

Social Media Domain Analysis (SoMeDoA) A Pharmaceutical Study

David Bell and Sara Robaty Shirzad

Brunel University, Information Systems and Computing, London, U.K.

Keywords: Social Media Network (SMN), e-Marketplace, Business to Business Environment.

Abstract: Social media data is increasingly becoming a valuable asset for marketing teams, and businesses are regularly coming up with new and innovative ways to make use of this data. A social media network (SMN) is able to connect enterprises with their customers, partners and even competitors. Public trading and relations-oriented structures of social media networks (SMN) have encouraged organizations to engage more actively with other transactional partners. Organizations are seeking to tap into the relationship development potential these websites offer, especially the network effect of each individuals or organisations social graph. It is recognized that these relationships (when utilised) are able to create value for network participants. This paper discusses SMN tools and outlines a methodology and procedure that supports the identification of domain specific networks within a global business-to-business environment. Research is carried out using SMN data about firms in the pharmaceutical industry. We use our own methodology to uncover market participants, linkages and prominent issues that may help new firms to position themselves effectively in a new marketplace. SMNs provide a considerable source of information and new methods are required to fully leverage their potential value. This paper explores how SMNs can be used as an effective source of business intelligence by analysing a popular SMN platform.

1 INTRODUCTION

As markets increase globally, smaller firms, whether they want to be or not, find themselves to be part of social networks (Pitt *et al.*, 2006). Consequently, increased connectivity to customers results in increased competition with rivals from around the world. Being in the social networks could even be an opportunity for them to survive and compete with larger counterparts (Copp and Ivy, 2002; Masarel and Janszen, 1998; Lipparini and Sobrero, 1994). In September 2012, CNBC, a recognized leading website in business news, reports that about 15.2 million site members in LinkedIn (one of the social network websites for professionals) or about 8.7 per cent of the site's 175 million members worldwide are small-business professionals. The strategic choice of new members of social networks is simple: how to understand these social networks in order to get them work for their own business. There is lack of empirical studies on social networks that try to answer questions around knowledge of discrete business networks and the advent of internet provides a unique opportunity to study these

business interrelationships. The Internet itself could be considered as the largest social network. Moreover, the internet is becoming the most popular vehicle for business to business (B2B) commerce. In this article, we examine the social network facing small to medium firms (SMEs) in the pharmaceutical industries. To achieve this aim, we analyse activity of a number of large organizations in the same domain to understand how large companies are cultivating relationships in SMN's so that small-medium organizations could explore their relationships to customers and counterparts.

We use the framework to recognize prominence that helps the new firms to position themselves effectively in a social network to leverage considerable value. This paper aims to improve the understanding on how SMNs can be used as a reliable source by analysing the temporal on *Twitter*¹ activities. Twitter was chosen for the popularity, according to a recent study (Skeels and Grudin, 2009) one third of employees in enterprises are in Twitter company networks. Professionally oriented

¹ <http://www.twitter.com>

structures within Twitter make it popular among organizations. Although the methodology outlined from the study has only involved one domain (albeit a large one), it is argued that the richness of the information provided by users from different backgrounds will provide generalizable outcomes to a range of scenarios. The paper is structured as follows: First, we provide a brief overview of social media commerce. Next we describe a methodology for exploring the data. Third, we describe the results of the studies. After identifying the obvious limitations of the research, we conclude and discuss future opportunities for new entrants to the e-Marketplace.

2 SOCIAL MEDIA COMMERCE: A PLATFORM FOR FINDING CONNECTIONS

With the large amount of information potentially available to organizations, the internet constitutes an important platform for information exchange between the consumers and industry suppliers, intermediaries, as well as organizations which are not have experience of being in the e-marketplaces. Different technology interfaces such as search engines, intermediaries facilitate the marketing information exchange between online organizations. Social media, which enables interaction among virtual organizations, has emerged as an integral element of communication. Ellison (2007) defines social media sites as Web-based services that allow individuals to (1) construct a public or semi-public profile within a bounded system, (2) articulate a list of other users with whom they share a connection, and (3) view and traverse their list of connections and those made by others within the system. Noted by Harris (2009), there are hundreds of different social media platforms such as text messaging, shared photo, social network, wikis, and discussion group. According to Alexa (2012), a Web information company that provides website traffic rankings, top 5 global social media websites by late 2012 that have significant presence on enterprises are: (1) Facebook, (2) Twitter, (3) LinkedIn, (4) MySpace and (5) Google Plus+. Some of these websites are more likely accessed by youths with IM experience such as Facebook. Some of them like LinkedIn and Twitter target professional use from the start.

In general, social media websites are beneficial and valuable for the network participants in that they

promote activities and the use of resources (Michaelidou, Siamagka and Christodoulides, 2011; Walter *et al.*, 1997). A study published on February 2010 by the Small Business Success Index (SBSI) indicates that 75% of the surveyed small businesses in the USA have already created a company page on a social networking site. A number of researchers are focused on social media theory to study the social networks of firms in B2B marketplaces (e.g. (Michaelidou, Siamagka and Christodoulides, 2011; Björkman and Kock, 1995)). Further, number of studies on small firms argues that social networks are important for the survival of small firms and critical in competing with large businesses (Pitt *et al.*, 2006; Copp and Ivy, 2002). Indeed social media websites are suited for collecting information/feedback from customers, developing relationships between customers through interaction. Some of the interactions from organizations and supporters in social media websites summarized by Carrera *et al.* (2008) are: posting links to external news item about the organization or its causes; using messaging board or discussion wall to post-announcements and answer questions.

Adopting new channels of technology such as social media websites may not be attractive to many companies. Many researches stress that many organizations have been slow to adopt new technologies due to perceived barriers such as lack of money, time and training, negative views about usefulness, as well as unfamiliarity with technology (Buehrer, Senecal and Bolman Pullins, 2005; Venkatesh and Davis, 2000). Additionally, Frambach and Schillewaert (2002) argue that organization size is also important in the adoption decision. They further suggest that smaller organizations are more innovative and are therefore expected to be more receptive to new changes in technologies. Copp (2002) and Pitt (2006) highlight that social network websites are important for the survival of small firms and critical in competing against large businesses. Increasing the number of socially active organizations change the strategic view of other partners to consider social networks as important arena to consider for competing in virtual marketing world. Figure 1, illustrates a conceptual framework of social media commerce that an organization seeks to participate in B2B marketplace and getting information about that marketplace through social media networks. Social media networks in this framework are a knowledge source for newly entering organizations.

One of the advantages of being active in social media websites is notable in their design in relation

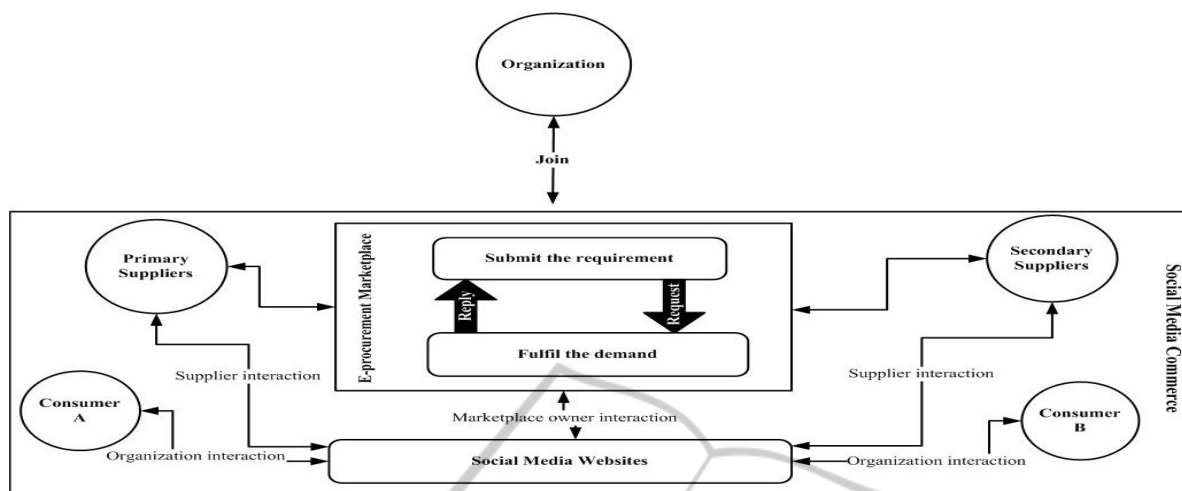


Figure 1: Social Media Commerce.

to both time and space (De Longueville, Smith and Luraschi, 2009). For instance, tweets on Twitter are organised in timelines (i.e. series of tweets sorted and displayed in reverse chronological order) (De Longueville, Smith and Luraschi, 2009). On the other hand, Social media websites are structured architecturally so as to communicate a space (e.g. geo-location in twitter) that is publicly accessible. Thus, private and public boundaries are employed to situate the network geographically (Papacharissi, 2009). Longueville et al. (2011) explored the role of Twitter as source of spatio-temporal information for retrieving, validating and filtering spatio-temporally referenced images from Flickr within the context of flooding in the UK to advance existing capabilities for monitoring natural hazards. They emphasized that graphical representation of place names as spatial references (e.g., town, county, etc.), resulted in space accuracy from a gazetteer viewpoint. In contrast, temporal referencing is the time-stamp when something posted on social media's. Lee et al. (2011) (Lee *et al.*, 2011) plotted information from twitter on earthquakes on a world map. They used city names to define positions of tweets.

Although, the body of scientific literature about social media websites are growing (as an example, the online scientific literature database Scopus.com provides 88 results for 2009, 123 results for 2010, 194 results for 2011 and 259 results for 2012 while searching for the keyword social media websites), its potential for spatio-temporal information has still to be fully assessed. Most of the research is focused on social aspects by analysing users satisfactions (Java *et al.*, 2009), interactions of users with social websites (Huberman, Romero and Wu, 2008) or adoption (Hughes and Palen, 2009). Pharmaceutical

e-marketplace adoption has been volatile over the past decade (Shirzad, 2013). This paper aims to provide further exploration of the role of social media websites such to identify the motivating factors for pharmaceutical organizations to participate in e-marketplaces and to consider the valuable temporal and geospatial components they can contain. Our research question requires us to advance spatio-temporal analysis methods in order that suitable knowledge can be extracted from social media networks for newly entering organizations to the market. In the remainder of this paper, possible methods are described, using a real-life datasets as an illustration. This work can serve as a theoretical base for future works aiming at performing web-based event, behaviour and strategy detection.

3 SOCIAL MEDIA DOMAIN ANALYSIS - SoMeDoA APPROACH

The temporal model considers the visit activity of people to specific times (including intervals) (Yoo and Hwang, 2008). The idea of harvesting temporal information from the Web has been of interest in recent years (De Longueville, Smith and Luraschi, 2009). Li et al. (2005) proposed a probabilistic model to detect retrospective news events. They explained the generation of "four Ws" - who (persons), when (time), where (locations) and what (keywords), from each news article. However, they considered time and location for discovering the reoccurring peaks in events. Mei et al. (2006) produced a model for spatio-temporal analysis for

Table 1: SoMeDoA Research Framework.

Phase	Description	Resulting output
Data Selection	Social media Web sites are selected as suitable sources for the domain of study.	List of social media platforms and associated search terms.
Data Gathering	Data gathering tools are selected and run against the selected Social Media sites.	List of software tools. Generated data files.
Temporal Separation	Public information, news and communications are extracted in order to determine the public activities of organisations (with associated timelines)	DateTime lists files for each organisation.
Temporal Coding	Further analysis of temporal data in order to uncover topics of importance (with timeline)	Keyword lists and domain ontology DateTime Data lists for each keyword, code or category

weblog data. In contrast to previous work, we apply the temporal model to describe organizations activities on Twitter with a more explorative motivation.

Our high level research framework is presented Table 1 and titled Social Media Domain Analysis (SoMeDoA). It involves core elements of social media data gathering and data analysis (including Grounded Theory approaches). Data from specific social media web sites is extracted using domain specific search terms that each target particular organisations temporal data sets. The generated data files are then analysed using a mix of visualisation and analytical tools.

Our research approach “in action” consisting of two main stages: *temporal separation* and *temporal coding*, converting data from Twitter into temporal aspects of organizations information. Twitter is selected in order to effectively detect real-time activity of organizations within a domain (the “what” and “when”).

We use the name of the organizations as a query term to get the tweets, they (or others) publish. Subsequently, Tweetcatcher2 (an application developed on the MATCH project at Brunel University) is used to retrieve tweets and related data such as published date, user ID, tweet, number of following, number of followers, time zone, number of users tweets, retweet count, expanded links and sentimental analysis (Sentiment analysis assigns scores indicating positive, neutral and

negative opinion to each distinct entity in the text (Pak and Paroubek, 2010). The *temporal separation and coding analysis* activities are developed for handling Twitter message streams, and categorized them with respect to the number of tweets published and frequency of occurrences in the selected timeslots. Temporal separation analysis was carried out in Microsoft Excel 2010. The dataset is visualized based on the time and number of tweets generated. The second part of Twitter analysis was temporally coding. The approach taken for the temporal coding analysis of content was based on the grounded theory method (GTM). GTM is the process of generating a theory from collected data (Glaser and Strauss, 1967). We used Nvivo9 (as software to support grounded theory techniques) in order to analyse tweet data. It was used for content analysis: 1) Storage and categorizing datasets, 2) conducting searches for further analysis in order to generate reports about frequency of word occurrences and associated categorisations and 3) creation of categories through computer-assisted coding. For example, a financial innovation category was created with associations to acquisition, finance and investor. Tweet frequency was used as a guide to category and sub-category importance. Both temporal separation and coding continue on with sentimental analysis with respect to time and wording. The data used in this study were collected in November 2012. A total of 54365 tweets were captured posted by selected organizations. Social

media data (including data interfaces) offer structure to data not found with traditional Web mining. Field descriptors in the Web sites' data interface or annotation (e.g. hashtags) in the data both offer opportunities for improved analyses.

4 EXPERIMENTS AND RESULTS

The pharmaceutical domain has actively used e-marketplaces over the past decade with platform volatility in line with evolutions on the Web (Shirzad, 2013). Further exploration is required into the role of social media websites can play in identifying the motivating factors for pharmaceutical organizations to participate in e-marketplaces. The first step is to fully understand the business network in which they operate and to consider the value of temporal components that are accessible.

4.1 Twitter Temporal Separation

Temporal analysis deals with time components (Lauw *et al.*, 2005). So far we have been monitor tweets posted daily from 11th to 29th of November 2012. As mentioned before, the top five pharmaceutical organizations were selected from Fortune Global which is an official website for ranking pharmaceutical organizations. 54,365 tweets were posted and subsequently downloaded for analysis. In order to calculate the proportion of organization activities on Twitter, we divide the dataset into three weekly time slots. Later on we decide to analyse Twitter activity on each Wednesday of weeks in November.

4.1.1 Tweets per Week

The first step was to analyse the overall number of tweets and how the numbers of tweets published are varying week by week. Figure 2 shows a graph with the total number of tweets between 14.11.2012 and 28.11.2012. The first and last tweet mentioning in all three time slots were published at 00:00 and 23:59 respectively. The columns are positioned over a label that represents the date and time that tweets posted. The height of the column indicates the number of tweets that the chosen organizations (under analysis) posted, defined by the column label. The highest rise in number of tweets was on 21th of November (Figure 2-B) where the total number of 38,272,550 tweets was published. Whereas the number of tweets on 14th (Figure 2-A) and 28th

(Figure 2-C) were 33,992,110 and 10,001,050 respectively. The other interesting issue in figure 2 is that the number of Tweets per minute on figure 2-C is more than the other time slots. For example on 28th of November, the peak tweets content were about online buying medicine from Roche (one of the chosen organizations). Most of the tweets were about new way of doing the online buying the medicines through the new portal. Also the highest number of re-tweet is on the same organization. The motivation behind this is investigating how useful is the sentiment lexicon developed for tweets in these three time slots. Peaks can be investigated for additional knowledge. In one example, a peak includes tweets about Warren Buffet's Berkshire Hathaway, whose sale of Johnson & Johnson shares are reported in the mainstream media.

4.1.2 Sentimental Average per Week

To assign numerical scores to sentiments of an individual sentence, we use the SentiStrength 7 tool developed by Thelwall *et al.* (2010) (Thelwall *et al.*, 2010) and used by Brunel's Tweetcatcher. This tool simultaneously assigns both a positive and a negative score to bits of English text, the idea being that users can express both types of sentiments at the same time, such as in "I love you, but I also hate you" (Kucuktunc *et al.*, 2012). Positive sentiment strength scores range from +1 (not positive) to +5 (extremely positive) and similarly, negative sentiment strength scores range from -1 to -5 (Kucuktunc *et al.*, 2012). The final positive sentiment strength for a bit of text is then computed by taking the maximum score among all individual positive scores. The negative sentiment strength is similarly calculated. Figure 3 shows the distribution of sentiment.

The vast majority of sentences are assigned a neutral +1/-1 sentiment score. Slightly negative (+1/-2) and slightly positive (+2/-1) scores are also common. Table 2 shows the percentage of each score in each time slot.

Table 2 clearly shows that the percentages of positive and negative tweet sentiments are much higher in first and last weeks of the month. The weekly sentiment analysis doesn't indicate the actual content of the tweets (e.g. Berkshire Hathaway's share sale mentioned earlier) beyond weekly sentiment analysis. Therefore temporal coding was conducted. This leads us to investigate more about tweets content by counting a word frequency for all the dataset and undertake sentiment analysis on frequently occurring words.

4.2 Twitter Temporal Coding

4.2.1 Tweet per Words

In the first instance, frequently used words or topics need to be identified in order to get a picture of the actual tweet content. Data needs to be subject to careful scrutiny and interpretation, which is largely achieved through a coding process. The approach taken for analysing content makes use of the grounded theory method (GTM). GTM is the process of generating a theory from collected data (Glaser and Strauss, 1967). The process was conducted by counting the word frequency for the dataset of tweets using Nvivo9. The most frequent words were “http” followed by other parts of URLs that appeared in most of the tweets (which should be discounted). After excluding articles and other terms

that did not provide meaningful context, table 3 shows the most frequent words.

As table 3 presents, the most frequent words are Roche followed by Johnson that appears in the most of the tweets. At this level we can get a general impression of key players and typical work associations (e.g. news, sales). Sentiment analysis on the most frequent words will help us to understand more on positive or negative tweets about those words over time.

4.2.2 Sentimental Average per Word

The pharmaceutical organizations are unsurprisingly most occurring words (as they are the search terms in question). Therefore, we first decide to do sentiment analysis on tweets to see how many positive and negative tweets published on each organization name. Table 4 presents the overall view

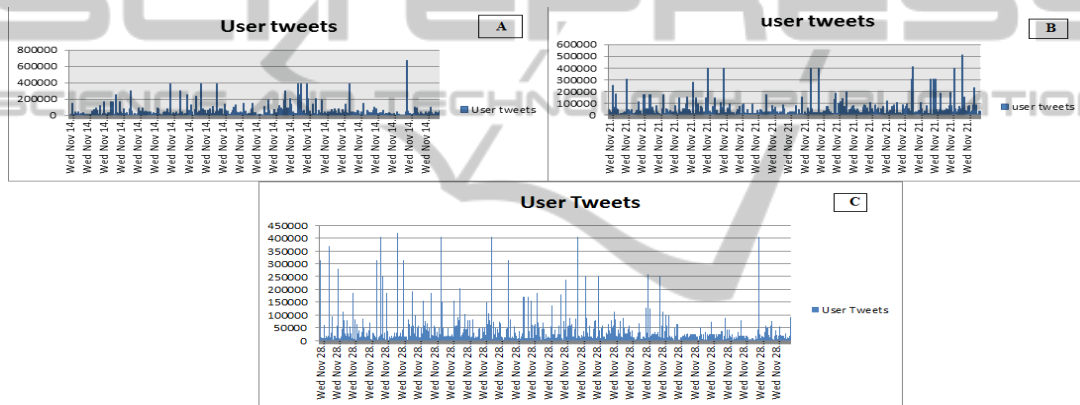


Figure 2: User Tweets per week.

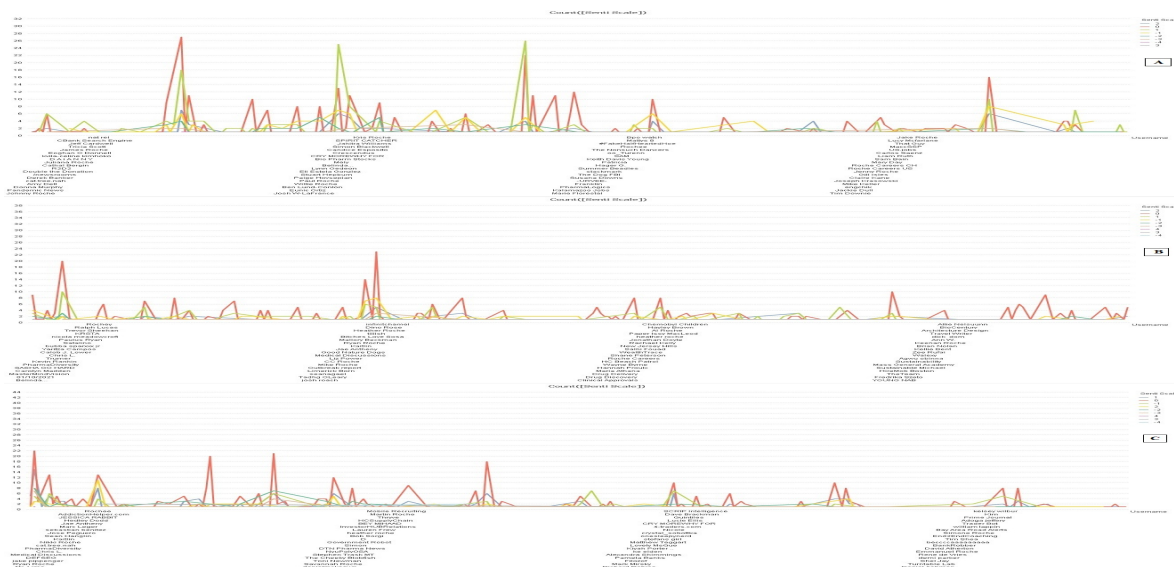


Figure 3: Tweets-Sentimental Average.

Table 2: The distribution of sentiment scores.

Time Slot	Senti-Positive	Senti-Neutral	Senti-Negative
14 th November 2012	32%	47%	21%
21 th November 2012	25%	53%	22%
28 th November 2012	27%	48%	25%

Table 3: Frequent words in tweets.

Frequent words in tweets	
Between 3000-4000 occurrences	Roche
Between 2000-3000 occurrences	Johnson & Johnson
Between 1000-2000 occurrences	Pfizer, Novartis
Between 0-1000 occurrences	Glaxo, GSK, innovations, news, marketing, yahoo, finance, healthcare, acquisition, city, advertising, business, development, manufacturing, products, research, investors, competition, consumer, customer, email, auction, collaboration, communication, contract, social network, supplier, Facebook, Google, government, outsourcing, technology, e-pharma, distributors, economics, twitter

on the number of positive and negative tweets published and the average frequency of occurrence for each organisation.

The same sentiment analysis is carried out on these terms to determine their respective time lines of sentiment. The timelines (Figures 2) can then be generated for both the companies and the categories. Tweet frequency was used as a guide to category and sub-category importance. In order to infer some of the socially based relationships, the content analysis is carrying out. The knowledge structure derived from tweets (ontology) is presented in next section.

4.2.3 Ontology and Concept Network

Ontology network is relating ontologies on the basis of explicit import relationships or implicit similarity. Concept network is relating concepts on the basis of

explicit ontological relationships or implicit similarity. As mentioned in section 3, GTM taken in order to analyse the tweets content. Figure 4 presents the process of storing, categorizing datasets.

The coding of content resulted in a number of categories and subcategories, including: Technology, Finance, Innovation, Suppliers, Government, Healthcare, Investors. For example, a financial innovation category was created with associations to acquisition, finance and investor. Subsequently, we have extracted the above mentioned ontology by folding the graph using the Protégé 4.2 OntoGraf. The same sentiment can then be report by key codes and categories.

Table 1: Senti-average per frequent Word (Organisations).

Name	#Senti-Pos.	#Senti-Neg.	Senti-Pos. percentage	Senti-Neg. percentage	Senti-Pos. Average	Senti-Neg. Average
Roche	5304	4444	56%	48%	1.33195	-1.56133
Johnson & Johnson	2134	848	23%	9%	1.363636	-1.35142
Pfizer	1058	2653	11%	28%	1.413043	-1.39224
Novartis	747	1205	8%	13%	1.570281	-1.62905
GSK	166	163	2%	2%	1.150602	-1.20245

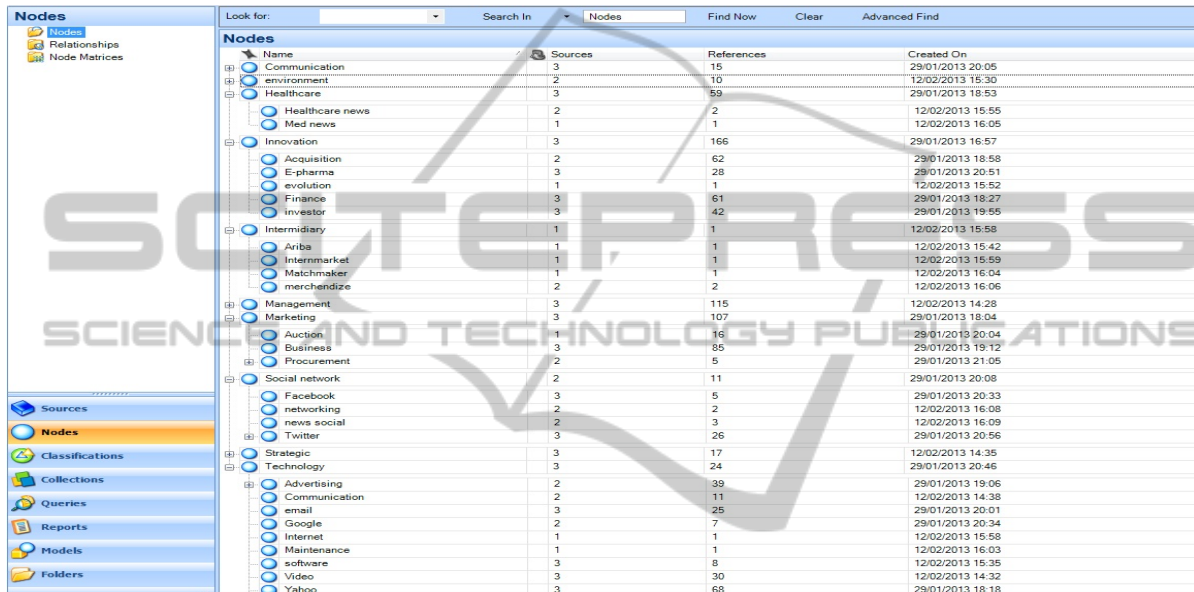


Figure 1: Tweets coding and categorizing.

5 CONCLUDING DISCUSSION

Before attempting to interpret and make sense of the findings, a caveat should be highlighted with regards to the dataset being analysed. The exploratory nature of this study and its focus on methodology meant that it was not possible to examine in further depth the actual content of the over 54,365 tweets captured. A more elaborate classification of tweet patterns is an issue of further research. Bearing this limitation in mind, it seems that pharmaceutical organizations are at the process of building an extensive LinkedIn and Twitter networks that gives them access to a diverse group of stakeholders. These same networks are freely available and provide a useful source of business intelligence – especially for small or new entrant organisations. This is indicated not only by the promising number of tweets per account, but also by the frequency of occurrence of the words, as well as the sentiment

analysis across accounts, which are usually quite influential. This diversity confirms that, beyond interactivity with organizations, LinkedIn and Twitter accounts indeed attempt to satisfy quite complicated information needs for organizations in terms of activity on social networks.

The findings of our study outline certain mixed conclusions about this ability. Firstly, as far as the content produced by the accounts is concerned, there are two main observations. The first is that the content is obviously temporal as indicated by the popular timeslot and keywords identified in the analysis. Temporal separation (weekly based) analysis on the content of Twitter dataset was carrying on in order to compare the number of tweets published in each time slot (figure 2). The result over the three snapshot times show the rise in number of tweets on second week of the month (21 November). Whereas the number of tweets per minutes in the last timeslot (28 November) is more

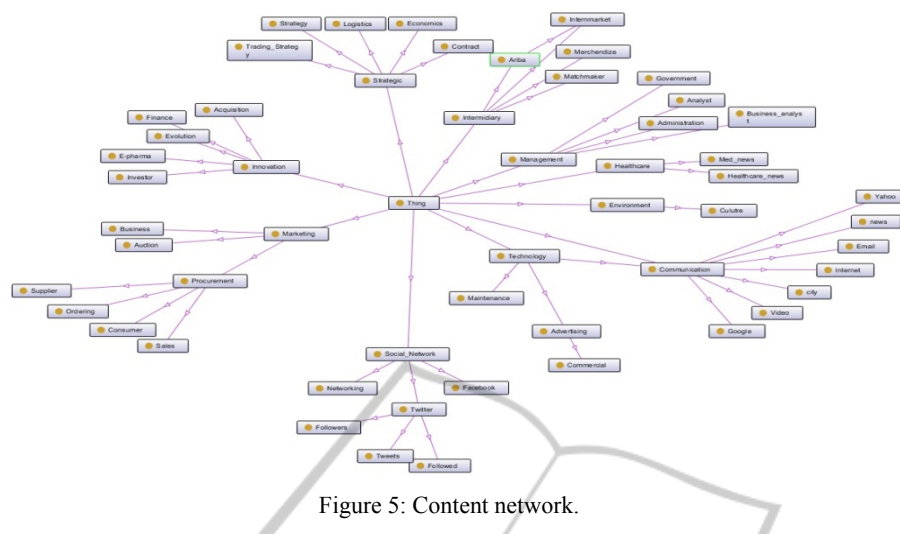


Figure 5: Content network.

transparent. Additionally our temporal separation analysis followed by sentiment information analysis on the time-stamps. We demonstrated the average of number of positive and negative tweets per each timeslot (figure 3, table 2). Sentiment analysis experiment shows the vast majority of tweets are assigned to neutral. The analysis shows tweets are slightly more positive in November 14th whereas towards end of the month the sentiments are more negative. Thirdly, further analysis was carried out on tweet content. Analysis was carried out using our own novel temporal coding technique drawing in part from the grounded theory method. In temporal coding, initially we count the number of word occurrences using Nvivo9. Word frequency analysis revealed 5 major words of those who tweet: Roche, Johnson & Johnson, Pfizer, Novartis and GSK. This categorization proved important in better understanding the type of contents those re-using such as content will be faced with. In particular we highlighted that unsurprisingly, few tweets are recognized as rich in content and therefore valuable in their own right. Alternatively, in aggregated form (by organisational, time, word or coded grouping) an interesting picture can be uncovered. This is because twitter users provide a mix of information which can't be easily distinguished by automatic means. Sentiment analysis (coupled with Grounded Theory categorisation) offers a number of opportunities to better understand the wider business networks and their language. It can be seen that Twitter offers a promising starting point for crawlers to collect related data, where time and location matter – a domain picture.

This paper has analysed datasets from a popular social media network. Such a subtle difference in social networks leads us to think more about

semantic integration to achieve interoperability among SMNs and ultimately content integration facilities on the Web. Semantic integration can provide an enhanced view of individual or organizations activities in distributed social networks. Therefore more intuitive semantic methods are required for presenting and navigating around SMN data. On the other hand, analysing data published on social network provide a unique opportunity for an SME or new entrant to the market to observe the dynamics of community development of large organizations as the data is easy, cheap and accessible.

In this study, the research approach and social networking analysis provided the opportunity to investigate the pharmaceutical organizations Twitter accounts beyond sampling a particular set of tweets or focusing only on a small subset of those accounts. Such an investigation could not be possible by applying standard research methods, which are not able to follow the pace of Internet change as organisations develop their online presence (Karpf, 2012). Therefore, research methods need to be informed accordingly so that the complex interactions being formed on social networks can be adequately understood. The Social Media Domain Analysis (SoMeDoA) framework is a key contribution of this paper and attempts to motivate more rigorous and integrative approaches to social media data analysis. The framework provides an effective approach to both selecting (and integrating) social media platforms and subsequent analysis of the data they hold. Temporal categorisation of data (including the addition of key domain concepts) before frequency and sentiment analysis provides an effective means of researching a domain of interest, marketplace or industry. The SoMeDoA approach

can be practically applied to further domains of study (or lines of inquiry), providing a flexibly method for mining new insight from readily available data.

ACKNOWLEDGEMENTS

The authors acknowledge support of this work through the MATCH Programme (UK Engineering and Physical Sciences Research Council grants numbers GR/S29874/01, EP/F063822/1 and EP/G012393/1).

REFERENCES

- Björkman, I. and Kock, S. (1995) "Social relationships and business networks: The case of western companies in China", *International Business Review*, vol. 4, no. 4, pp. 519-535.
- Buehrer, R. E., Senecal, S. and Bolman Pullins, E. (2005) "Sales force technology usage—reasons, barriers, and support: An exploratory investigation", *Industrial Marketing Management*, vol. 34, no. 4, pp. 389-398.
- Copp, C. B. and Ivy, R. L. (2002) "Networking Trends of Small Tourism Businesses in Post-Socialist Slovakia", *Journal of Small Business Management*, vol. 39, no. 4, pp. 345-353.
- De Longueville, B., Smith, R. S. and Luraschi, G. (2009) "Omg, from here, i can see the flames!: a use case of mining location based social networks to acquire spatio-temporal data on forest fires", *Proceedings of the 2009 International Workshop on Location Based Social Networks* ACM, , pp. 73.
- Glaser, B. G. and Strauss, A. L. (1967) *The discovery of grounded theory: Strategies for qualitative research*, AldineTransaction.
- Huberman, B., Romero, D. and Wu, F. (2008) "Social networks that matter: Twitter under the microscope", *Available at SSRN 1313405*, .
- Hughes, A. L. and Palen, L. (2009) "Twitter adoption and use in mass convergence and emergency events", *International Journal of Emergency Management*, vol. 6, no. 3, pp. 248-260.
- Java, A., Song, X., Finin, T. and Tseng, B. (2009) "Why we twitter: An analysis of a microblogging community", *Advances in Web Mining and Web Usage Analysis*, , pp. 118-138.
- Karpf, D. (2012) "Social Science Research Methods in Internet Time", *Information, Communication & Society*, vol. 15, no. 5, pp. 639-661.
- Kucuktunc, O., Cambazoglu, B.B., Weber, I. and Ferhatosmanoglu, H. (2012) "A large-scale sentiment analysis for Yahoo! answers", *Proceedings of the fifth ACM international conference on Web search and data mining* ACM, , pp. 633.
- Lauw, H. W., Lim, E.P., Pang, H. and Tan, T. T. (2005) "Social network discovery by mining spatio-temporal events", *Computational & Mathematical Organization Theory*, vol. 11, no. 2, pp. 97-118.
- Lee, C. H., Yang, H.C., Chien, T.F. and Wen, W.S. (2011) "A Novel Approach for Event Detection by Mining Spatio-temporal Information on Microblogs", *Advances in Social Networks Analysis and Mining (ASONAM), 2011 International Conference on* IEEE, , pp. 254.
- Lipparini, A. and Sobrero, M. (1994) "The glue and the pieces: entrepreneurship and innovation in small-firm networks", *Journal of Business Venturing*, vol. 9, no. 2, pp. 125-140.
- Masurel, E. and Janszen, R. (1998) "The relationship between SME cooperation and market concentration: Evidence from small retailers in the Netherlands", *Journal of Small Business Management*, vol. 36, pp. 68-73.
- Michaelidou, N., Siamagka, N. T. and Christodoulides, G. (2011) "Usage, barriers and measurement of social media marketing: An exploratory investigation of small and medium B2B brands", *Industrial Marketing Management*.
- Pak, A. and Paroubek, P. (2010) "Twitter as a corpus for sentiment analysis and opinion mining", *Proceedings of LREC*.
- Papacharissi, Z. (2009) "The virtual geographies of social networks: a comparative analysis of Facebook, LinkedIn and ASmallWorld", *New Media & Society*, vol. 11, no. 1-2, pp. 199-220.
- Pitt, L., van der Merwe, R., Berthon, P., Salehi-Sangari, E. and Caruana, A. (2006) "Global alliance networks: A comparison of biotech SMEs in Sweden and Australia", *Industrial Marketing Management*, vol. 35, no. 5, pp. 600-610.
- Shirzad, S. R., Bell D. (2013) "A Systematic Literature Review of Flexible E-Procurement Marketplace", .
- Skeels, M. M. and Grudin, J. (2009) "When social networks cross boundaries: a case study of workplace use of facebook and linkedin", *Proceedings of the ACM 2009 international conference on Supporting group work* ACM, , pp. 95.
- Thelwall, M., Buckley, K., Paltoglou, G., Cai, D. and Kappas, A. (2010) "Sentiment strength detection in short informal text", *Journal of the American Society for Information Science and Technology*, vol. 61, no. 12, pp. 2544-2558.
- Venkatesh, V. and Davis, F. D. (2000) "A theoretical extension of the technology acceptance model: Four longitudinal field studies", *Management science*, vol. 46, no. 2, pp. 186-204.
- Walter, A., Ritter, T., Germunden, H. and Gemunden, H. (1997) *Relationships and networks in international markets*, Pergamon Press.
- Yoo, J. and Hwang, J. (2008) "A framework for discovering spatio-temporal cohesive networks", *Advances in Knowledge Discovery and Data Mining*, , pp. 1056-1061.