Crowd Sensing System for Public Participation

M. Tenney[1], R. Sieber[1], G.B. Hall[2]

1 McGill University, Montreal, QC, Canada
Email: matthew.tenney@mail.mcgill.ca, renee.sieber@mcgill.ca

2 Esri Canada, Toronto, ON, Canada
Email: bhall@esri.ca

Abstract

We propose a crowd sensing system to capture certain dynamics of public participation in a city. Crowd sensing systems (CSS) attempt to capture the opinions of local publics from web-resources. We define our CSS using a spatially-situated social network graph where users along with different variables, such as time, location, social interaction, service usage, and human activities can be studied and used to identify experts or influential citizens who are relevant to municipal affairs.

1. Introduction

In the fields of engineering and the computational sciences, the term crowd sensing represents a popular area of research (Cardone et al. 2013). Similar to citizen sensing, urban sensing and participatory sensing, we broadly define crowd sensing systems (CSS) as being an integrated hardware and software architecture designed to collect user-generated content for a specified topic, issue or theme. In this paper, we introduce a portion of our conceptual CSS, which describes several social and spatial interactions within a local population (i.e., connections between individuals and locations of communities-of-interest), establishes place-based topics across a city from user-generated content (e.g., geotagged posts from social media), and identifies various forms of activity across specific geographies (e.g., patterns of urban travel). The CSS combines methods of natural language processing, spatial analysis, and graph theory to create a data structure with possible value when used to inform local decision makers.

Our work builds on smart city initiatives, data-science and Web 2.0 literature that seek to revise traditional forms of public participation (Cardone et al. 2013). In particular, these difficulties can be assuaged by integrating data-driven techniques that automatically extract "similar" information (i.e., topical pubic opinions) from user-generated content.

2. Crowd Sensing Systems as Tripartite Network

Public participation in municipal affairs is often seen as a product of stakeholders and interest groups, which are spatially distributed across a city. Choosing to model public participation digitally, requires representing relationships between structured and unstructured content, deriving explicit and implicit social interactions, and inferring frames of context through shared interests and co-location. Like other social networks, our CSS network graph (G) contains nodes and edges $G = (N, E)$. The graph is further divided into three subgraphs containing unique node and edge types defined as $U = (Un, Ue)$, $C = (Cn, Ce)$ and $T = (Tn, Te)$, where Un, Ue are nodes and edges of the user profiles in subgraph U; user-generated content forms the subgraph C consisting of nodes Cn and edges Ce; and a "geotopics" subgraph T contains spatially located nodes Tn with temporally weighed edges Te. These connected subgraphs as shown in Figure 1, can form spatially-situated networks constructed from what we call Social Signals (SocSigs, see below). SocSigs information contained in each of the connected sub-graphs include a user-network U (i.e., social-network among citizens with ties representing relationships from social affiliations or topical/interest similarities), content-network *C* (i.e., unstructured text) and a topic-network *T* (i.e., content nodes derived from secondary data). This latter subgraph contains geographical ties to relevant locations within the city, as well as two extra sets of bridge edges representing semantic similarity to data in the content network and/or related to the interests of users in the user network.

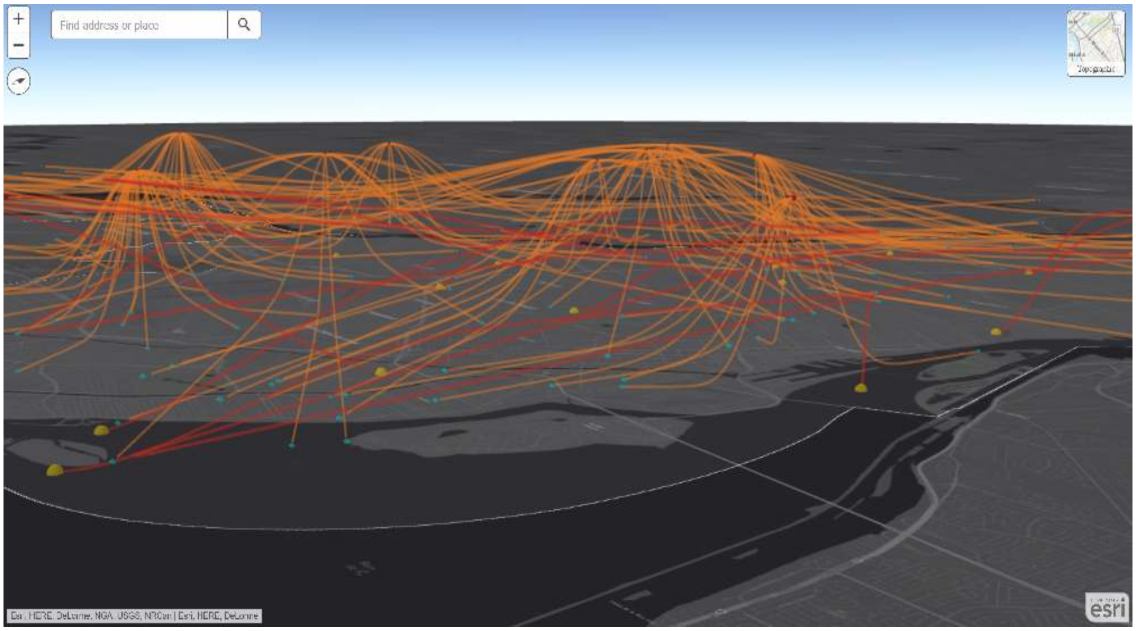

Figure 1 Example of CSS data-structure situated over Montreal, QC. Red nodes are Users in subgraph *U*, with content connections to subgraph *C* (blue) shown in orange, and aggregated 'geotopic" nodes of *T* seen in yellow.

All edges within a single sub-graph are undirected. Bridge edges connecting different sub-graphs are defined as directed-edges so to limit the connectivity of *G.* Using both undirected and directed edges allows for calculating different network measures on specific graph components (De Meo et al. 2014). For example, network communities can be found either by using only connections in the *U* sub-graph or by including all "in-edges" to the content-graph *C* as available paths between users. A mixed-edge graph design containing undirected and directed edges should provide a network-structure that is sensitive to variances in interaction-flows at both local and global levels (De Meo, Ferrara, Fiumara, & Provetti, 2014).

2.1 Relevant Social Signals: A Balance of Context and Content

SocSigs are informative signals that can directly or indirectly provide contextual meaning for interactions, relationships, and behaviours observed from user-generated content (Sheth 2009; Golbeck 2013). We chose four SocSigs variables relevant to our efforts of capturing interests held by a local network members. The first is *content*, which contains unstructured text (e.g., social media posts and status updates). The second is *users*, who are seen as the producers and consumers of the content. We include available characteristics about them, as well as connections to other people and content (e.g., "friends" or "followers", "Likes" or "Shares"). The third is *space-time*, which joins location and time attached to a collected dataset where geotagged content is collected at (x,y). In this instance, the locational content is time stamped with the time it is created. Fourth, is *strength* which includes the frequency of spatial and social connections among different users, their locations, and topics found in their content.

2.2 Citizen Centrality and Social Influence

Centrality, in graph theory, is a measure of a node's importance to the structure of a graph (Scott and Carrington 2011). Centrality has been used for identifying influential people (i.e., opinion leaders) in various kinds of social networks (Golbeck 2013). It is relevant to public participation because participatory activities can be understood, in part, as multiple actors interacting in relational systems. The centrality of the interactions about a user-node within a network resembles concepts found in participation literature like "opinion-leaders", gatekeepers, and stakeholders (Dubois and Gaffney 2014).

Centrality measures within our CSS represent the importance of a node in each of the sub-graphs without the inclusion of bridge links to other graph components. We use centrality to find salient content, users, and locations. The more central a node is, the more it can be said to represent important topics or people relative to other topics and people. Node influence determines how a combination of connections both within a sub-graph and to other components represent the leading topics important to citizens and opinion-leaders (Dubois and Gaffney 2014), and provides a means to estimate how the opinions of one user may affect the views of others.

2.3 Community Detection and Description

Community is often viewed as possessing a certain physicality, for example, a jurisdictional bounding of city blocks that comprise a neighbourhood. Individuals in this CSS can become community members by expressing shared interests, behaviours, and affiliations throughout the evolution of a network (e.g., increased social similarity between content or co-location patterns). In social network analysis, a community is a set of nodes with strong connections and that contain frequent interactions between members (Fortunato 2010). Community detection in our CSS attempts to decompose a complex network into groups of nodes and edges that are densely connected. Using either direct edge-connections or by including similarity measures between graph-components, grouped elements are considered to have similar interests including topics, activities, and locations (Clauset, Newman, and Moore 2004). Communities can overlap if they comprise a threshold number of nodes or edges that are members of two or more communities.

3. Conclusion

Automatically connecting citizens and governments is a form of automated public participation, which can be provided by a CSS (Cardone et al. 2013). We see a CSS as a computational instrument, composed of computers, sensors, software and algorithms. The instrument can automatically harvest posts, locations, times, and connections among streams of citizen's data to derive insights on the public intent. This process aligns with a big data and smart city vision as data provide access to localized "Citizen Sensor Networks" to computationally facilitate public participation (Koch et al. 2013). Future work will be the critical investigation of these systems, the algorithms applied to these network structures and datasets, and the implications of transitioning to a "coded form" of public participation.

Acknowledgements

We are grateful for the support from the following funders: SSHRC grant 895-2012-1023 "How the geospatial web 2.0 is reshaping government-citizen interactions" and Mitacs Accelerate PhD fellowship co-funded by Esri Canada Limited.

References

Burke, J. A., Estrin, D., Hansen, M., Parker, A., Ramanathan, N., Reddy, S., & Srivastava, M. B. (2006). Participatory sensing. Center for Embedded Network Sensing.

Cardone, G., Foschini, L., Bellavista, P., Corradi, A., Borcea, C., Talasila, M., & Curtmola, R. (2013). Fostering participaction in smart cities: a geo-social crowdsensing platform. IEEE Communications Magazine, 51(6), 112–119.

Clauset, A., Newman, M. E. J., & Moore, C. (2004). Finding community structure in very large networks. Physical Review E, 70(6), 66111.

De Meo, P., Ferrara, E., Fiumara, G., & Provetti, A. (2014). Mixing local and global information for community detection in large networks. Journal of Computer and System Sciences, 80(1), 72–87.

Dubois, E., & Gaffney, D. (2014). The Multiple Facets of Influence Identifying Political Influentials and Opinion Leaders on Twitter. American Behavioral Scientist.

Fortunato, S. (2010). Community detection in graphs. Physics Reports, 486(3–5), 75–174.

Gabrys, J. (2014). Programming Environments: Environmentality and Citizen Sensing in the Smart City. Environment and Planning D: Society and Space, 32(1), 30–48.

Golbeck, J. (2013). *Analyzing the social web* (First edition). Waltham, MA: Morgan Kaufmann is an imprint of Elsevier.

Koch, F., Cardonha, C., Gentil, J. M., & Borger, S. (2013). A Platform for Citizen Sensing in Sentient Cities. In J. Nin & D. Villatoro (Eds.), Citizen in Sensor Networks (pp. 57–66).

Scott, J., & Carrington, P. J. (2011). The SAGE Handbook of Social Network Analysis. SAGE Publications.

Sheth, A. (2009). Citizen Sensing, Social Signals, and Enriching Human Experience. IEEE Internet Computing, 13(4), 87–92.