Qurb: Qatar Urban Analytics

Sofiane Abbar, Laure Berti-Equille, Javier Borge-Holthoefer,
Sanjay Chawla, Hossam M. Hammady, Jaideep Srivastava
Qatar Computing Research Institute, HBKU, Qatar
{sabbar, lberti,jborge, schawla,hhammady, jsrivastava}@qf.org.qa

Doha is one of the fastest growing cities of the world with a population that has increased by nearly 40% in the last five years. There are two significant trends that are relevant to our proposal. First, the government of Qatar is actively engaged in embracing the use of fine-grained data to "sense" the city for maintaining current services and future planning to ensure a high standard of living for its residents. In this line, QCRI has initiated several research projects related to urban computing to better understand and predict traffic mobility patterns in the city of Doha [1]. Second trend is the high degree of social media participation of the populace, providing a significant amount of time-oriented social sensing of the all types of events unfolding in the city. A key element of our vision is to integrate data from physical and social sensing, into what we call *socio-physical sensing*. Another key element of our vision is to develop novel analytics approaches to mine this *cross-modal data* to make various applications for residents smarter than they could be with a single mode of data. The overall goal is to help citizens in their every-day life in urban spaces, and also help transportation experts and policy specialists to take a real time data-driven approach towards urban planning and real time traffic planning in the city.

Fast growing cities like Doha encounter several problems and challenges that should be addressed in time to ensure a reasonable quality of life for its population. These challenges encompass good transportation networks, sustainable energy sources, acceptable commute times, etc. and go beyond physical data acquisition and analytics.
In the era of Internet of Things [5], it has become commonplace to deploy static and mobile physical sensors around the city in order to capture indicators about people's behaviour related to driving, polluting, energy consumption, etc. The data collected from physical as well as social sensors has to be processed using advanced exploratory data analysis, cleaned and consolidated to remove inconsistent, outlying and duplicate records before statistical analysis, data mining and predictive modeling can be applied.

Recent advances in social computing have enabled scientists to study and model different social phenomena using user generated content shared on social media platforms. Such studies include the spread of diseases on social media [3] and studying food consumption in Twitter [4]. We envision a three layered setting: the ground, physical sensing layer, and social sensing layer. The ground represents the actual world (e.g., a city) with its inherent complexity and set of challenges. We aim at solving some of these problems by combining two data overlays to better model the interactions between the city and its population.

QCRI vision is twofold:

- From a data science perspective: Our goal is to take a holistic cross-modality view of urban data acquired from disparate urban/social sensors in order to *(i)* design an integrated data pipeline to store, process and consume heterogeneous urban data, and *(ii)* develop machine learning tools for cross-modality data mining which aids decision making for the smooth functioning of urban services;

- From a social informatics perspective: Use social data generated by users and shared via social media platforms to enhance smart city applications. This could be achieved by adding a semantic overlay to data acquired through physical sensors. We believe that combining data from physical sensors with user generated content potentially leads to the design of better and smarter lifestyle applications such as "evening out experience" recommenders that optimize for the whole experience including driving, parking and restaurant quality; Cab finder that takes into account the current traffic status, etc.

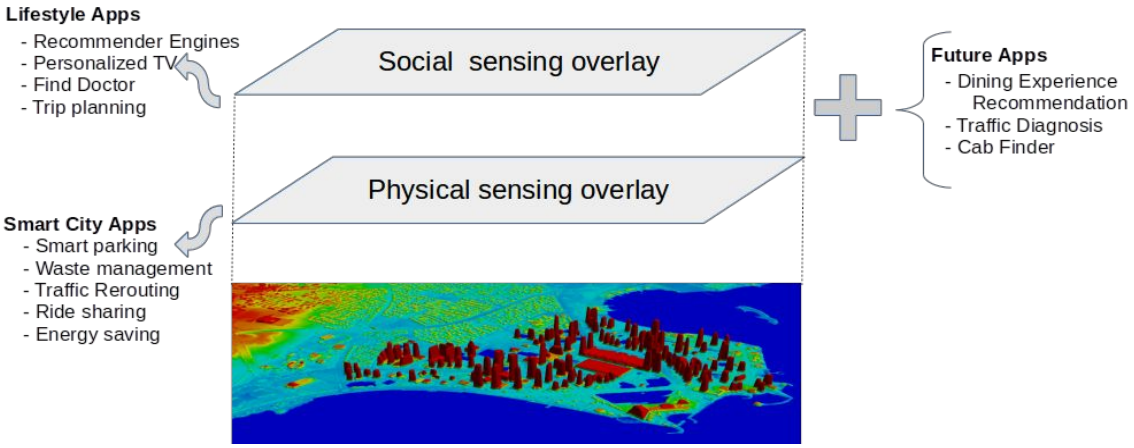

Figure 1. Overview of Proposed Approach.

In Figure 1 we provide a general overview of our cross-modality vision. While most of the effort toward building applications assisting people in their everyday life has focused on only one data overlay, we claim that combining the two overlays of data could generate a significant added value to applications on both sides.

References

1. Chawla, S., Sarkar, S., Borge-Holthoefer, J., Ahamed, S., Hammady, H., Filali, F., Znaidi, W., "On Inferring the Time-Varying Traffic Connectivity Structures of an Urban Environment", Proc. of the 4th International Workshop on Urban Computing (UrbComp 2015) in conjunction with KDD 2015, Sydney, Australia.
2. Sagl, G., Resch, B., Blaschke, T., "Contextual Sensing: Integrating Contextual Information with Human and Technical Geo-Sensor Information for Smart Cities". *Sensors* **2015**, *15*, 17013-17035.
3. Sadilek, A., Kautz, H. A., Silenzio, V.. "Modeling Spread of Disease from Social Interactions." *ICWSM*. 2012.
4. Sofiane Abbar, Yelena Mejova, and Ingmar Weber. 2015. You Tweet What You Eat: Studying Food Consumption Through Twitter. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (CHI '15). ACM, New York, NY, USA, 3197-3206.
5. Atzori, L., Iera, A., Morabito, G.. "The internet of things: A survey." *Computer networks* 54.15 (2010): 2787-2805.