TOWARD AN ANALYTICAL MODEL OF SOCIAL-MEDIA MARKETING CAPABILITIES

Hajer KEFI
Professeur, Paris School of Business / National University of Singapore
h.kefi@psbedu.paris

Talel ABDESSALEM
Professeur, LTCI & IPAL, CNRS, Télécom ParisTech, Université Paris-Saclay / National University of Singapore
talel.abdessalem@telecom-paristech.fr

Sitesh INDRA
Ingénieur de recherche, LTCI & IPAL, CNRS, Télécom ParisTech, Université Paris-Saclay / National University of Singapore
sitesh.indra@telecomparistech.fr

RESUME

In this research, we adopt a design science approach to develop an analytical framework of social media capabilities of the firm. The approach has been applied in the case of five brands in the cosmetics industry using social network analysis, consumer engagement, descriptive statistics and sentiment analysis. The findings allow us to enlighten the competitive positioning of these brands during the period under study.

INTRODUCTION

Social media platforms such as Facebook, Twitter or Instagram are now widely recognised as largely influencing social, political and business communication practices (Aral, Dellarocas & Godes, 2013). For instance, new and innovative marketing opportunities using these platforms seem limitless. Brand fan pages published by companies on social networking sites represent now a popular way to build communities around a given brand. Managers of social media marketing who operate these brand fan pages are now concerned with the popularity of their posts and with the necessity to increase the number of likes and comments. They are also highly solicited to enhance the engagement of users with their own brands comparably to competing ones.

In this article, we propose a framework adapted from Okazaki & Taylor (2013) conceptual model in which we consider that the effective use of social media in advertising and marketing depends on three capabilities on the international market place: networking capability, leveraging consumer engagement capability and leveraging consumer sentiment capability. This framework is applied within a design science approach to assess and compare the effectiveness of brand fan pages deployed on Facebook by five leading brands in the beauty industry. We use observational data related to the real activity generated on these pages by the firms and by their users. Considering the number of active users of these pages (hundreds of thousands and even millions) and the huge number of posts, we have to deal with big data analytics issues and challenges. The three components of our model are aimed to be complementary and help address the following issues: What leads a post to generate engagement within the fans' community? Who are the most influential users and posts? Are there any specific sub-communities that evolve within the fan pages and what are the motives behind their construction? And how is the brand emotionally perceived?

We focus on Facebook because of its undeniable popularity among users (2.2 billion users, STATISTICA 2018) and also because it is the most plebiscited by marketers (Stelzner, 2013). We have chosen the beauty industry as a case-study to apply our approach because the industry is recognized as one of the most dynamic sectors in ecommerce, with an ongoing increasing rate of online sales for the last decade and a large emphasis on social media communication (Łopaciuk & Łoboda, 2013).

We start hereafter with a literature review in which we define the principle domains of research we are mobilizing. Then, we present our framework and how it has been applied to fan pages belonging to five brands in the beauty industry. We finally discuss our results and their implications.

THEORETICAL BACKGROUND ON SOCIAL MEDIA MARKETING CAPABILITIES

The emergence of advertising and marketing practices on social media and their increasing popularity have opened the way to many research avenues. For Xie & Lee (2015), social media is likely to serve within a multi-channel perspective of marketing actions and is rarely deployed exclusively or separately from traditional marketing. This multi-perspective view (Trusov, Bucklin, & Pauwels 2009) has helped identifying the specificities of the social media vehicle for marketing compared to the traditional one: it is (1) reputed to be less expensive, (2) more likely to generate consumer's confidence, and (3) interactive (*versus* unidirectional for traditional vehicles). From a business analytics perspective, the required capabilities to leverage social media data are recognized as different from the conventional business analytics capabilities which take advantage mostly from business internally generated or earned data. In fact, research studies that specifically focus on social media analytics are still needed. Many recent articles have pointed out the potential usages and implications of social media but not explicitly in terms of social media analytics capabilities, with a notable exception: the Okazaki &Taylor (2013) conceptual model. The authors draw on an extensive literature review to identify three theoretical perspectives: networking capability, image transferability and personal extensibility, in order to provide powerful insights and directions concerning how to leverage social media data.

Networking capability is built on network theory and is tightly related to the social patterns of interactions initiated and enhanced by web 2.0 and social media particularly. In the case of brand fan pages, networking capability assessment has also been related in prior research to qualitative analysis of posts and comments and has been nurtured by a multidisciplinary-based (including different domains, such as sociology, computer science, mathematics and marketing) and prolific research streams dedicated mostly to the mechanisms through which different contents could be produced and shared, virally or through electronic Word-of-Mouth (e-WOM) for example. Stephen & Galak (2012) have reminded that three types of media activity can be used by marketers, offline and/or online: Paid media (advertising); Owned media (generated by the firm or its agents via controlled channels: posts, brochures, web official website, etc.); and Earned media (produced by other entities like consumers, journalists, bloggers, possibly in reaction to paid and/or owned media). We consider in this article that this joint FGC (Firm Generated Content) and UGC (User Generated Content) analytical perspective is particularly insightful to investigate brand fan pages because of their inherent structure and functionalities constructed upon firms' posts and fans reactions within a network-based community.

The second capability in Okazaki & Taylor (2013) model is image transferability, which is closely linked to the Social Media Analytics (SAM)-based ability to build a sustainable brand image. Within social media, the question here is about how to place content in order to potentially evoke favourable emotions and perceptions in a way that is beneficial for the brand sustainability. The authors have mainly discussed the issue of building a global/standard *versus*a local specific brand image in cross-national settings. We argue in this article that image transferability capabilities in the context of social media and brand fan pages especially are mainly about how to transfer a desired meaning which could be captured and shared by a community built around a given brand. If we assume, following (Laroche, Habibi & Richard, 2013), we suggest to focus on the capacity of the firm to leverage what we call consumer sentiment, i.e. the perceptions and opinions of brand fan pages' users about the promotional and advertising content displayed by the brand. We can also take profit of an increasing body of research (Chamlertwat & Bhattarakosol, 2012; Lipizzi, Iandoli & Ramirez Marquez, 2015)aimed at collecting opinions in large online audiences and communities, using for example sentiment analysis or opinion mining techniques.

Personal extensibility is the third capability in Okazaki &Taylor (2013) model and is aimed at capitalizing on consumer's desire for more interaction with the brand at any time and at any place. It is therefore related to distance factors (geographic and also cultural, economic and psychic) and to time (intensity, regularity, timeliness of interactions). In their literature review, Okazaki and Taylor have particularly discussed the cultural factors. They certainly are of a tremendous importance but have to be considered amidst other factors explaining the degree of consumers 'stickiness' to the brand. In this research, we have chosen to address the concept of consumer engagement. In the marketing/branding literature, there is a growing interest on this concept. Hollebeek & Chen (2014) define consumer's engagement as the "cognitive, emotional and behavioral investment in specific brand interactions" (p. 62). They have drawn on netnographic methodology to address the antecedents and consequences of positively *versus* negatively oriented engagement.

We have also to notice that Okazaki & Taylor (2013) have not empirically tested their model. They have defined future avenues and called for empirical studies to develop focused concepts and measures for the three components of their models. This article is a response to their call.

METHODOLOGY. IMPLEMENTATION AND FINDINGS

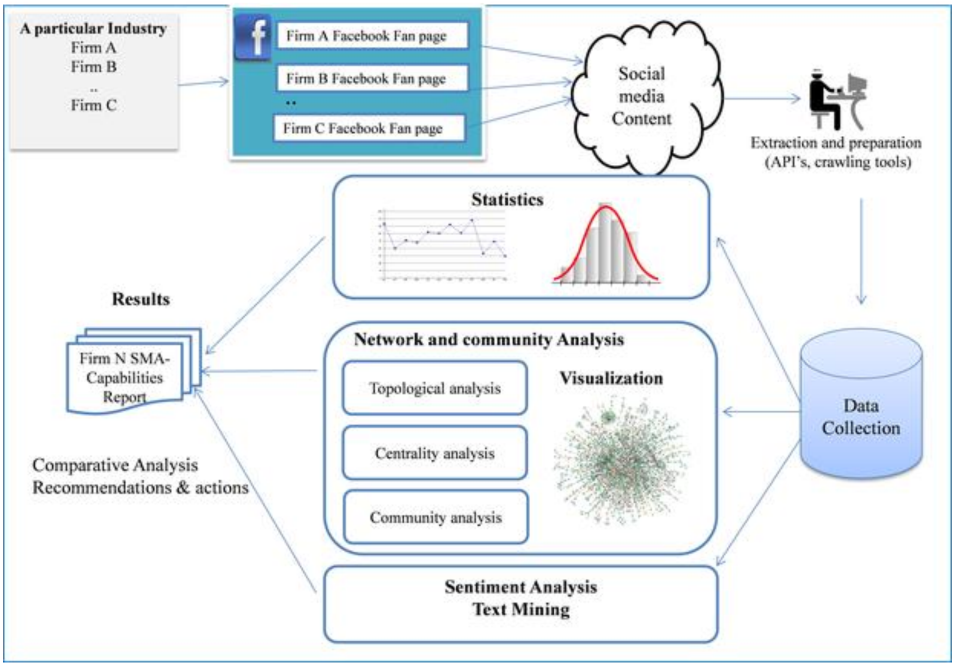
We adopt the design science methodology, defined by Hevner, March, Park & Ram (2004) as a problem-solving paradigm. According to their view, understanding a problem and finding out a solution for it are achieved in parallel during and by the construction and application of an artefact designed for this purpose, i.e. an idea, a practice, a product, a method or an analytical schema through which the analysis, development, management and use of a given information system or information technology device could be properly achieved. For Holmström, Ketokivi & Hameri (2009), a design science approach is more appropriate to exploratory than to explanatory research and might be more willing to bridge practice to theory. Hence, we argue that this view is relevant in our research.

Our framework, displayed in figure 1, consists of the following steps:

1. Identify the industry we intend to study and more specifically a group of competing brands which present a high interest in social media marketing and possess official Facebook fan brand pages.

2. Extract the data generated by all the activity realized on these pages (UGC and FGC) during a given period.
3. Analyse the data collected using social network analysis (SNA), post's life statistics and sentiment analysis (SA).
4. Interpret the findings for each brand and compare their capabilities in terms of networking, consumer engagement and consumer sentiment.

Figure 1 – SMA-based capabilities Framework



Our research is mandated by Yves Rocher (noted YR in this article) a French leading beauty brand, ranked 1st in natural and organic beauty and cosmetics products in France. We have asked the digital marketing executives of this brand to identify their direct competitors in France. Consensually, they pointed out 4 brands: L'Occitane, BodyShop, L'Oréal and Kiko (Noted respectively LOC, BS, LOR and KI). This information is confirmed by the statistical studies published in this sector. The analyses we are presenting in this article are processed using data extracted from the official international Facebook brand fan pages of these brands (referred to hereafter as: YR, LOC, BS, LOR and KI) and covering a period of six months of activity (from January to June 2016).

A data pre-processing step was needed to transform the bipartite user-post Graph (G0) provided by our data extracting tool into a user-user graph (G1), as $G_1 = G_0G_0^T$ (where G_0^T is the standard matrix transpose of G_0). This transformation insures edge weight in the generated user-user graph.

As discussed above, the networking capability of the five brands could be investigated in terms of topological, centrality and community detection analyses.

First, topological analysis gives a broad idea of the network structure. Some of the key elements in this analysis are: the number of nodes and edges, graph density, average path length and diameter. We have chosen graph density as a topological metric to compare and rank the activity of our brands. A higher density means a more intense activity generated by the fan page, and also more active users.

According to this criterion, LOR with more than 20 million registered fans is ranked third in terms of graph density, while YR and BS are ranked 1st and 2nd, with respectively more than one million and approximatively 674 200 fans. In effect, graph density is related to users' real interactions in the network and seems more appropriate to assess the volume of the activity generated by the brand pages than the number of registered fans. Identifying key roles within these real interactions through centrality analysis could also give more information about this activity.

Centrality analysis indeed aims to determine the nodes which play important roles in the network. In our case, we remind that nodes are either posts (by the page owner) or users. Centrality for a post means that it has generated an important activity performed by users, in other words, it can be considered as popular or successful. On the other hand, a central user, is a member of the community who comments, likes, shares, etc. more than the other members, and hence s/he can be considered as (more) influential. Our first intuition was to use the Eigenvector centrality for both categories because this indicator deals with the influence of a given network member in terms of weighted connectedness (Opsahl, Agneessens, & Skvoretz, 2010). The main characteristic of social network user-post graph is that they are often directed and bipartite. Straight forward implementation of Eigenvector centrality will therefore be misleading. Furthermore, if we implement Eigenvector centrality on our user – post graph, where edge is directed from user to post, it assigns zero to all the users. To deal with this issue we implement Page rank algorithm to identify the most popular posts and Eigenvector centrality on transformed user-user undirected graph to identify the most influential users.

By doing so, a given brand could choose to target the most influential users in order to more efficiently diffuse advertising contents within the network. It is also possible to identify the more successful posts (and also the less successful ones), with regard to their content (a given advertising campaign for example) or format (photo, video, quiz, etc.) and take the appropriate decisions. Consequently, making sense of the members' real interactions within the community and their respective roles is important to assess the networking capability of the brand.

Furthermore, detecting clusters within brand communities could also provide information about marketing campaigns impact. From a technical point of view, choosing the right algorithm for a specific network could be problematic. In our case, Louvain Algorithm (De Meo, Ferrara, Fiumara & Provetti, 2011), which works on modularity maximization is an acceptable solution because it fulfils two purposes. First, it offers an optimized modularity, a value that measures the density of links inside communities compared to links between communities. Second, it provides hard clustering. That is, it assigns nodes into distinct communities and hence distinct sets of interest and different themes. An important assumption we make here is that a community is likely to be built upon shared content and interests arising from some specific posts and their related comments. For the needs of interpretation, we have therefore to examine the thematic content of the posts and comments related to each community.

Results show that for all the brands, the five top clusters identified represent at least 70% (90% for LOR) of the total nodes. This means that they capture the majority of the population of users/posts within the graph and that the interests shared between these actors could be concentrated in a limited number of themes. We actually identified two broad theme types: (1) Product lines interests emerging from brand advertising campaigns and users' reactions; and (2) social interests emerging from what we can define as 'socialization' posts aimed to create and sustain social links between the brand and the users and where no product to advertise or sell are discussed. Figure 2 shows a visualization of the clusters of YR brand community using Gephi 0.9.2: the data visualization tool used in this study.

Figure 2 – Clusters Visualization representing sub-communities within the graph (Gephi 0.9.2)

Extracted data from Facebook fun pages provide us with temporal data concerning the posts and all the activities they generate, including comments, likes, etc. This gives us the possibility to measure the impact of each post by defining: (1) its life cycle, i.e. the duration between the moment it has been published and the last activity it has generated; and (2) the engagement of users, which we measure (as processed by Gephi) as the sum of the total number of comments/comments' replies and likes.

We can therefore assess the performance of each post using these two metrics, compare the results obtained by type of post (photo, video, link, etc.), with regard to the day of the week when the post is published, and this for the five brands under study.

Our results show that for all the brands, the most successful posts (in terms of duration or life cycle) are those published during the week-end (mostly on Saturday and Sunday). Whereas, on average, photos seem to generate the highest engagement. If we take a closer look at the results and rank the brands by type of post generated engagement, it appears for example that LOR is the most successful brand for photos, and that globally all the brands have difficulties to spread their advertising videos. If we consider the escalating power of video blogging in the beauty sector, this counter-performance could be considered as critical or alarming especially for YR.

Finally, sentiment analysis (SA) puts the focus on the assessment of value judgments about a given object (Chamlertwat & Bhattarakosol, 2012; Thelwall, Buckley, & Paltoglou, 2011). The analysis leads to a valence-based (i.e. positive, negative, and neutral) sentiment (or opinion) classification, which can be applied at a given semantic level (e.g. a document, a text or a sentence). In our case, it is applied to each single post, comment and comment-reply. The textual material is handled as a raw text which is put into an annotation object and sequentially annotated within an analysis pipeline, to be tokenized, cleaned, lemmatized, etc. until a score corresponding to a positive, neutral or negative status is affected. The command-line for SA at the sentence level provided by Stanford CoreNLP toolkit is particularly helpful in our case. SA scores are the following: 4 for positive, 2 for neutral and 0 for negative.

Due to space limitations, we do not present the entire procedure and all our findings. We have developed a SA data base including the sentiment polarity of each single post and comment for the five brands under study during the analysis period. It could serve to calculate the number of positive, neutral and negative contents (posts/comments/comments' replies) published by the ten most influential users. We can also calculate for any given community the number of positive, neutral and negative comments and comments' replies it has generated.

Amidst our most salient results, we have particularly noticed that top ranked posts (in terms of centrality) are sentiment-charged (i.e. they are not neutral), and that most of them are positively-charged. Concerning YR for example, if we look at the 100 highest ranked posts, we can find less than 10 negative posts (most of the time labelled as questions). We can therefore argue following previous studies (e.g. Stieglitz and Dang-Xuan 2013), that sentiment charged posts are more likely to spread than neutral ones. With regard to clusters analysis, comments exchanged within a given community can be positive, negative or neutral (no clear trend has emerged).

Moreover, it could be insightful to focus on the negative content related to a specific community and try to find out the reasons of these negative perceptions about products functionalities, delivery conditions, etc. These findings are also congruent with the literature on social media marketing and Business Intelligence (Kefi, Indra, & Abdessalem, 2016; Lipizzi et al., 2015) in which it is highly recommended to pay attention to the positive as well as to the negative value judgements of consumers. YR is for example the most commented brand and at the same time the one presenting the highest proportion of negative posts. At the contrary, KI is the least commented brand and has the lowest proportion of negative posts. This result confirms that online engagement with a brand does not necessarily mean satisfaction.

DISCUSSION AND CONCLUSION

Firms are increasingly using social media and especially brand fan pages to promote their brands and interact with their customers. Brand fan pages published on social media can provide these firms with more than online stores to display their brands. They offer the opportunity to better understand the market evolutions, to collect customer opinions and to benefit from a considerable amount of user-generated content. The urge to develop and enhance durable competitive advantages has become a day-to-day concern for many companies. Hence their need to monitor what their competitors are doing on these platforms (He, Wu, Yan, Akula, & Shen, 2015). The framework we have developed in this paper is a response to this need. It is conceived as an industry-specific SMA-based capabilities assessment tool applied to Facebook brand fan pages. Our approach uses social network analysis, sentiment analysis and statistical analysis in a complementary manner to produce business (internal) and competitive (inter-brands comparative) reports. It brings therefore a valuable contribution to the research on social media analytics, besides its practical implications as a competitive intelligence analytical tool that could be delivered to marketing analysts and decision-makers.

Our framework has been applied in the beauty industry and on Facebook brand pages. Deploying it in other industries and other social media platforms is feasible. We have already started data collection from Twitter and Instagram brand fan pages in the luxury fashion industry to apply our framework.

Finally, we strongly advocate in our research for the use of social data analytics in marketing and management information systems. More research is certainly needed to harness data generated from online and offline activities to model and predict consumers' and other social actors' attitudes and behaviors.

REFERENCES

Aral, S., Dellarocas, C., & Godes, D. (2013). Introduction to the Special Issue—Social Media and Business Transformation: A Framework for Research. *Information Systems Research*, *24*(1), 3–13. https://doi.org/10.1287/isre.1120.0470

Chamlertwat, W., & Bhattarakosol, P. (2012). Discovering Consumer Insight from Twitter via Sentiment Analysis. *J. Ucs*, *18*(8), 973–992. https://doi.org/10.1016/j.pragma.2013.12.003

De Meo, P., Ferrara, E., Fiumara, G., & Provetti, A. (2011). Generalized Louvain method for community detection in large networks. In *International Conference on Intelligent Systems Design and Applications, ISDA* (pp. 88–93). https://doi.org/10.1109/ISDA.2011.6121636

He, W., Wu, H., Yan, G., Akula, V., & Shen, J. (2015). A novel social media competitive analytics framework with sentiment benchmarks. *Information & Management*, *52*(7), 801–812. https://doi.org/10.1016/j.im.2015.04.006

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). Design Science in Information Systems Research 1. *Design Science in IS Research MIS Quarterly*, *28*(1), 75–105. https://doi.org/10.2307/25148625

Hollebeek, L. D., & Chen, T. (2014). Exploring positively- versus negatively-valenced brand engagement: A conceptual model. *Journal of*

Product and Brand Management, *23*(1). https://doi.org/10.1108/JPBM-06-2013-0332

Holmström, J., Ketokivi, M., & Hameri, A. (2009). Bridging practice and theory: a design science approach. *Decision Sciences*, *40*(1), 65–88. https://doi.org/doi:10.1111/j.1540-5915.2008.00221.x

Kefi, H., Indra, S., & Abdessalem, T. (2016). Social media marketing analytics: A multicultural approach applied to the beauty & cosmetics sector. *Pacific Asia Conference on Information Systems, PACIS 2016 – Proceedings*, *Paper 176*. Retrieved from https://www.scopus.com/inward/record.uri?eid=2-s2.0-85011103223&partnerID=40&md5=7a7fceb7990736f0102f9c82d0e64fff

Laroche, M., Habibi, M. R., & Richard, M.-O. (2013). To be or not to be in social media: How brand loyalty is affected by social media? *International Journal of Information Management*, *33*(1), 76–82. https://doi.org/10.1016/j.ijinfomgt.2012.07.003

Li, X., Xie, H., Chen, L., Wang, J., & Deng, X. (2014). News impact on stock price return via sentiment analysis. *Knowledge-Based Systems*, *69*(1), 14–23. https://doi.org/10.1016/j.knosys.2014.04.022

Lipizzi, C., Iandoli, L., & Ramirez Marquez, J. E. (2015). Extracting and evaluating conversational patterns in social media: A socio-semantic analysis of customers' reactions to the launch of new products using Twitter streams. *International Journal of Information Management*, *35*(4), 490–503. https://doi.org/10.1016/j.ijinfomgt.2015.04.001

Łopaciuk, A., & Łoboda, M. (2013). Global Beauty Industry Trends in the 21st Century. *Knowledge Management & Innovation Knowledge and Learning*, 1079–1087. Retrieved from http://www.toknowpress.net/ISBN/978-961-6914-02-4/papers/ML13-365.pdf

Okazaki, S., & Taylor, C. R. (2013). Social media and international advertising: theoretical challenges and future directions. *International Marketing Review*, *30*(1), 56–71. https://doi.org/10.1108/02651331311298573

Opsahl, T., Agneessens, F., & Skvoretz, J. (2010). Node centrality in weighted networks: Generalizing degree and shortest paths. *Social Networks*, *32*(3), 245–251. https://doi.org/10.1016/j.socnet.2010.03.006

Stelzner, M. A. (2013). *2012 Social media marketing industry report*. *Social Media Examiner*. Retrieved from http://www.socialmediaexaminer.com/social-media-marketing-industry-report-2013/

Stephen, A. T., & Galak, J. (2012). The Effects of Traditional and Social Earned Media on Sales: A Study of a Microlending Marketplace. *Journal of Marketing Research*, *49*(5), 624–639. https://doi.org/10.1509/jmr.09.0401

Stieglitz, S., & Dang-Xuan, L. (2013). Emotions and Information Diffusion in Social Media—Sentiment of Microblogs and Sharing Behavior. *Journal of Management Information Systems*, *29*(4), 217–248. https://doi.org/10.2753/MIS0742-1222290408

Thelwall, M., Buckley, K., & Paltoglou, G. (2011). Sentiment in Twitter events. *Journal of the American Society for Information Science and Technology*, *62*(2), 406–418. https://doi.org/10.1002/asi.21462

Trusov, M., Bucklin, R. E., & Pauwels, K. (2009). Effects of Word-of-Mouth Versus Traditional Marketing: Findings from an Internet Social Networking Site. *Journal of Marketing*, *73*(5), 90–102. https://doi.org/10.1509/jmkg.73.5.90

Xie, K., & Lee, Y. (2015). Quantifying the Impact of Earned and Owned Social Media Exposures in a Two-stage Decision Making Model of Brand Purchase. *Journal of Management Information Systems*, *32*(December), 204–238. https://doi.org/10.1080/07421222.2015.1063297