Demo Abstract: CitySense: A Data Collection Approach for City Computing Applications

Yue Mao, Zeyu Zhang, Hua Sun, Yang Chen
School of Computer Science, Fudan University, China
{ymao14,zeyuzhang15,hsun15,chenyang}@fudan.edu.cn

ABSTRACT

Today's cities are filled with sensors to get real-time data on all aspects of the city. Each of these data sources reflects one aspect of the city and provides support for city computing. However, these data are owned by different organizations or companies and stored in different locations of the Internet. It takes lots of efforts for researchers to obtained all of these data. At the same time, a single data set can only represent one side of a city's activities, but combining multiple data allows us to stand in a better perspective overlooking the city we study. We designed a city data aggregation platform to aggregate several data sources about city computing to get a more comprehensive understanding of the city, to discover pattern that are not reflected in a single data set, and facilitate researchers to access data from multiple sources.

CCS CONCEPTS

• **Human-centered computing**;

KEYWORDS

city computing, city sensing, data aggregation

ACM Reference Format:
Yue Mao, Zeyu Zhang, Hua Sun, Yang Chen. 2018. Demo Abstract: CitySense: A Data Collection Approach for, City Computing Applications. In *The 16th ACM Conference on Embedded Networked Sensor Systems (SenSys '18), November 4–7, 2018, Shenzhen, China.* ACM, New York, NY, USA, 3 pages. https://doi.org/10.1145/3274783.3275211

1 BACKGROUND AND MOTIVATION

Many organizations are collecting and publishing city-related data. In this paper, we take the city of Shanghai as an example.Take the road traffic index data released by the Shanghai Municipal Government as an example. Speed and traffic flow throughout Shanghai road network can be obtained through road cameras, ground coils, and GPS-equipped taxis all over the city to calculate the level of road congestion. Some online review services such as Dianping[1] contains lots of restaurants' data, the website shows the data of all restaurants including how customers feels about the environment, taste and service. People can also check the reviews of customers so that they can know the restaurant [1]. the weather data of Shanghai can also be obtained via website so that weather can be taken into account, which serves a significant role.

The platform can provide convenience in many aspects. First, different organizations focus on different areas. For instance, the Dianping regards more on the business activities in the city, while the Shanghai government publishes more data on urban infrastructure, transportation and other fields. One single data set is not enough to depict the integral view of city, which greatly reduces its value. For example, with the single traffic data set, we can only detect the congested area, but the cause of the congestion cannot be acquired. Combining public review data and urban POI data can better understand the underlying reason behind the phenomenon; Second, the data is owned by different organizations, and the barriers between data are so high that it is not easy to integrate multiple data sources [2]. For researchers, specific crawler needs to be implemented for each data set, which is very time-consuming. Therefore, we designed the city data aggregation platform to integrate data on all aspects of Shanghai and provide a download interface and data visualization functions to facilitate research in Shanghai.

This platform is mainly designed for researchers and the government. For researchers, they can download data easily from the platform, saving unnecessary effort. At the same time, the platform can be used to perform simple statistical analysis and visualization of data. In the early stages of research, researchers can rely on our platform to gain an intuitive understanding of the data, and draw some simple conclusions from it. By this they can be more focus on exploring complex urban patterns. The fusion of multiple data sets may also prompt researchers to find new research directions and achieve groundbreaking results. This platform also has great significance for the government [3]. It brings together the data generated in every corner of the city, and served just like the dashboard of a car, giving the government a more direct and in-depth understanding of the status of the city.

2 DESIGN AND IMPLEMENTATION

The city data aggregation platform is designed to integrate various data sources about Shanghai and present it to users in an intuitive way, so we decide to use the webpage as a carrier

[1] from https://www.dianping.com

SenSys '18, November 4–7, 2018, Shenzhen, China

ACM ISBN 978-1-4503-5952-8/18/11...$15.00
DOI: 10.1145/3274783.3275192

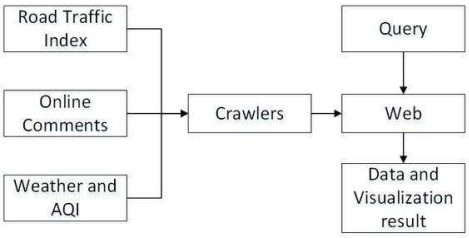

Figure 1: the overall architecture of our platform

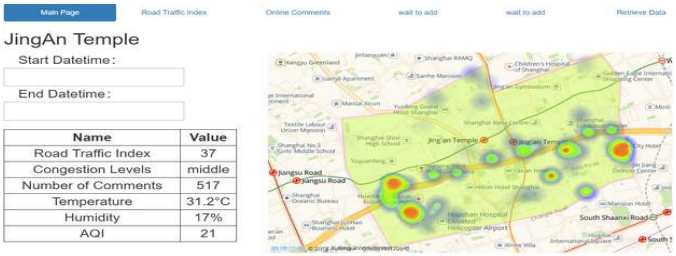

Figure 2: the UI of the website while user checking the region of Jing'an Temple

to make it more accessible [4]. Furthermore, we will open the data download interface for those who needs these data. The overall architecture of the platform is shown in Figure 1: Multiple crawlers are running on the back-end server continuously to collect data from data-sources. Then the data obtained will be cleaned and organized and finally published on our platform.When a query request is received, the corresponding data and visualization results will be returned. We offer two ways to view the data: 1) View detailed data of a certain area: We divide the city of Shanghai into 68 areas as the way in [2]. Users can select one of them to view detailed data of the area, including the road traffic index, online reviews of the shops in the area, and weather records for the area. Organizing all the data together can help us understand the area in all aspects, and it is also convenient for researchers to discover the connection between the data. 2) View data from a single data set: In this way, only data from a single data set is displayed to the user. Although connections between data sets are not included, users can check a specific type of data more easily. Users are allowed to set more complex data filtering conditions, display multiple regions at the same time, and get more detailed data visualization results.

Currently, the data integrated on the platform is as follows: 1) Road traffic index data: released by the Shanghai Municipal Government. The road traffic index of 68 areas in Shanghai has been recorded since July 19, 2018. 2) Online review data: Currently crawled from Dianping, a representative online review service in China. The data of other e-commerce platforms will continue to be integrated in the future. The reviews of the merchants under the Shanghai Food Classification starting from July 10, 2018 were recorded. 3) Weather data: Published by the Shanghai Meteorological Service, recorded real-time meteorological data (updated hourly) from various administrative districts in Shanghai starting from August 20, 2018, including temperature, humidity, visibility, and air quality index-AQI.

The city data aggregation platform operates a website built on the Flask framework. The data is stored in the MySQL database. As shown in Figure 2, users can smoothly switch between the main page and different data sets through the navigation bar at the top of the page. In the main page, the user views the data in the first way described above. Taking the region of Jing'an Temple as an example, the left side of the page displays various data about the region. Through the two date input boxes in the upper left corner of the page, the user can select the time period of the displayed data. On the right is a heatmap drawn on the map of Shanghai, which contains the distribution of all data on spatial scale. Users can easily drag or zoom in to view details. When switching to another dataset through the navigation bar, the user will view the data in the second way.

[2]from https://www.jtcx.sh.cn

3 CONCLUSIONS AND FUTURE WORK

In this paper, we build a city data aggregation platform. By using this platform, people can access the real-time information of different regions of the city of Shanghai from various aspects. We provide an intuitive interface for users to view the collected data in a convenient way. In the future, we plan to build some city computing-related applications on top of this platform. One possible application is to understand the difference between weekdays and weekends under the context of city computing. Also, we aim to explore the relationship between different types of collected data, for example, the connections between traffic congestion and the weather.

ACKNOWLEDGEMENT

This work is sponsored by National Natural Science Foundation of China (No. 61602122, No. 71731004), Natural Science Foundation of Shanghai (No. 16ZR1402200) and Shanghai Pujiang Program (No. 16PJ1400700). Yang Chen is the corresponding author.

REFERENCES

[1] E. Çelikten, G. Le Falher, and M. Mathioudakis. Modeling urban behavior by mining geotagged social data. *IEEE Transactions on Big Data*, 3(2):220–233, 2017.

[2] J. Koh, S. Sandha, B. Balaji, D. Crawl, I. Altintas, R. Gupta, and M. Srivastava. Data hub architecture for smart cities. In *Proceedings of the 15th ACM Conference on Embedded Network Sensor Systems*, page 77. ACM, 2017.

[3] T. H. Silva, P. O. Vaz de Melo, J. M. Almeida, J. Salles, and A. A. Loureiro. Revealing the city that we cannot see. *ACM Transactions on Internet Technology (TOIT)*, 14(4):26, 2014.

[4] S. Suakanto, S. H. Supangkat, R. Saragih, et al. Smart city dashboard for integrating various data of sensor networks. In *ICT for Smart Society (ICISS), 2013 International Conference on*, pages 1–5. IEEE, 2013.