AMPLIFYING CITIZENS' VOICES IN SMART CITIES

An Application of Social Media Sentiment Analysis in Urban Sciences

IMAN SAYAH¹ and MARC AUREL SCHNABEL²

^{1,2}*Victoria University of Wellington*

¹*iman.sayah@gmail.com* ²*MarcAurel.Schnabel@vuw.ac.nz*

Abstract. In the past decade, urban researchers have paid significant attention to the emergence of computer science and urban planning. According to the literature, social media as a pool of real-time citizen feedback can be investigated to inform smart city synergies. However, the success factors of such an approach have not been thoroughly investigated. In this study, various factors were derived from an extensive literature review to create an efficient e-participation platform. It is explained how our proposed platform 1) complies to the data protection regulations 2) uses advanced text analysis and natural language processing (NLP) tools to identify opinions and emotions 3) maintains persistent communication between citizens and city planners 4) incorporates creative visualisation techniques 5) is informative for its target audience 6) takes into consideration the socio-cultural diversity and 7) can be used as an informing tool in combination with offline methods of participation.

Keywords. Sentiment analysis; e-participation; social media mining; big data analytics; citizen engagement.

1. Social media mining in smart cities

Social media provides a valuable resource of real-time data and has become the centre of attention in a wide range of decision-making studies (Lin et al., 2017). Researchers have investigated the potentials in collecting digital feedback from the community by practising various digital techniques. A review of the case studies shows that social media analysis seems to be the dominant approach in e-participation initiatives (Boudjelida et al., 2016). The emergence of cell phones and geo-based social media significantly facilitate a better understanding of the community's behaviour as these types of social media are supported by users' geo-located data. Considering this approach, some experts have highlighted the role of social media analysis in urban planning. Social media networks have been investigated by researchers to identify their possible potentials for providing meaningful feedback to urban planners (Zhang and Feick, 2016, Leszczynski, 2016, DiSalvo and Jenkins, 2017).

The literature shows the importance of investigating the limitations and potentials to social data mining and various techniques have been used to evaluate

the success of these projects. They have mostly followed a case study approach to investigate the significant impact of social media apps such as Facebook, Twitter, and Instagram on public participation (Schafer, 2012, Moser, 2016, Swim and Bloodhart, 2015). Social media has proven to be useful even in a broader range of fields. For instance, social media as “*a database of electronic word of mouth*” (Chen and Yang, 2014) has proven to be an effective media for participation in political involvement (Rahimi, 2011, Gregory, 2013, Olorunnisola and Martin, 2013, Alkazemi et al., 2014).

The objective of this research is to create a smart application for the administrative bodies to evaluate the impact of their decisions and actions by analysing citizens’ emotional feedback on Twitter. The novelty of this approach is in the implementation of a real-time timeline of events rather than classification and correlation maps proposed in previous studies (Jongman et al., 2015, Himmelboim et al., 2017, Grandjean 2016). It is explained how the timeline visualisation of the sentiments provides an effective way of comparing actions with people’s opinions and as a result proves to be a more engaging and potent tool for social media analysis.

2. Method

The following steps were used to identify the most relevant research in this area. First, the main elements of the study were used as search terms. *Scopus* and *Web of Knowledge* were chosen as the scientific data repositories for conducting this literature review. The abstracts of the search results were examined based on two main factors: 1) Inclusion of a case study and 2) Indication of the contributing factors to the success or failure of the case study. The search terms were revised and fed into the first step (search in data repositories). As a result, the most relevant papers were derived using the following combination of terms:

```
eparticipation AND ("smart city" OR "urban planning")
AND ("social media" OR Facebook OR Twitter)
```

The full text of the final selected papers was studied for the identification of the success factors.

3. Success factors

3.1. ETHICS

Privacy concerns might prevent stakeholders from accepting ubiquitous monitoring initiatives (Shilton, 2012, Ertio and Bhagwatwar, 2017). While concerns about privacy are the most common issue in public data mining, there seems to be a sufficient set of regulations and conducted research on this topic (Townsend and Wallace, 2017). The EU GDPR (European General Data Protection Regulations) along with DPA (Data Protection Act) has undergone a

series of changes in 2018 which indicates significant attention to user's privacy in data science. According to the EU GDPR (2018), "Personal data is any information that relates to an identified or identifiable living individual. Different pieces of information, which collected together can lead to the identification of a particular person, also constitute personal data." As a result, our innovative approach does not disseminate any form of entity (piece of text, image, video, or hashtag) that would result in the identification of a person or a Twitter profile.

Figure 1. The sentiment timeline only records the number of positive (green), negative (red) or neutral (grey) tweets on a specific topic without revealing an individual's identity.

3.2. TECHNOLOGY

The data used for this research were obtained by connecting to the Twitter API while a suite of data analysis and visualisation tools were used to stream, pre-process, analyse and visualise data (JONES et al., 2001, McKinney, 2010).

Natural Language Processing (NLP) is an attempt to process natural language with the help of a computer. It may be used to count single word or multiple words occurrences to identify topics, removing the words that do not directly pertain to the intention of the statement (known as stop-words), or implications on the emotions that a written piece might hold (Bird et al., 2009). The latter extent has led to the establishment of a technique in text mining which is known as opinion mining or sentiment analysis. "*Sentiment analysis or opinion mining is the computational study of people's opinions, appraisals, attitudes, and emotions toward entities, individuals, issues, events, topics, and their attributes.*" (Liu and Zhang, 2012).

Sentiment analysis allows for analysing a large body of uploaded text by the citizens and deriving valuable measures based on public feedback. Researchers have proposed an assessment framework for translating unstructured online public opinions into sentimental and topical indicators (Jiang et al., 2016, McGetrick et al., 2017). In similar studies, geo-located text from Twitter was analysed to identify public signals related to urban subjects (Zhang and Feick, 2016, Jongman et al., 2015, Panteras et al., 2015, Anselin and Williams, 2016, Shelton,

2017). Additionally, sentiment analysis enables users to keep track of community emotions.

Figure 2. A streaming timeline showing the result of sentiment analysis of tweets on traffic in Wellington NZ during a week.

Various supervised algorithms such as Support Vector Machine (SVM) and Artificial Neural Networks (ANN) have been used as a sentiment mining technique. According to the literature, SVM has been used widely by researchers and yielded acceptable results (Moraes et al., 2013).

3.3. PERSISTENT COMMUNICATION

E-participation brings the opportunity to the city to collect voluntarily generated feedback from citizens. However, participants have not always been involved in every stage of an urban initiative. Keeping the community engaged at every step of the project seems to yield better results (Commodore et al., 2017). Therefore, a wide range of innovative technologies has been incorporated in smart city initiatives to facilitate interaction but was not well received by the people due to inaccessibility or climate issues (Ylipulli et al., 2014). In particular, social media allows for persistent communication as well as convenient access by a large number of stakeholders as they are widely available on smartphones, tablets, TVs, and PCs.

3.4. CREATIVITY AND PLAYFULNESS

A key element here is to provide a user-friendly interactive medium to encourage public engagement. Visual toolkits have the potential to improve communication in value-embedded participatory urban projects (Leslie et al., 2007, Forlano and Mathew, 2014, Fechner and Kray, 2016). A common approach to visualise the relationship between tweets is using network graphs (Grandjean, 2016, Himmelboim et al., 2017). A network graph consists of nodes which are represented in circles in various radii and colours. Nodes are connected with one/two-headed arrows which indicate the subjectivity or objectivity of the nodes. Size of each circle usually identifies the larger quantity or, the higher quality of a feature attached to the nodes.

Figure 3. Mapping the social affinities of a group of individuals, Moreno (1934)'s first sociograms visualise the relationships between pupils in a classroom: who wants to be sitting next to whom? Adapted from Grandjean (2015) Social network analysis and visualisation: Moreno's Sociograms revisited. Source: Wikipedia (2018).

Ball-Rokeach et al. (2001) have proposed that storytelling as a concept uses ubiquitous data to let the community share their concerns and issues about their neighbourhood via a communication infrastructure. While graph networks provide an informative overview of the relationships between pieces of data, they are not suitable for representing a story. Having that said, a suite of various visualisation types (interactive timeline, word cloud, search bar) have been used to create the proposed interface. The Actions timeline at the bottom of the Sentiment timeline contains a history of actions of a decision-making organisation. For example, the administrative body may upload their record of actions on 'traffic' and then observe public feedback on 'traffic' through time. A possible insight from observing this timeline would be the fact that after the proposal of Plan X.3.33, the number of positive posts has decreased and more negative feedback has been posted by the citizens.

Figure 4. Sentiment tab of the developed interface - Interactive timeline.

3.5. TECHNOLOGY EDUCATION

The level of familiarity with the technologies used in smart city initiatives affects the acceptance of such projects (Ylipulli et al., 2014). Thus, the appropriateness of the e-participation method may decrease if there is no education involved in the early steps of the project. Using education as an incentive not only boosts the level of participation but also results in a greater civic virtue (Kelty et al., 2014). The familiarity of the urban dwellers with social media in this context reduces the need for conducting training sessions in smart city projects as the number of social media users is increasing rapidly. Compared to IoT devices and AR/VR technology, social media stands out as a cost-effective e-participation strategy.

3.6. SOCIO-CULTURAL DIVERSITY

Community members have diverse age, sex and cultural backgrounds. Taking into account all the various mindsets and beliefs they bring can contribute to the effectiveness of urban development strategies (Gudowsky et al., 2017). Also, web applications can decrease power dynamics as they provide the opportunity for every individual to share their opinions without the fear of speaking in front of a large audience (MacMillan; Klenk and Hickey, cited in Pagatpatan and Ward 2017). Young people seem to construct 60% of the urban population in future (“The Role of Youth in the Urban Future”, cited in Rexhepi et al., 2016). This projection has led to the investigation of motivating factors to encourage youth participation in urban initiatives (Shakeri et al., 2015, Rexhepi et al., 2016).

3.7. ONLINE AND OFFLINE METHODS

There is evidence that transitioning from offline to online methods has facilitated citizen engagement in the smart city processes. Therefore, the internet can be a promising tool for attracting a large number of participants in our cities. (Spil et al., 2017). The simultaneous use of online and offline methods has been suggested in cases where the project deals with people who do not have the required technical skills (Kleinhans et al., 2015).

4. Conclusion

In this research, NLP tools and creative visualisation techniques were employed to create a sentiment analysis interface for urban planning. While privacy concerns might discourage using social media mining, studies showed that there are enough guidelines and regulations for ethical use of online public feedback. The EU GDPR (2018) and the Twitter Developer Agreement, along with the instructions from the literature, enabled reading, analysing and visualising public data without breaching the users’ privacy. While there is ongoing research on the efficiency of opinion mining techniques, our approach revealed that a widely-used method (SVM) is adequately capable of detecting the emotion of tweets. When a new e-participation platform is introduced, stakeholders need instructions to interact with the new system. However, using a popular social media network like

Twitter decreases the learning time because it is user-friendly and easy to learn. Twitter also allows users of diverse backgrounds and ages to engage in city decision-making actively. While e-participation may discourage using offline methods such as face-to-face meetings or workshops, it seems to work effectively as a complementary tool to add to the potentials of these methods.

4.1. FUTURE STEPS

There are various elements embedded in the Twitter interface that allows for the evaluation of the contents of a tweet. The most common Twitter elements are: 1) Hashtags 2) Mentions 3) Likes 4) Retweets and 5) Replies. Depending on how much users want to engage in a conversation they can choose either to 1) Like 2) Retweet (with/without text) and 3) Reply to a tweet. To further our research we are planning to work on calculating the weight of each of these reactions since it can simplify the comparison of various tweets, e.g. a tweet with ten likes and a tweet with one retweet and one reply. How much a tweet has been engaging (E) can then be calculated by using an average weight (Equation 1).

$$E = \alpha \cdot lk + \beta \cdot rt + \gamma \cdot rp \quad (1)$$

where lk is the number of likes, rt is the number of retweets, and rp is the number of replies. The constants α , β and γ determine the weight of each three interactions. Additionally, the impact of retweets with text shows a greater amount of involvement than the retweets without any opinions attached to them. Thus, rt can be separated into two elements rt and rt' which leads to (Equation 2)

$$E = \alpha \cdot lk + \beta \cdot rt + \beta' \cdot rt' + \gamma \cdot rp \quad (2)$$

where rt is the number of retweets with text, and rt' is the number of retweets without text and the weight factors are α , β , β' and γ .

References

- “UK Data Protection Act” : 2018. Available from <http://www.legislation.gov.uk/ukpga/2018/12/pdfs/ukpga_20180012_en.pdf> (accessed December 2018).
- “Twitter Developer Agreement” : 2018. Available from <<https://developer.twitter.com/en/developer/terms/agreement-and-policy.html>> (accessed May 2018).
- Alkazemi, M., Bowe, B. and Blom, R. 2014, Facilitating the Egyptian uprising: A case study of facebook and Egypt, in U.S.A. Information Resources Management Association (ed.), *Human Rights and Ethics: Concepts, Methodologies, Tools, and Applications*, IGI Global, 27.
- Anselin, L. and Williams, S.: 2016, Digital neighborhoods, *Journal of Urbanism: International Research on Placemaking and Urban Sustainability*, **9**(4), 305-328.
- Ball-Rokeach, S.J., Kim, Y.C. and Matei, S.: 2001, Storytelling Neighborhood: Paths to Belonging in Diverse Urban Environments, *Communication Research*, **28**(4), 392-428.
- Bird, S., Klein, E. and Loper, E.: 2009, *Natural Language Processing with Python*, O'Reilly

- Media, Inc., Sebastopol, USA.
- Boudjelida, A., Mellouli, S. and Lee, J.: 2016, Electronic Citizens Participation: Systematic Review, *9th International Conference on Theory and Practice of Electronic Governance (Icegov 2016)*, **9**, 31-39.
- Chen, L.S. and Yang, T.Y.K.: 2014, Increasing Customer Loyalty in Internet Marketing, *Intelligent Data analysis and its Applications*, **II**, 95-103.
- Commodore, A., Wilson, S., Muhammad, O., Svendsen, E. and Pearce, J.: 2017, Community-based participatory research for the study of air pollution: a review of motivations, approaches, and outcomes, *Environmental Monitoring and Assessment*, **189**, 378.
- DiSalvo, C. and Jenkins, T.: 2017, Fruit Are Heavy: A Prototype Public IoT System to Support Urban Foraging, *Proceedings of the 2017 Conference on Designing Interactive Systems*, **17**, 541-553.
- Ertio, T.P. and Bhagwatwar, A.: 2017, Citizens as planners: Harnessing information and values from the bottom-up, *International Journal of Information Management*, **37**(3), 111-113.
- Fechner, T. and Kray, C.: 2016, Presenting Citizen Engagement Opportunities Online: The Relevancy of Spatial Visualization, *Geospatial Data in a Changing World*, **19**, 105-121.
- Forlano, L. and Mathew, A.: 2014, From Design Fiction to Design Friction: Speculative and Participatory Design of Values-Embedded Urban Technology, *Journal of Urban Technology*, **21**(4), 7-24.
- Grandjean, M.: 2016, A social network analysis of Twitter: Mapping the digital humanities community, *Cogent Arts & Humanities*, **3**, 1-14.
- Gregory, D.: 2013, Tahrir: Politics, Publics and Performances of Space, *Middle East Critique*, **22**(3), 235-246.
- Gudowsky, N., Sotoudeh, M., Capari, L. and Wilfing, H.: 2017, Transdisciplinary forward-looking agenda setting for age-friendly, human centered cities, *Futures*, **90**, 16-30.
- Himmelboim, I., Smith, M.A., Rainie, L., Shneiderman, B. and Espina, C.: 2017, Classifying Twitter Topic-Networks Using Social Network Analysis, *Social Media + Society*, **3**(1), 1-13.
- Jiang, H.C., Qiang, M.S. and Lin, P.: 2016, Assessment of online public opinions on large infrastructure projects: A case study of the Three Gorges Project in China, *Environmental Impact Assessment Review*, **61**, 38-51.
- Jones, E., Oliphant, T. and Peterson, P.: 2001, "SciPy: Open Source Scientific Tools for Python". Available from <<https://www.scipy.org/>> (accessed 10th May 2018).
- Jongman, B., Wagemaker, J., Romero, B. and de Perez, E.: 2015, Early Flood Detection for Rapid Humanitarian Response: Harnessing Near Real-Time Satellite and Twitter Signals, *ISPRS International Journal of Geo-Information*, **4**(4), 2246.
- Kelty, C., Panofsky, A., Currie, M., Crooks, R., Erickson, S., Garcia, P., Wartenbe, M. and Wood, S.: 2014, Seven dimensions of contemporary participation disentangled, *Journal of the Association for Information Science and Technology*, **66**, 474-488.
- Kleinhans, R., Van Ham, M. and Evans-Cowley, J.: 2015, Using Social Media and Mobile Technologies to Foster Engagement and Self-Organization in Participatory Urban Planning and Neighbourhood Governance, *Planning Practice and Research*, **30**(3), 237-247.
- Leslie, E., Coffee, N., Frank, L., Owen, N., Bauman, A. and Hugo, G.: 2007, Walkability of local communities: Using geographic information systems to objectively assess relevant environmental attributes, *Health & Place*, **13**(1), 111-122.
- Leszczynski, A.: 2016, Speculative futures: Cities, data, and governance beyond smart urbanism, *Environment and Planning A*, **48**(9), 1691-1708.
- Lin, Y.P., Lin, W.C., Li, H.Y., Wang, Y.C., Hsu, C.C., Lien, W.Y., Anthony, J. and Petway, J.R.: 2017, Integrating Social Values and Ecosystem Services in Systematic Conservation Planning: A Case Study in Datuan Watershed, *Sustainability*, **9**(5), 718.
- Liu, B. and Zhang, L.: 2012, A Survey of Opinion Mining and Sentiment Analysis, in C. Aggarwal and C. Zhai (eds.), *Mining Text Data*, Springer, Boston, MA, 415-463.
- McGetrick, J.A., Bubela, T. and Hik, D.S.: 2017, Automated content analysis as a tool for

- research and practice: a case illustration from the Prairie Creek and Nico environmental assessments in the Northwest Territories, Canada, *Impact Assessment and Project Appraisal*, **35**(2), 139-147.
- McKinney, W.: 2010, Data Structures for Statistical Computing in Python, *Proceedings of the 9th Python in Science Conference (SciPy 2010)*, 51-56.
- Moraes, R., Valiati, J.F. and Neto, W.P.G.: 2013, Document-level sentiment classification: An empirical comparison between SVM and ANN, *Expert Systems with Applications*, **40**(2), 621-633.
- Moser, S.C.: 2016, Reflections on climate change communication research and practice in the second decade of the 21st century: what more is there to say?, *Wiley Interdisciplinary Reviews-Climate Change*, **7**(3), 345-369.
- Olorunnisola, A.A. and Martin, B.L.: 2013, Influences of media on social movements: Problematising hyperbolic inferences about impacts, *Telematics and Informatics*, **30**(3), 275-288.
- Pagatpatan, C.P. and Ward, P.R.: 2017, Understanding the factors that make public participation effective in health policy and planning: a realist synthesis, *Australian Journal of Primary Health*, **23**, 516-530.
- Panteras, G., Wise, S., Lu, X., Croitoru, A., Crooks, A. and Stefanidis, A.: 2015, Triangulating Social Multimedia Content for Event Localization using Flickr and Twitter, *Transactions in Gis*, **19**(5), 694-715.
- Rahimi, B.: 2011, The Agonistic Social Media: Cyberspace in the Formation of Dissent and Consolidation of State Power in Postelection Iran, *The Communication Review*, **14**(3), 158-178.
- Regulation, E.U.: 2018, 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), *Official Journal of the European Union*, **L119**, 1-88.
- Rexhepi, A., Filiposka, S. and Trajkovik, V.: 2016, Playful e-participation with minecraft as development tool for urban redesign: a case study, *IADIS International Conference Big Data Analytics, Data Mining and Computational Intelligence 2016 (part of MCCSIS 2016)*, Funchal, Portugal, 49-56.
- Schafer, M.S.: 2012, Online communication on climate change and climate politics: a literature review, *Wiley Interdisciplinary Reviews-Climate Change*, **3**(6), 527-543.
- Shakeri, M., Kingston, R. and Pinto, N.: 2015, Towards a Community Support System: Social Media Culture, Games, and Planning Tools, *Digital Placemaking: Augmenting Physical Places with Contextual Social Data: Papers from the 2015 ICWSM Workshop*, 18-22.
- Shelton, T.: 2017, Spatialities of data: mapping social media 'beyond the geotag', *GeoJournal*, **82**(4), 721-734.
- Shilton, K.: 2012, Participatory personal data: An emerging research challenge for the information sciences, *Journal of the American Society for Information Science and Technology*, **63**(10), 1905-1915.
- Spil, T.A., Effing, R. and Kwast, J.: 2017, Smart City Participation: Dream or Reality? A Comparison of Participatory Strategies from Hamburg, Berlin & Enschede, *Digital Nations – Smart Cities, Innovation, and Sustainability, IFIP International Federation for Information Processing 2017*, Delhi, India, 122-134.
- Swim, J.K. and Bloodhart, B.: 2015, Portraying the Perils to Polar Bears: The Role of Empathic and Objective Perspective-taking Toward Animals in Climate Change Communication, *Environmental Communication-a Journal of Nature and Culture*, **9**(4), 446-468.
- Townsend, L. and Wallace, C.: 2017, The Ethics of Using Social Media Data in Research: A New Framework, in K. Woodfield (ed.), *The Ethics of Online Research*, Emerald Publishing Limited.
- Ylipulli, J., Suopajarvi, T., Ojala, T., Kostakos, V. and Kukka, H.: 2014, Municipal WiFi and interactive displays: Appropriation of new technologies in public urban spaces,

- Technological Forecasting and Social Change*, **89**, 145-160.
- Zhang, S.Q. and Feick, R.: 2016, Understanding Public Opinions from Geosocial Media, *Isprs International Journal of Geo-Information*, **5**(6), 74.