2014): 1-9.

[7] Yafei Wang, Rudolf de Groot, Frank Bakker, Heinrich Wörtche, and Rik Leemans. "Thermal comfort in urban green spaces: a survey on a Dutch university campus." International journal of biometeorology 61, no. 1 (2017): 87-101.



2. Screenshot of our green space recommendation system. The various components are: (1) input of user preference; (2) output of recommendations; (3) ranked list of recommended green spaces; and (4) location of top recommended green space. Retrieved from (Wang et al., 2018).