

Using the SPENCE Model of Online/Offline Community to Analyse Sociality of Social Machines

Caroline Halcrow
University of Southampton
University Rd.
Southampton SO17 1BJ
+44 023 8059 5000
cah4g12@soton.ac.uk

Leslie Carr
University of Southampton
University Rd.
Southampton SO17 1BJ
+44 023 8059 5000
lac@ecs.soton.ac.uk

Susan Halford
University of Southampton
University Rd.
Southampton SO17 1BJ
+44 023 8059 5000
susan.halford@soton.ac.uk

ABSTRACT

Online/offline community (O/OC), the *integrated performance of community in a blend of online/offline activities* is increasingly prevalent as online systems organise, mediate and broadcast forms of communal engagement. O/OCs are social machines where the focus is on the social achievement, rather than the computational outcomes, of the combined human-technical infrastructure. An O/OC model *SPENCE* is proposed as an analytical tool for describing social machines from the perspective of sociality.

Twitter is a technical infrastructure and social network of shared online/offline community phenomena that is also a social machine combining social participation with conventional forms of machine-based computation. Drawing from the extensive Twitter research literature, a sample of papers are analysed against *SPENCE*, demonstrating the clarity of the organisation of inter-relating themes of a range of perspectives in current Twitter research. It is concluded that *SPENCE* provides a lens of synthesis for the *sociality* dimension of a social machine and can be used in taxonomic activities (such as the social machines observatory) to differentiate social machines.

Keywords

SOCIAM, social machines, sociality, online/offline community, models, twitter, web observatory, literature classification.

1. INTRODUCTION

1.1 Online/offline community and social machines

Giddens [1] asks whether the use of online social networking has led to ‘the decline or reinvention of “community”’. O/OC is the *sequential or concurrent performance of community in a blend of online/offline*. It is proposed as a reinvention of community, a new ‘social representation’ [2] in which community is performed in a complex balance of online/offline. In an *online/offline community* (O/OC), people, in physical and/or virtual common places, ‘communicate via different modalities that include blending online and offline interaction’ [3].

O/OC is encompassed by the definition of social machine studied in the SOCIAM project¹: ‘loose collectives of people connected

by computational communication substrates at their core’²; or ‘...systems that combine some form of social participation with conventional forms of machine-based “computation”’ [4]. The O/OC view, i.e. the topographical, interpersonal online/offline blended communications perspective, is inherent in the social participation and interaction or *sociality* dimension of social machines such as Twitter. The definition of *Sociality* used here is the extent to which the system supports social interaction [4], communication, participation and exploration; and provides ‘landscape’ for ‘inhabiting’ to create ‘social fabric’ [5].

The model *SPENCE* frames the analysis of O/OC. It is a faceted, dynamic model, derived from an assembly of social phenomena allied to community theory, which is being iteratively evaluated in empirical observation including Focus Groups and study of London local area Twitter social data.

It is proposed that the *SPENCE* model offers a coherent, synthesised analysis of the *sociality* dimension of social machines that could be deployed as a method in the SOCIAM *Web Observatory*³ set of functions. It could provide granularity in SOCIAM’s taxonomic framework [4]. The approach to testing the feasibility of *SPENCE*, as an *Observatory* method for taxonomy use, involved its application in a literature review of the social machine, Twitter. Secondary was chosen over primary analysis of social machine, as an initial feasibility method. A single social machine was selected to provide a constant in the literature sample, affording cross-comparison.

The definition of O/OC evolves from Rheingold’s [6] study of an early online community - WELL (Whole Earth ‘Lectronic Link) blended with IRL (In Real Life). The concept of O/OC is built on the use of a multiplexity of channels [7] in community communications. Wellman and Gruzd [8] set out to define online community in Twitter and created the first *implicit* definition of O/OC. Their model is based on Anderson’s ‘Imagined communities’ [9], McMillan’s exploration of the concept ‘sense of community’ [10] and Jones’s concept of ‘virtual settlement’ [11].

Copyright is held by the International World Wide Web Conference Committee (IW3C2). IW3C2 reserves the right to provide a hyperlink to the author’s site if the Material is used in electronic media.

WWW 2016 Companion, April 11-15, 2016, Montréal, Québec, Canada.

ACM 978-1-4503-4144-8/16/04.

<http://dx.doi.org/10.1145/2872518.2890596>

¹ <http://sociam.org/>

² <http://sociam.org/socm2016/>

³ <http://sociam.org/content/updates-sociam-web-observatory>

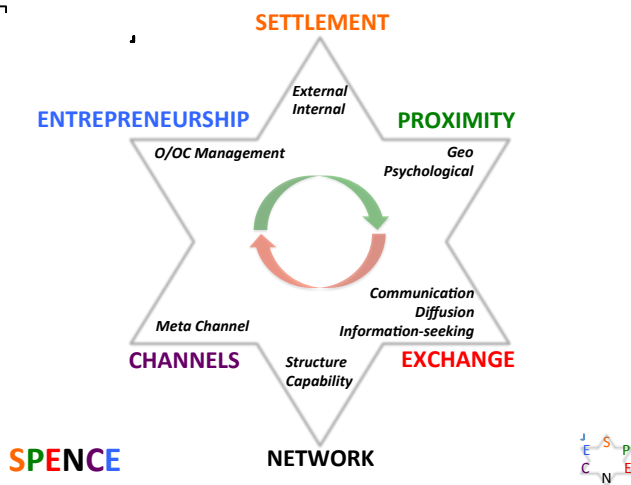


Figure 1. SPENCE – A model of online/offline community.

1.2 SPENCE

SPENCE developed from Gruzd and Wellman's starting point [8]. The theory was formulated through literature review of community and basic network theory that yielded an assembly of social phenomena that surfaced facet⁴ definitions. The assembly of social phenomena includes: the 'propinquity effect' or 'proximity principle' [12], [13]; the concept of social capital coined by Hanifan [14], and developed by Putnam [15], Bourdieu [16], Coleman [17]; the dynamics of trust [18], [19]; social cohesion [20]; and patterns of information diffusion [21]. Each of the facets of the model described below has core theory encapsulated in a general definition and key concepts.

SETTLEMENT is populated virtual, physical or cognitive space, with featured affordances, in which personal or collective community is performed.

: External; Internal

With **Settlement**, there is both an external and internal expression of space. A person engages with a bounded physical or virtual *External* settlement with affordances of offline public facilities and online social infrastructure. The engagement in *Internal* settlement is cognitively realised.

PROXIMITY is physical or psychological 'nearness' between people.

: Geo; Psychological. *Sub-concepts*: Interests, Needs, Values, Status.

With **Proximity**, people are physically or virtually 'near' each other in *geo*-physical or virtual neighbourhoods. *Psychological proximity* is also known as homophily and involves 'nearness' through sharing values, interests, needs and demographic status.

EXCHANGE is deliberate communication, diffusion and seeking of content, aiming for interaction, understanding, solution creation and link-making.

: Communication; Information-seeking; Diffusion – *Sub-concepts*: Direct, Indirect

Exchange is the meaningful, intended *Communication*, *Diffusion* and *Seeking* of content, which generates and substantiates the O/OC network in 'sociable and supportive social relations' [8].

NETWORK is the interpersonal, capital structure of trust, influence and information/data capability in confirmed ties.

: *Capability* – *Sub-concepts*: Influence, Trust,

Information/data; *Structure* – *Sub-concepts*: Node view, Cluster view

Network comprises in its *Structure* i.e. the topology of its entities and their relationships, analysed by node and cluster; and it encompasses the resources inscribed within it i.e. its *Capability*. *Capability* is made up of social and cultural capital, in 'sociable and supportive social relations' [8], whose outcomes are social benefit, mutual bonds and information/data solutions.

CHANNELS in multiplexity are the communications mechanisms across which the traffic of Exchange flows and the social phenomenon of Proximity operates.

: Meta channel (institutional O/OC)

Channels transform **Settlement** so that **Proximity** operates and **Exchange** is enabled. The *Meta channel* is an O/OC, e.g. Facebook, Twitter, that has achieved *Institutional* status due to global popularity.

ENTREPRENEURSHIP (social) is the design, creation, management and maintenance of an O/OC by an agent or organisation.

: O/OC Management; *Sub-concept*: Content management

Entrepreneurship necessarily facilitates social impact and change. The **Entrepreneur** acts as service manager, ensuring the **Settlement** is shaped for **Proximity**, **Channels** instrument the **Settlement**, and content in **Exchange** is stimulated, curated and moderated.

1.2.1 Synthesis of facets

The **Entrepreneurial** driver leads to the creation of **Settlement** using **Channels**, which enables **Proximity** that drives **Exchange** that creates **Network**. When **Proximity** is engendered through the *featured affordances* of the **Settlement**, **Exchange** is generated that develops into **Network**. **Entrepreneurship** maintains the lifecycle of the O/OC by managing the O/OC development to ensure new and adjusted elements in **Settlement** that drive **Proximity** etc.

1.3 Applying SPENCE in secondary analysis of Twitter as an O/OC and social machine

Twitter is the 'basis for a real community' Gruzd and Wellman [8] assert, centering their argument on experiments that refer to a model of online community determinants, that draws from and combines core community theory. Java [22] supports this view: 'Communities are the building blocks of any social network tools'. Quercia et al [23], apply 'sociological theories of real-life networks' to test whether social media communities resemble real-life communities. Murthy [24] applies sociological theory, i.e. Goffman's interactionist analysis, to Twitter data because he argues Twitter gives 'a certain level of richness' not offered by other mediated communication. This paper argues that Twitter involves the combination of 'social media communities' and 'real-life communities' [23] in the performance of online/offline community encompassed by the definition of a social machine. It attempts a similar exercise to Gruzd/Wellman, Quercia and Murthy in applying sociological mini-theory in the form of SPENCE, in secondary analysis of the O/OC within Twitter's

⁴ The facet labels, e.g. **Exchange**, **Network**, aim to encapsulate the core argument of the definition.

social machine. A sample of the significant⁵ body of twitter primary research literature was selected and *SPENCE* was used as a classification approach to secondarily analyse the Twitter research.

2. METHOD

A sample of twitter articles, with distinctly different foci, filtered by high citation counts, was selected using ACM Digital Library and Google Scholar. The articles were subject analysed and assigned a classification using the *SPENCE* model facets/concepts in Figure 1. The analysis gauged if the main research area(s) of the article matched wholly as a *focus* or in a *peripheral* or partial way with the theory of a facet. Additionally it was noted by facet (using the marker *goal*) if the research was concerned with the ‘goal, task and process’ of Twitter social operations. A number of *foci* or *peripheral* markers could be attributed to each article.

The subject analysis method is demonstrated using the paper ‘The Social World of Twitter: Topics, Geography, and Emotions’ [23]: 1) the main foci of the article were identified; 2) brief discursive notes were made on the assumptions and the implications of the foci; 3) the foci were aligned with the core definitions of the *SPENCE* facets. The foci-alignments were identified as: network metrics, aligned with **Network**; the effectiveness of topic diversity, aligned with **Proximity**; the significance of geographic locality for strong ties, aligned with **Settlement** and **Network**; and the silent communities of shared emotions, aligned with **Proximity**. Quercia’s paper contains 3 facet foci, according to the *SPENCE* model.

3. RESULTS

The Facet with the highest relevance is **Exchange**: 13/15 of the articles had this focus; the next most relevant facet was **Network** with 9/15. **Proximity**, **Channels** and **Settlement**, ranked in order, scored significantly lower than **Exchange** and **Network**. The most ‘goal, task and process’ concerned research was in the area of **Exchange** (5) followed by **Settlement** (3). The number of classifications made for the 15 articles was 52. They were 9/52 *goal* classifications. The most ‘comprehensive’ articles with the most foci are: Gruzd [8] with 4 foci; Quercia [23], Himmelboim [25], Huberman [26] and Takhteyev [27] with 3 foci; followed by Kwak [28] with 2 foci and 2 peripheral markers.

Table 1. Classification results

FACETS	Focus	Peripheral	Goal	Totals
Settlement	3	2	3	8
Proximity	6	2		8
Exchange	13		5	18
Network	9	1	1	11
Channels	4	2		6
Entrepreneurship		1		1
Totals	35	8	9	52

⁵A search of the ACM Digital Library using the keyword ‘Twitter’ yielded a results set of 15,549 in October 2015.

4. DISCUSSION

The classification exercise revealed a heterogeneous mix of research in terms of focus. The two facets, encompassing established sociological theory - **Exchange** and **Network** - appeared the most relevant in this selection of Twitter-based research. The less established strands of theory in the facet of **Proximity**, scored significantly with 5/15 articles aligned. The goal-based research proportion is roughly 1/5. There is one article - Gruzd [8] - that has foci that align with 4/5 of the facets, evidencing a rounded analysis of Twitter. A possible application of *SPENCE* would be to assess the breadth of a research perspective by number of facets as focus or peripheral. The discussion of the results is organised using the *SPENCE* facets and the classification categories in Table 1. The authors are listed by category in 7. Appendix. The discussion implicitly tests if *SPENCE* can effectively organise themes in a selection of Twitter research literature.

4.1 Settlement

The facet of **Settlement** is the focus in 3 articles. Gruzd [8] gives a detailed analysis of ‘virtual settlement’ (after Jones, [11]) from which the facet conceptualisation in part derives. Takhteyev [27] observes how **Network** and **Settlement** are blended: ‘...a substantial share of ties lies within the same metropolitan region’ and Quercia [23] connects **Settlement** with **Proximity** in proposing that ‘...geographically-constrained twitter users share information about diverse topics...’. Kwak [28], with a peripheral interest in **Settlement**, discovers that users with less than 2000 reciprocal followers in a **Network** structure are likely to be geographically close. Quercia [23] contributes novel ideas about the design of O/OC **Settlement** e.g. ‘self-reflecting user interfaces that make people aware of their emotions’.

4.2 Proximity

There are 6 articles classified with **Proximity** as a focus. Gruzd [8] in observing that ‘The notion of “community” has often been caught between concrete social relationships and imagined sets of people perceived to be similar’ gives clarity to how the facets of **Exchange**, **Network** and **Proximity** interplay. Kwak [28] observes that reciprocated relationships (i.e. **Network**) exhibit ‘some level of homophily’, i.e. **Proximity**. Murthy [24] has a focus on the follower/followed relationship in Twitter. This interest is classified as **Proximity** as the follower/followed relationship is not always mutualised in concrete social relationships of **Exchange** and **Network**. Gruzd [8] has a 4 foci facet interest while Murthy [24] is less concerned with **Exchange** and **Channels** – they are of peripheral interest. The examination by Gruzd [8] appears rounded, if not cohesive, in its multi-faceted understanding of the social phenomenon of Twitter. Himmelboim [25] discovers in his experiments that ‘Political content...was overall confined to like-minded clusters of users’. The ‘like-minded cluster’ is an effective description of **Proximity**. Java [22] also finds the community structure in microblogging is based on the **Proximity** of shared interests. Quercia [23] considers ‘communities’ distinguished by shared positive or negative emotions, operating in sparser networks. The observed ‘silent’ communities of emotion align with the **Proximity** facet. Huberman’s [26] argument demonstrates **Proximity** inversely: his interest is in the mutual interactions ‘that matter’ through significant actualisation within ‘linked structures’ of **Network** and **Proximity**. But this assumes the ‘silent’ linked structure of the follower/followed. The insight that Huberman’s [26] view invites is that the ‘personal networks’ of e.g. Facebook, depend on ‘social networks that matter’ evidenced in actual **Exchange**. Other O/OC

platforms such as Twitter in which both personal and collective O/OC is performed, depend on a balance of **Exchange** and **Proximity**.

4.3 Exchange

The number of articles with **Exchange** as their focus is 13. This 13/15 proportion indicates that the study of *Communication/Diffusion*, is core to Twitter research and is at least one of the major foci in articles. This accords with the social communications-based nature of the Twitter platform. Gruzd [8], Murthy [24], Himelboim [25], Quercia [23] and Huberman [26], discussed above, also have *Exchange* as foci. Asur [29], Jansen [30] and Bollen [31] are 3 of the 5/13 researchers that have a goal-based perspective. They share an interest in sentiment, mood or opinion analysis. **Exchange** in Twitter appears to particularly serve ‘goal, task and process’. Asur [29], Jansen [30] and Bollen [31] consider using sentiment/mood analysis for general forecasting, marketing and stock market prediction, respectively. Pak [32] investigates a novel method of sentiment analysis but does not consider industrial applications. In the *SPENCE* model, the conceptual account of tweet content divides into *Communication* i.e. reciprocated meaning (e.g. a hashtag with a twitter account mention, @) and *Diffusion* i.e. information (e.g. URLs), although they can be combined in a tweet. The concept of sentiment is assumed within *Communication*: the use of sentiment implies a reciprocated connection with a strong tie or a meaningful utterance aiming for understanding. Tsur’s [33] article, with the 2 foci **Exchange** and **Network**, centres on method i.e. predicting how the interestingness and sentiment of the meme/idea/hashtag within its tweet context impacts on its *Diffusion* in the **Network** in a given time. Java’s early pre-hashtag analysis of Twitter, supports the *SPENCE* division of **Exchange** into *Communication* (‘conversation’ using @) and *Diffusion* (URLs as ‘information’). Kwak et al [28] examine the entire ‘Twittersphere’, considering *Diffusion* in the *Network*, and discover that diffusion through retweets does not operate by power law: any twitter user, regardless of the number of their followers, who is retweeted is likely to have a readership of 1000.

4.4 Network

9/15 articles are concerned with the **Network** facet of Twitter. This facet is the second most relevant after **Exchange**. 8/9 of the articles have both foci, showing the interdependency of these facets. Kwak’s [28] interest is in Twitter’s ‘...power as a new medium of information sharing’. This combines the *Diffusion* of **Exchange** with the topology or *Structure* of **Network**. The study is comprehensive as in addition to these 2 foci it also has peripheral markings of **Proximity** and **Settlement**. Gruzd [8] in his ‘rounded’ article of 4 foci has a fundamental concern with network mutuals. Tsur [33], Himelboim [25], Java [22] and Huberman [26] also address the correlation between **Exchange** and **Network Structure**. Content correlated with network analysis is clearly a core axis of interest in this sample of twitter research. Quercia [34] like Kwak [24] is interested in ‘influentials’. The facet **Network** is divided into the concepts of **Structure** and **Capability**, which are intended to encapsulate the dual topology and social capital nature of **Network**. With the latter, there are the sub-concepts of *Influence* and *Trust*. But it is not the *collective* influence capability of a network that is of core interest in Kwak [28] or Quercia [34] rather the *individual* influence capability score. The study of the collective influence capability of a social network appears to be a gap in Twitter research literature, based on this selection, although Kwak’s *implicit* research goal

addresses the potential macro influence capability of the Twittersphere.

4.5 Channels

Takhteyev [27] examines ‘the mechanisms through which distance and ties relate’ i.e. the mechanism of air travel. A **Channel** is a mode of conveyance for communications, most obviously e.g. internet/web or face-to-face. Takhteyev [27] argues that airlight in facilitating long-distance face-to-face interaction - as a Channel - is ‘the best predictor of non-local Twitter ties’. This argument that transport infrastructure facilitating face-to-face communications significantly relates to ties in an O/OC marks the importance of **Channels** to **Network** in O/OCs. Zhao [35] studies Twitter as a news medium, comparing it to the ‘traditional news medium’ of the ‘New York Times’. In *SPENCE*, a ‘traditional news medium’ has institutional status as a *Meta channel* and as such equates to a Channel. Following this argument, Twitter, an institution in its huge popularity, is a *Meta channel*: it is a Channel-platform for O/OC. URLs in tweets to news articles in the ‘New York Times’ are a Channel within the Channel of Twitter. So *Meta channels* in O/OCs, include the Twitter-platform, airlight, and the ‘New York Times’.

4.6 Entrepreneurship

Of the 15 articles selected in this classification experiment, there is no mention of the entrepreneurial ownership or management of Twitter as a Channel-platform for O/OC. This facet is only of peripheral interest in Quercia’s piece [23] when he indirectly suggests that user interface design could be adjusted - i.e. *by the O/OC owner* - to improve emotional intelligence in users.

5. CONCLUSION

The classification reveals a heterogeneous mix of research that is in key cases focused on the utility of Twitter, or in SOCIAM taxonomy terms, its ‘goal, task and process’ [4]. Twitter serves functionally as a news delivery service, a viral marketing instrument, a stock market mood detector, a celebrity or influential broker/hubs ranker, or an information diffusion pattern alerting tool. Crucially, however, it also demonstrates the *performance* of online/offline community: the *sociality* dimension is fundamental.

The *SPENCE* classification effort evidences clarity in the secondary analysis of themes in twitter research. It shows the strength of the *SPENCE* model as a method. The Model offers a lens for synthesis that reveals a coherence of social phenomena. This is exemplified in the approach to the Quercia [23] article. Quercia, though demonstrating a roundedness of multi-faceted focus, does not cohere his 4 main contributions: network metrics; the effectiveness of topic diversity; the significance of geographic locality for strong ties; and the silent communities of shared emotions. Using *SPENCE* in secondary analysis, these themes can be viewed as the interdependence of **Settlement**, **Proximity** and **Network**. The review did not intend to critique research direction or identify gaps but it is suggested that it could be used in the future for these functions and for research motivation.

The lens of synthesis offered by *SPENCE* could add a particular quality to the analysis of *sociality*, driven by basic network logic implicit in the facet relationships. *SPENCE* could support the SOCIAM goal to ‘...describe and differentiate current social machines when viewed as a collective’⁶ by faceting and describing the degrees and interdependent elements of *sociality* in social machines.

⁶ <http://sociam.org/socm2016/>

It is suggested that Giddens's freighted observation of the commonplace '...people will fit the new social media into their everyday routines and use them alongside their face-to-face relationships' [1] touches on an emergence of a new social reality in which social machines are fundamental societal infrastructure. *SPENCE* could offer a view on the implications of social machines assuming the status of 'social representation' [2]. It is proposed that *SPENCE* in contributing to the social representation view and to other taxonomical activities would work as an effective method in the SOCIAM Web Observatory⁷ set of functions.

6. REFERENCES

- [1] Giddens, A. and Sutton, P. 2013. *Sociology*, 7th ed. Polity, Cambridge.
- [2] Moscovici, S. 1981. On Social representations. *Social Cognition: Perspectives on Everyday Understanding*, 181-209.
- [3] Preece, J. and Maloney-Krichmar, D. 2005. Online Communities: Design, Theory, and Practice. *Journal of Computer-Mediated Communication* 10, 4.
- [4] Smart, P., Simperl, E. and Shadbolt, N. 2014. A Taxonomic Framework for Social Machines. In *Social Collective Intelligence*, D Miorandi, Ed. Springer, Cham, Switzerland, 51-85.
- [5] Murray-Rust, D., Tarte, S., Hartswood, M. and Green, O. 2015. *On Wayfaring in Social Machines*. In *Proceedings of the 24th International Conference on World Wide Web* (Florence, Italy, 18-22, 2015).
- [6] Rheingold, H. 1993. *The Virtual Community: Homesteading on the Electronic Frontier*. MIT Press, Cambridge, Mass.
- [7] Haythornthwaite, C. 2001. Exploring Multiplexity: Social Network Structures in a Computer-Supported Distance Learning Class. *The Information Society* 17, 3, 211-226.
- [8] Gruzd, A., Wellman, B. and Takhteyev, Y. 2011. Imagining 'Twitter' as an Imagined Community. *American Behavioral Scientist* 55, 10.
- [9] Anderson, B. 1983. *Imagined Communities: Reflections on the Origin and Spread of Nationalism*. Verso, London.
- [10] McMillan, D. W. 1986. Sense of Community: A Definition and Theory. *Journal of Community Psychology* 14, 1, 6-23.
- [11] Jones, Q. 2006. Virtual-Communities, Virtual Settlements and Cyber-Archaeology: A Theoretical Outline. *Journal of Computer-Mediated Communication* 3, 3, 0-0.
- [12] Festinger, L., Back, K. W. and Schachter, S. 1950. *Social Pressures in Informal Groups: A Study of Human Factors in Housing*. Stanford University Press, Stanford, California.
- [13] Newcomb, T. M. 1956. The Prediction of Interpersonal Attraction. *American psychologist* 11, 11, 575.
- [14] Hanifan, L. 1916. Rural School Community Center. *Annals of the American Academy of Political and Social ...*67, 130-138.
- [15] Putnam, R. D. 2000. *Bowling Alone: The Collapse and Revival of American Community*. Simon and Schuster, London.
- [16] Bourdieu, P. 2008. The Forms of Capital. *Readings in Economic Sociology* 4, 280.
- [17] Coleman, J. 1989. *Social Capital in the Creation of Human Capital*. University of Chicago Press, Chicago.
- [18] Sherchan, W., Nepal, S. and Paris, C. 2013. A Survey of Trust in Social Networks. *ACM Computing Surveys (CSUR)*, 45, 4, 47.
- [19] Fukuyama, F. 1995. *Trust: The Social Virtues and the Creation of Prosperity*. Free Press, New York.
- [20] OECD. 2012. Perspectives on Global Development 2012 *Social Cohesion in a Shifting World*. OECD Publishing, Paris.
- [21] Weng, L., Ratkiewicz, J., Perra, N., Gonçalves, B., Castillo, C., Bonchi, F. and Flammini, A. 2013. The Role of Information Diffusion in the Evolution of Social Networks. In *Proceedings of the 19th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*.
- [22] Java, A., Song, X., Finin, T. and Tseng, B. 2007. Why We Twitter: Understanding Microblogging. In *Proceedings of the 9th WebKDD and 1st SNA-KDD 2007 Workshop on Web Mining and social Network Analysis*, 56-65.
- [23] Quercia, D., Capra, L. and Crowcroft, J. 2012. The Social World of Twitter: Topics, Geography, and Emotions. In *Proceedings Sixth International AAAI Conference on Weblogs and Social Media*, 298-305.
- [24] Murthy, D. 2012. Towards a Sociological Understanding of Social Media: Theorizing Twitter. *Sociology*, 46, 6, 1059-1073.
- [25] Himelboim, I., McCreery, S. and Smith, M. 2013. Birds of a Feather Tweet Together: Integrating Network and Content Analyses to Examine Cross-Ideology Exposure on Twitter. *Journal of Computer-Mediated Communication* 18, 2, 40-60.
- [26] Huberman, B. A., Romero, D. M. and Wu, F. 2008. Social Networks That Matter: Twitter Under the Microscope. In *Proceedings of the International Conference on Computer-Supported Cooperative Work*, 107-114.
- [27] Takhteyev, Y., Gruzd, A., and Wellman, B. 2012. Geography of Twitter Networks. *Social Networks* 34, 1, 73-81.
- [28] Kwak, H., Lee, C., Park, H. and Moon, S. 2010. What is Twitter, a Social Network or a News Media? *The International World Wide Web Conference Committee (IW3C2)*, 1-10.
- [29] Asur, S. and Huberman, B. 2010. Predicting the Future with Social Media. *Computing* 1, 492-499.
- [30] Jansen, B. J., Zhang, M., Sobel, K., and Chowdury, A. 2009. Twitter Power: Tweets as Electronic Word of Mouth. *Journal of the American society for Information Science and Technology* 60, 11, 2169-2188.
- [31] Bollen, J., Mao, H. and Zeng, X. 2011. Twitter Mood Predicts the Stock Market. *Journal of Computational Science* 2, 1, 1-8.
- [32] Pak, A. and Paroubek, P. 2010. Twitter as a Corpus for Sentiment Analysis and Opinion Mining. *Lrec*, 1320-1326.
- [33] Tsur, O. and Rappoport, A. 2012. What's in a Hashtag?: Content Based Prediction of the Spread of Ideas in Microblogging Communities. In *Proceedings of the Fifth ACM International Conference on Web Search and Data Mining*, 643-652.
- [34] Quercia, D., Kosinski, M., Stillwell, D. and Crowcroft, J. 2011. Our Twitter Profiles, Our Selves: Predicting Personality with Twitter. In *Proceedings - 2011 IEEE International Conference on Privacy, Security, Risk and Trust and IEEE International Conference on Social Computing, PASSAT/SocialCom 2011*, 180-185.
- [35] Zhao, W. X., Jiang, J., Weng, J., He, J., Lim, E.-P., Yan, H. and Li, X. 2011. Comparing Twitter and Traditional Media Using Topic Models. In *Advances in Information Retrieval*, P. Clough, Ed. Springer-Verlag, Berlin, 338-349.

7. APPENDIX

FACETS	Focus	Peripheral	Goal
Settlement	Gruzd [8] Quercia [23] Takhteyev [27]	Java [22] Kwak [28]	Quercia [23] Quercia [34] Java [22]
Proximity	Gruzd [8] Murthy [24] Himmelboim [25] Quercia [23] Huberman [26] Java [22]	Takhteyev [27] Kwak [28]	
Exchange	Asur [29] Gruzd [8] Himmelboim [25] Jansen [30] Murthy [24] Pak [32] Quercia [23] Quercia [34] Tsur [33] Kwak [28] Zhao [35]		Asur [29] Himmelboim [25] Jansen [30] Java [22] Bollen [31]

	Java [22] Huberman [26] Bollen [30]		
FACETS	Focus	Peripheral	Goal
Network	Gruzd [8] Himmelboim [25] Quercia [23] Quercia [34] Takhteyev [27] Tsur [33] Kwak [28] Java [22] Huberman [26]	Murthy [24]	Himmelboim [25]
Channels	Gruzd [8] Quercia [34] Zhao [35] Takhteyev [27]	Himmelboim [25] Murthy [24]	
Entrepreneurship		Quercia [23]	