1

Understanding the user generated content and interactions on a Facebook brand page

Irena Pletikosa Cvijikj* and
Florian Michahelles

Information Management,
ETH Zurich,
Scheuchzerstrasse 7, 8092 Zurich, Switzerland
Fax: +41-44-632-1740
E-mail: ipletikosa@ethz.ch
E-mail: fmichahelles@ethz.ch
*Corresponding author

Abstract: Social networks provide the technological platform for individuals to connect, produce and share content online. As such, they are becoming an additional marketing channel that could be integrated with the traditional ones as a part of the marketing mix. In order to contribute to the companies' understanding of utilisation of social networks as a marketing platform, our study approaches the problem from the perspective of understanding the user generated content and the interactions. First we analyse the topics, intentions for participation and emotions shared by the users on a Facebook brand page. Second, we analyse the activities and interactions, in terms of their evolution over time and dependency on the community size. Finally, we discuss the implications of the obtained results for social media marketing.

Keywords: social networks; Facebook; social media marketing; SMM; content analysis; social network analysis; SNA.

Reference to this paper should be made as follows: Cvijikj, I.P. and Michahelles, F. (xxxx) 'Understanding the user generated content and interactions on a Facebook brand page', *Int. J. Social and Humanistic Computing*, Vol. X, No. Y, pp.000–000.

Biographical notes: Irena Pletikosa Cvijikj is a PhD candidate and Senior Researcher at the Swiss Federal Institute of Technology Zürich (ETH Zürich) since December 2009. She holds an MSc in Electrical Engineering from the SS. Cyril and Methodius University in Skopje, Macedonia. She has ten years industry experience in the area of software engineering, working as Line Manager, Project Leader and Technical Coordinator for numerous software projects, mostly in the area of telecommunication systems. Her current research is focused on usage of Facebook as a platform for social media marketing.

Florian Michahelles is the Associate Director of the Auto-ID Labs Zurich/St. Gallen and Manager of the Labs of Professor Fleisch at the Department of Management, Technology, and Economics at ETH Zürich. He holds a PhD from ETH Zurich for his research in participative design of wearable computing applications and the development of innovative business cases for ubiquitous computing. His research interests include following topics centred on the internet of things: RFID applications, extending the EPC architecture framework for sensing capabilities, technical approaches against anti-counterfeiting, and RFID applications for the end-consumer.

1   Introduction

Marketing has recently undergone significant changes in the way information is delivered to the customers (Brandt, 2008). Social networks (SN), as a form of social media (SM), provide the technological platform for individuals to connect, produce and share content online (Boyd and Ellison, 2007). The rise and continued growth of SNs have attracted the interest of companies who see the potential to transmit their marketing messages to customers, enter into a dialogue with them using word-of-mouth (WOM) principles and to use the SNs to gain a better understanding of their customers.

Social media marketing (SMM), also known as WOM marketing or viral marketing, is the intentional influencing of consumer-to-consumer communications by professional marketing techniques. This is not to be seen as a replacement for the traditional marketing techniques but rather as an additional marketing channel that could be integrated with the traditional ones as a part of the marketing mix. The advantage of this new electronic channel is that it can be used to communicate globally and to enrich marketing toward consumers at the personal level (Brandt, 2008). Through users' feedback or by observing conversations on SM, companies can learn about customers' needs, potentially leading to involvement of members of the community in the co-creation of value through the generation of ideas (Palmer and Koenig-Lewis, 2009).

Companies, across all industries, are starting to understand the possibilities of SMM. They have evolved their approach to customers, shifting from traditional one-to-many communication to a one-to-one approach and offering assistance at any time through SNs, such as Facebook, Twitter, etc. (Gordhamer, 2009). Still, viral marketing on SNs has not yet reached the high expectations set (Clemons et al., 2007). Although many brand pages have already been created on SNs, how these pages are being used, what their potentials are and how consumers interact, remains largely unknown (Richter et al., 2011).

To contribute in this direction we present the results of a case study aiming at understanding the

1   user generated content

2   interactions on a Facebook brand page.

We examine the topics, participation intentions and sentiment shared on a brand page. In addition, we explore the interaction patterns and their evolution over time. From our results we derive implications for SMM.

The continuation of this paper is structured as follows: Section 2 provides an overview of the related work. Section 3 introduces a Facebook brand page used as a source of data. Section 4 describes the methodology. Section 5 presents the results, while

Section 6 derives the implications for SMM. We conclude the whole paper with Section 7.

2   Related work

A SN can be defined as "web-based services that allow individuals to (1) construct a 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" [Boyd and Ellison, (2007), p.1]. Since their introduction in 1997 with SixDegrees.com, SNs have attracted millions of users, becoming an integral part of their daily routines (Cassidy, 2006). At the time of writing, Facebook is the largest SN with more than 800 million active users (Facebook Statistics).

SNs and Facebook have been studied from different perspectives such as usage patterns (Lampe et al., 2006), motivations (Joinson, 2008), characteristics of communities (Nazir et al., 2008), etc. However, little has been published about the use of SNs in the context of companies, though SNs can be applied in three distinct areas: "(1) recruiting and professional career development, (2) relationship facilitation in distributed work contexts, and (3) business-to-customer interactions" [Richter et al., (2011), p.97].

According to Harris and Rae (2009) SNs may play a key role in the future of marketing; they may increase customers' engagement, as well as help to transform the traditional focus on control into a collaborative approach more suitable to the modern business environment. User participation as a main feature of the SNs imposes challenges to the traditional one-way marketing, resulting in companies experimenting with many different approaches, thus shaping a successful SMM approach based on the trial-and-error experiences (Coon, 2010).

In order to provide insights to practitioners looking to use SNs to benefit their brands, previous studies in the field have focused on users by trying to identify the most influential target group (Li, 2007) or explain their relation to the SM (Agozzino, 2010). Others have addressed the challenges of SMM such as aggressive advertisement and over-commercialisation, lack of e-commerce abilities, invasion of user privacy and transparency control (Bolotaeva and Cata, 2010; Harris and Rae, 2009). An inappropriate approach to these challenges could lead to fan loss and exposing the company to the risk of destroying its own credibility.

Apart from the challenges, many opportunities have also been recognised, such as raising public awareness about the company, community involvement and gathering experience for the future steps (Bolotaeva and Cata, 2010). Weston (2008) argues that SMM can also help find talent, find new customers, and help conduct brand intelligence and market research. In addition, as Javitch (2008) advises, free SMM is a good alternative to the costly traditional marketing campaigns.

Based on exploratory findings and practical examples scholars have tried to generate guidelines for successful SMM. Guidelines that apply for online WOM can also be applied to SMM:

1   sharing the control of the brand with consumers

2   engaging them in an open, honest, and authentic dialog (Brown et al., 2007).

Furthermore, Li (2007) recommends that companies need to build an engagement plan before diving into the SMM in order to appropriately approach the frequent users who are most likely to virally spread their enthusiasm for a new product or service. This might be achieved by

1   focusing on a conversation

2   developing a close relationship through brand pages

3   finding out what interactions and content will keep users coming back (Li, 2007).

Similarly, Kozinets et al. (2010) argue that classification of different types of content is the first step towards understanding the conversation.

Related to the understanding of the conversation within the SM platforms is the field of opinion mining and sentiment analysis. The term opinion mining describes tools for text processing used to generate a list of attributes and aggregate opinions about them (Dave et al., 2003). Approximately at the same time the term sentiment was used in reference to the automatic analysis of evaluative text and the tracking of predictive judgments (Das and Chen, 2007), thus the two disciplines are interconnected.

The importance of opinion mining for SMM has been recognised (Nasukawa and Yi, 2003). Opinion mining allows companies to conduct analyses of UGC to determine how the population perceives a given brand, product or feature, i.e., for market analysis and rumour detection. This resulted in a research over long structured text discussions from blog posts for prediction of sales, stock market activity, etc. (De Choudhury et al., 2008; Liu et al., 2007).

The change in SM towards short commentary, as introduced by SNs Twitter and Facebook, resulted in a significant difference in the comment structure and language that may affect the accuracy of opinion mining techniques (Simm et al., 2010). Recent studies focus on the SN Twitter to investigate the value of tweets as online WOM, possibilities for movie revenue prediction, etc. (Asur and Huberman, 2010; Jansen et al. 2009). Still, in the domain of Facebook there are many open questions on how different companies could fit in with and adhere to the unwritten rules of engagement with the SN communities (Richter et al., 2011).

In order to contribute to the company's understanding of utilisation of SNs as a marketing platform, our study goes beyond a practitioners report containing simple metrics such as number of fans, posts, etc. Instead, we approach the problem from the perspective of understanding the usage of these platforms. First we analyse the content of the posts shared on a Facebook brand page. We focus on identification of the topics referred to within the posts, intentions for participation, and emotions that people share within the posts. Second, we identify the patterns of user interaction which can be used as criteria for user classification. We investigate the changes in the activity level that occur over time with the growth of the community. Finally, we perform social network analysis (SNA) over the interactions network to see how it evolves over time. We discuss the implications of the obtained results for SMM.

3   Background: Facebook page of ok.- brand

The selection of Facebook as an underlying platform was based on the reasoning that Facebook is currently the largest SN. In this section we first provide a brief introduction

on the concepts and terminology relevant for Facebook as a marketing platform. Further, we introduce a case study of the ok.- brand page used as a source of data.

3.1   Facebook as a platform for SMM

Facebook provides four possibilities for companies to utilise the platform for marketing purposes:

1   Ads

2   pages

3   connect and open graph protocols

4   applications.

Of these, Facebook pages provide the possibility for direct engagement with the consumers through dialog, resulting in UGC which reveals consumers' interests, needs and opinions.

   In order to define the terminology, we will describe the concepts used in this paper based on the current definitions from Facebook. Although 'like page' is the official name for all Facebook pages which are not user profile pages, we will use the common terminology 'brand page' in order to distinguish pages created and operated by brand owners. The content shared on brand pages is referred to as 'posts' and appears on the central part of the page, known as the 'wall'. Each page might have one or more administrators responsible for creation and deletion of content, i.e., the 'moderators'. A brand page can have any number of members, in the continuation referred to as users or 'fans'.

Figure 1   The ok.- Facebook brand page (see online version for colours)

Note: Screen shot captured on 25.10.2011.

3.2 ok.- brand page

ok.- is a Swiss brand for fast moving consumer goods (FMCG), placed on market in 2009. This particular brand was selected for this study for the reason of having the possibility to access its full data since the first day of creation of its Facebook brand page. At the time of writing the page counts 62,164 fans, as illustrated on Figure 1.

4   The method

The method used for this study consists of two steps (1) data collection and (2) data analysis (see Figure 2). Detailed description is presented in the continuation.

Figure 2   The method overview (see online version for colours)
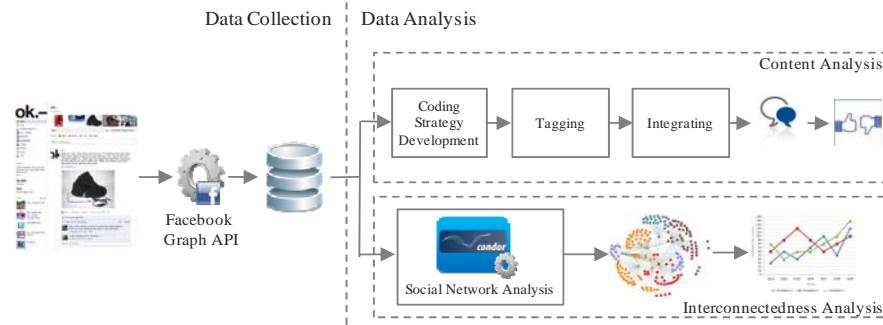
4.1 Data collection

The dataset used for this study consists of posts shared on the ok.- brand page. The data collection was employed from July 2010 to March 2011, and data from the official launch of the page in March 2010 to July 2010 were fetched retroactively to ensure a complete dataset. To guarantee accuracy of the data posts were fetched on a daily basis, using a script utilising the Facebook Graph API (Facebook Graph API).

The graph API provides access to Facebook social graph via a uniform representation of the objects in the graph (e.g., people, pages, etc.) and the connections between them (e.g., friends, shared content, etc.). For purposes of this study we have used the *feed* connection of the *page* object. *Feed* connection represents a list of all *post* objects containing the post details, i.e., the message, post type, likes, comments, time of creation, etc. The extracted elements were stored in a database for further investigation.

From the total of 759 posts shared on the ok.- brand page during the selected period, 134 posts were published by the page moderator, and the remaining 625 by the fans. Of those, three were removed due to the difficulty in recognising the used language and an additional 11 after being labelled as spam, leaving 611 user posts for qualitative analysis.

4.2 Data analysis

In order to perform the analysis of the data, two main methods were applied:

1   action-objet approach, for analysis of the content

2   SNA, for understanding the user interactions.

4.2.1 Content analysis

The ok.- brand is present only in the German speaking part of the Switzerland. Although German is the official language in this part of the Switzerland, in daily conversation a dialect is used, know as Swiss German, which significantly differs from High German, with large variations within different parts of Switzerland [e.g., 'breakfast' -> 'Frühstück' (High German) -> 'Z'Morge', 'Morgenässe' (Swiss German)]. In addition, there are no standardised rules for the written form. This introduces challenges to the existing automatic opinion mining techniques.

A supervised approach commonly used for opinion mining is based on the usage of a labelled training corpus to learn the relevant classification function based on the feature extraction and similarity calculation. In case when the same feature differs in form and writing, the precision of the results reduces. To overcome the described challenges we have applied the action-object approach for post classification (Zhang and Jansen, 2008) which represents a manual process of text analysis, following the coding development strategy (Glaser and Strauss, 1967). The process consisted of three steps:

- *Coding strategy development:* a manual coding scheme was created to define the classification rules, e.g. 'mention of a brand name -> '*brand*', 'post in form of question -> '*information inquiry*', etc.

- *Tagging:* using the defined coding strategy, each of the posts was assigned one or more tags to identify the key concepts in the content, e.g., '*my favourite: ok.- chocolate bar, when will there be a jumbo pack?*', was tagged as '*product*', '*positive affect*', '*information inquiry*', '*suggestion*' and '*package*'.

- *Integrating:* based on the tagging results, grouping of similar tags was performed in order to define groups of topics and categories. For the previous example, the resulting group descriptors were:

  1   'product: affect: positive'

  2   'sales: availability: launch date: information inquiry'

  3   'product: feature: package: suggestion'.

This approach captured the following aspects of the posts:

1   topics referred to within the posts

2   intentions for participation, further denoted as post categories

3   sentiment present in the content.

4.2.2 Interconnectedness analysis

The main research method applied for understanding the interactions between the users and the moderator on the ok.- brand page was SNA. The goal was to understand the evolution of the interaction network, as well as the dependency between the interaction characteristics and the size of the community, i.e., total number of fans. For that purpose

a dynamic social network analysis (DSNA) was employed, based on the utilisation of the tool Condor (Condor). DSNA provided the possibility for temporal visualisation of the network structure and measures, i.e., visualisation of the network evolution over time (Gloor et al., 2004).

To describe the characteristics of the interaction network we have applied the following measures form the SNA theory (Wasserman and Faust, 1994):

- *betweenness centrality*, is a measure of node's centrality, determined by the number of shortest paths between other nodes it belongs to, indicating node's importance to the network

- *degree centrality* is a number of ties a node has, indicating the level of interaction the node has with other nodes from the network.

When applied to the network as a whole, these measures are commonly referred to as a network centrality or centralisation. Network centrality quantifies the "tendency of the single point to be more central than all other points in the network" [Freeman, (1979), p.227]. As such, it is measured by the differences between the centrality values of the most central node and those of all other nodes, thus quantifying the variations among individual nodes. The closer the value is to 0, more uniform the users' behaviour is.

In addition, we have also measured the:

- *Group density*, i.e., the proportion of existing ties between nodes relative to the maximal possible number of ties in the network. The higher density (closer to 1), indicates existence of larger number of ties, thus also a greater degree of interaction among the fans.

The network elements, i.e., nodes and ties will be described later, once the related terminology has been defined.

5   Results

5.1   Content analysis

Content analysis revealed the

1   topics

2   intentions for participation

3   sentiment present in the user posts.

In the continuation we present the results obtained for each of these aspects.

5.1.1   Topics of conversation

In order to understand the interests of the fans on a Facebook brand page, we have analysed the posts to identify what is being talked about, i.e., to identify the topics of conversation. Our analysis identified seven major topic groups:

1   *product* (318 posts, 52% of the total)

2   *sales* (79, 13%)

3   *brand* (46, 8%)

4   *competitor* (26, 4%)

5   *Facebook contest* (20, 3%)

6   *company* (11, 2%)

7   *environment* (3, 0%).

The distribution of posts over topic groups is illustrated on Figure 3.

Figure 3   Topic groups (see online version for colours)
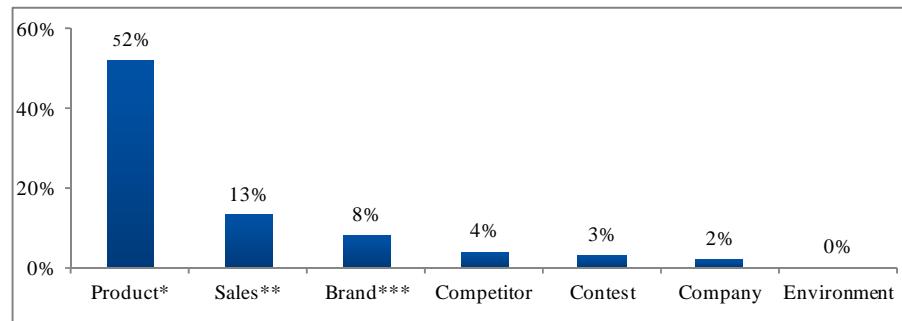
Notes: *$p < 0.0001$, **$p < 0.005$, ***$p < 0.05$

To evaluate the differences in proportions of posts belonging to different topic groups, we have applied the two-proportion Z-test for statistical analysis by comparing each topic group with the 'following' one, e.g., *product* vs. *sales*, *sales* vs. *brand*, etc.

Figure 4   Sub-topics (see online version for colours)
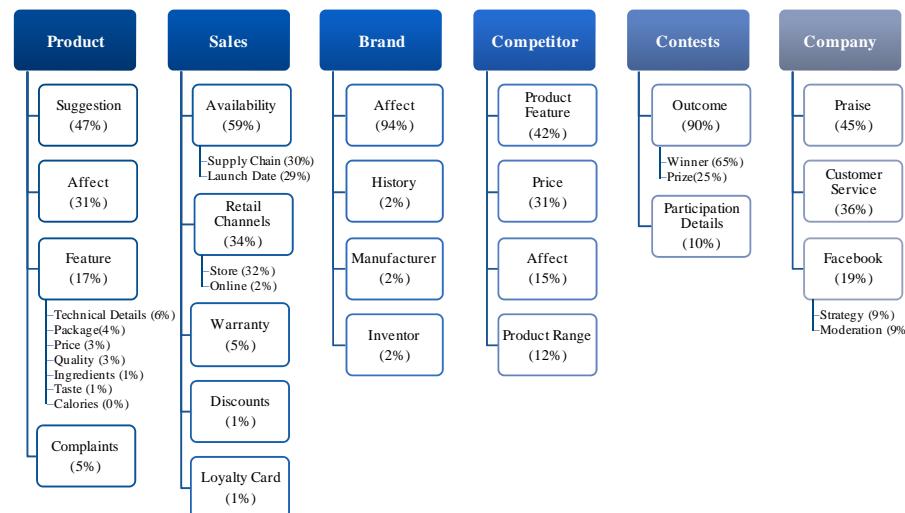
Note: Distribution values of sub-topics is relative to the topic group.

The results of the statistical analysis have shown that there is a significant difference (z = 14.539, p < 0.0001) in the proportion of *product* posts compared to the *sales*. Furthermore, the proportion of posts regarding *sales* is significantly larger (z = 3.02, p < 0.005) compared to the proportion of *brand* posts, while the number of *Brand* mentions is significantly larger (z = 2.306, p < 0.05) compared to the number of *competitor* posts. Thus, *product*, *sales* and *Brand* topic groups are larger compared to all of the topic groups with the smaller number of mentions. For the remaining topic groups, no significant difference was found to exist (p > 0.05).

The applied action-object approach provided the possibility to understand the content with a better granularity by dividing each topic group into more specific sub-topics as presented on Figure 4.

In the following text we explain the topic groups and discuss the most frequent subtopics and their value from the marketing perspective.

Product was the most frequently referred to topic group. Within this group, the two most common sub-topics were new *product suggestion* (149, 47%) and expressing *product affect* (99, 31%), e.g.: '*I want ok.- Apfelsaftschorle!*'. From the marketing perspective understanding both of these subcategories is important, the first one for sensing the needs of the customers and the second one as an indicator of product acceptance.

The two most discussed sub-topics within the *sales* group were *Availability* (47, 59%) and *retail channels* (26, 34%), e.g.: '*will you soon open an ok.- store?*' These topics can be used for evaluation of the purchase intention and can further be mapped to the specific geographical location.

Affect was the most common sub-topic (43, 94%) within the *brand* topic group, e.g.: '*i <3 ok.-*'. Since this is where users express their emotions, it can be used as the measure of the brand strength and awareness.

References to *competitors* occurred in 4% of the posts in total, mostly as comparison to specific *product feature*, but also to express *affect*, e.g.: '**Red Bull is not ok.-*'. This topic group could reveal the specific aspects of the products that users compare and expect to get from the brand, such as price, taste, etc.

Facebook contest refers to the responses by users to the entertainment actions undertaken by the moderators, e.g.: '*will there soon be another photo contest?*' The goal of organising contests is to engage the users, thus increasing the level of activity on the page, but is not of specific interest from the marketing perspective.

Company is an umbrella topic for the posts related to *customer service*. It includes *praise* for the company, as well as for the Facebook strategy and moderation style, e.g.: "*I think it's cool that you always write back to answer...*" It is an indication of the 'job-well-done' for the moderator and the company.

Environment topic occurred in only three posts, yet with a high intensity, indicating that it should be taken seriously, e.g.: "*Why doesn't one emerging brand like ok.- focus on environment friendly material???...*"

Not all of the posts were classified as belonging to a topic. Some of them (20%) were written in a form of a 'word play' or slogan, thus belonging to no particular topic, e.g., '*ok...in the end.*' These are further referred to as *general*.

5.1.2 Intentions for participation

When focusing on the 'action' element of the applied method, eight post categories, which served as indication for the participation intentions, were identified:

1 suggestions and requests (170, 27%)

2 affect expression (169, 27%)

3 sharing (165, 27%)

4 information inquiry (98, 16%)

5 complaints and criticism (23, 4%)

6 gratitude (22, 4%)

7 competitor comparison (22, 4%)

8 praise (5, 1%).

Distribution of the posts over categories is visualised on Figure 5.

Figure 5 Post categories (see online version for colours)
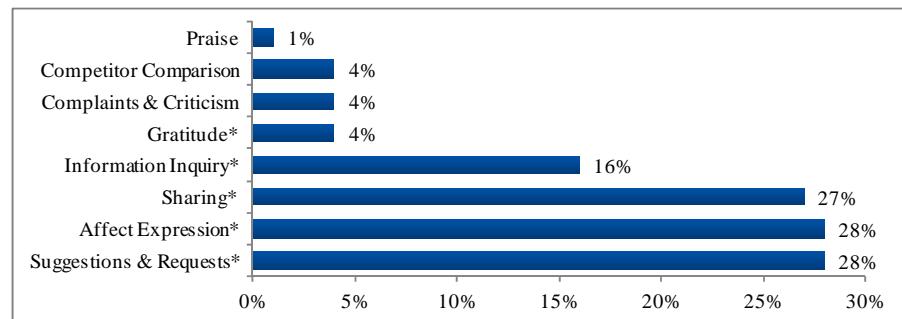
Note: *p < 0.0001

The most common intentions for posting were found to be *suggestions and requests*, *affect expression* and *sharing*, showing no significant differences in proportions. When compared to the *information inquiry*, *sharing* occurred in a significantly larger number ($z = 4.592$, $p < 0.0001$). Thus, *sharing* also occurred in a significantly larger number compared to all of its 'following' topics. Since *suggestions and requests* and *affect expression* have a slightly larger number of occurrences compared to *sharing*, the same reasoning applies. Furthermore, the proportion of *information inquiry* (16%) is significantly larger ($z = 7.09$, $p < 0.0001$) compared to *complaints and criticism* (4%), thus also compared to the remaining categories with smaller number of occurrences. Finally, the proportion of *gratitude* (4%) is significantly larger ($z = 3.111$, $p < 0.0001$) compared to the proportion of *praise*.

Each of the identified categories is explained and discussed in the following text.

Within *suggestions and requests*, two related topic groups that occur are *sales* and *product*. Fans tend to give suggestions for products, new ones as well as improvements of specific features of existing products, e.g.: '*please an ok.- energy drink pineapple*'.

Affect expression was mostly targeted towards the *brand*, a *product* or a specific *product feature*, e.g.: '*I don't like the ok.- mango*'. Posts belonging to this category give clear indication on how the brand is perceived by the users, what the popular products are and which specific features are the ones that are favoured by the consumers.

Our study indicates that within a Facebook brand page fans *share*:

1   activity

2   advice

3   opinion

4   intention

5   need

6   information

7   feelings

8   reflection on specific events

9   rhetorical questions.

A very common form of expression was a 'word play' or a slogan (92, 15%), e.g.: '*One ok.- a day, keeps the dok.-ter away*'.

Information inquiry is an important category from the organisational aspect of the Facebook brand page. It identifies possible domains of interest of the fans, e.g.: '*Who invented the ok.- brand?*' This insight reveals a need for different sources of specific information requested by the fans. Topics referred in this category include

1   brand

2   product

3   sales

4   company

5   contests.

Complaints and criticism offer the possibility for improvement that would result in greater customer satisfaction. The most complaints belonged to the *Product* topic, shared after the launch of a mobile phone product line which had several technical problems at the beginning, e.g.: '*Today I bought a ZTE San Francisco: dead on arrival...*' product availability, i.e. *Sales* was the second most referred topic in form of asking for a product delivery to a given location.

Gratitude was mostly shown in case of winning a prize in one of the organised *Contests*. However, those of interest from the marketing perspective expressed gratitude toward product launch (*Sales*) providing initial insight into the acceptance of the product, e.g.: '*I am happy that there is now an energy drink in the light version! Thank you ok.-*'

Finally, *praise* was targeted towards the subtopics of the *company* topic group.

Since topics (i.e., objects) and categories (i.e., actions) are interconnected, a matrix representation, as illustrated with Table 1, is suitable to identify most common combinations.

Table 1   Topic-category combinations and co-occurrence frequencies

	Product	*Sales*	*Brand*	*Competitor*	*Contests*	*Company*	*Environment*	*General*
Requests and suggestions	24%	3%					0%	
Expressing affect	20%		7%	0%				
Sharing								27%
Information inquiry	5%	8%	0%		2%	0%		0%
Complaints and criticism	3%	0%				0%	0%	
Expressing gratitude	0%	1%			2%	0%		0%
Praise						1%		
Comparison				4%				

Apart from *sharing* (27%), *product – requests and suggestion* (24%) and *expressing affect* towards the *products* (20%) were found to be the most dominant topic-category combinations. These numbers confirm that Facebook supports the identified goals of SMM, as presented in the related work section.

5.1.3 Sentiment analysis

A final view on the existing data was an analysis conducted to determine how users feel about the ok.- brand or the products. This understanding was gained by manual categorisation of the sentiment shared within the posts from the *Affect Expression* category.

Table 2 shows the frequency of occurrence of positive and negative sentiment in posts.

Table 2   Sentiment analysis contained within the affect expression category

Sentiment	*Occurrences (all)*	*Percentage (all)*	*Percentage (sentiment only)*
Positive	150	25%	93%
Negative	11	2%	7%
Neutral	450	73%	-

These results show that positive sentiment was shared far more often (25% of the total posts) compared to negative sentiment (2%).

5.2 Interaction analysis

Interaction analysis over the ok.- brand page enabled

1   classification of users based on the interaction patterns

2   insights into the evolution of the level of interaction

3   understanding the effect of the community size over the structural characteristics of
     the underlying brand community.

In continuation we present and discuss the obtained results.

5.2.1 *User classification based on interaction patterns*

The most general user classification on online platforms distinguishes between
participants who create content, i.e., *posters*, and those who might read the content but do
not post, i.e., *lurkers* (Nonnecke and Preece, 2000). Translated to Facebook, this
classification could further be expanded based on the available options for interaction on
brand pages, i.e., fans can

1   share posts on the wall

2   comment on existing post

3   'like' the post.

Based on this reasoning, active fans could be divided into the following three classes:

- *posters* are fans that contribute by sharing posts on the wall

- *commenters* respond to existing posts with a comment

- *likers* indicate interest in an existing post by pressing the 'like' button.

Since users can clearly belong to more than one class, for the analysis purposes the
highest one was selected, e.g., a user who interacted by liking a post, but also by writing a
comment was categorised as being a *commenter*.

Figure 6 illustrates the distribution of fans over different classes of users. *Lurkers*
were intentionally left out in this picture since they represented 98% of the fans at the end
on the observation period, which is high compared to the 90% predicted (Nonnecke and
Preece, 2000).

Figure 6   Number of active fans as a proportion of total number of fans (right axis) and individual
            classes of fans as a proportion of active fans (left axis) (see online version for colours)
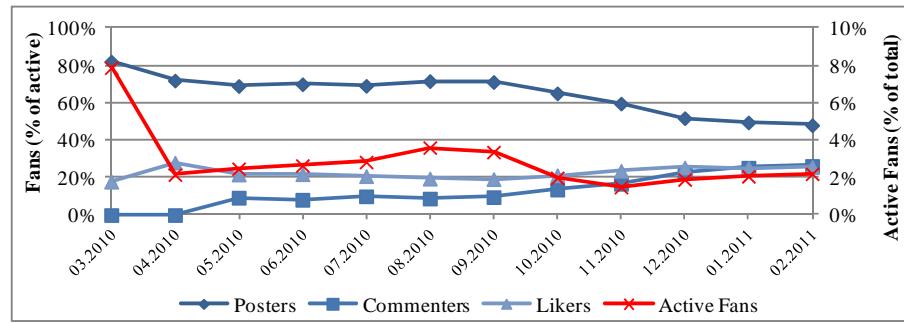
Measured by absolute numbers, the number of active users increased during the selected
period, from 28 in March 2010, to 853 in February 2011. However, percentagewise the
situation is the opposite. With the growth of the total number of fans, the percentage of
active fans was reduced, from 8% to 2% of total number of fans.

Furthermore, the level of activity measured in terms of fans distribution over classes of active users was also reduced. From initial 82% of *posters*, in relation to the number of active fans, this class was reduced to 48% at the end of the study. In turn, the percentage of *commenters* and *likers* slightly increased. From initial 0%, through 9% in June 2010, at the end of the study the *commenters* have reached the number of 26% of the active users. In case of *likers*, the numbers did not show large variations, ranging between 18% and 28% of the active users.

These results comply with the existing research from sociology which indicates that an increase in the size of the SN has negative effect over the interactions between individuals (Simmel and Wolff, 1950).

5.2.2 *The evolution of interaction level*

To understand if the same effect of the community size exists over the actions undertaken by the fans, we performed an additional analysis from the perspective of number of posts, comments and likes shared by the users. Similarly, when observed in relation to the total number of fans, the results showed that after an initial peak in March 2010, the level of interaction fell down and remained almost constant until the end of the observed period (see Figure 7).

Figure 7   Number of likes, comments and posts from fans in relation to total number of fans (see online version for colours)
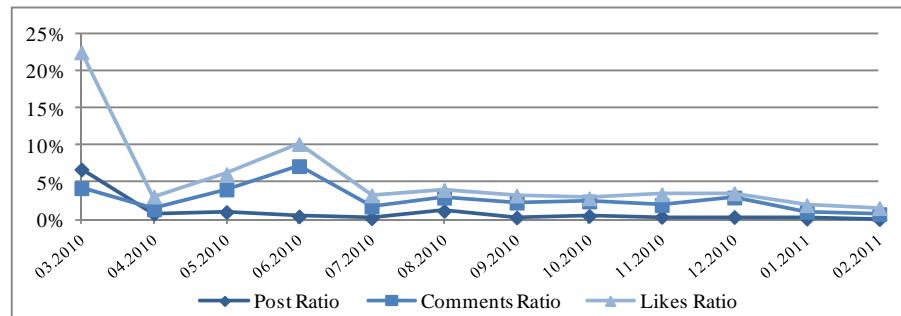

The likes' ratio showed the highest values throughout the observed period. This indicates that fans feel free to express themselves by performing the action of 'liking' which does not expose them to follow-up reactions by other users. On the contrary, the normalised numbers of posts show lowest values, which indicate that majority of fans don't feel free to share verbal opinions which might result in additional actions by other users or the moderator, such as comments.

Apart from the initial peak, an additional increase in the number of comments and likes can be observed starting from May 2010, with the maximum in June 2010. This period corresponds with two Facebook Ad campaigns organised by the company to increase the number of fans – first one on the week of April 15th and the second one in the week of May 25th (see Figure 8). These numbers and the increase in the number of fans clearly indicate that this action had positive outcome.

An additional measure for the interaction level can be obtained through the number of unique page views from the Facebook users (fans and non-fans). This measure can be obtained through a Facebook Insights platform which is provided to the page administrators in order to support them in high-level monitoring of the content and activities (Facebook Insights).

Figure 8 illustrates the obtained results in relation to the total number of fans.

Figure 8   Number of page views from Facebook users (fans and non-fans) in relation to the total number of fans (left axis) and the total number of fans (right axis) (see online version for colours)
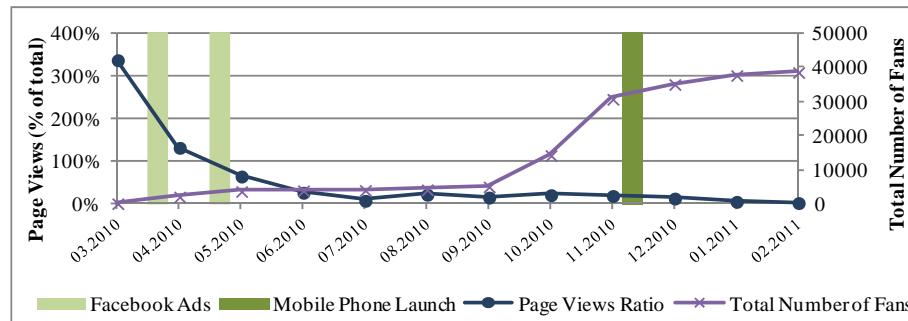
It can be seen that this measure also exhibits decrease over time. In addition, while at the beginning there were many page views by non-fans, as the brand page reached certain level of maturity, this number fell.

While the interaction of the active fans can be monitored, activities of *lurkers* remain unknown. Since page posts appear on fans' wall, *lurkers* don't need to access the brand page in order to consume the content, thus leaving no measurable trace of their activities.

Figure 9   Betweenness, degree centrality and group density of the commenting network, (a) with the moderator (b) without the moderator (see online version for colours)

	(a)		(b)
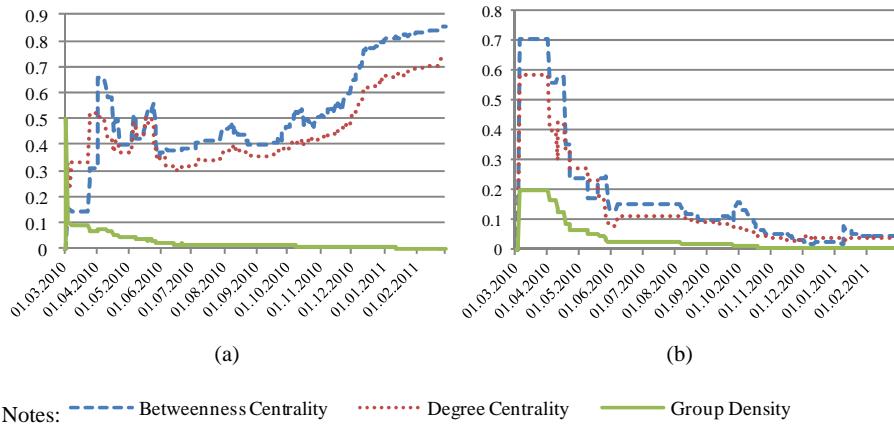
Notes: − − − Betweenness Centrality   · · · · · · Degree Centrality   ——— Group Density

5.2.3 *Structural characteristics of the interaction network*

A final overview over the interactions between the fans was obtained through structural analysis of the interaction network. To create the interaction network we have used fans as network nodes and commenting and liking activities as network ties. Posting activity was not taken in consideration since in the format provided by the graph API all posts shared by the users are addressed to the moderator, thus the resulting network has a perfect star shape and as such provides no insights.

To observe if the similar effect of the network size exists over the structural characteristics of the interaction network, we have applied DSNA over the selected period. Furthermore, since liking and commenting activities previously exhibited different characteristics, we have divided these activities into separate networks to see if they show different evolution over time. The results of the DSNA of the commenting network are illustrated on Figure 9.

Figure 9(a) shows that the group density exhibits initial peak in March 2010, then falls down and continues into a relatively stable phase. On the contrary, both centrality measures follow similar patterns of increase over time. This shows that some fans are more active than others. Since this network includes both, fans and the moderator, we have assumed that the obtained results are biased by the moderator. For that reason, we have filtered out the comments from and to the moderator. The obtained results are shown on Figure 9(b).

It becomes clear that commenting interaction between the users decreases over time. After the initial peak, both interaction measures decrease until June 2010, followed by a relatively stable period until October 2010, where an additional decrement can be seen before entering into the stable behaviour. Compared to the total number of fans illustrated on Figure 8, these fluctuations are exactly the opposite, confirming again the negative effect of the network size over the interactions between the users.

In terms of liking, the characteristics of the interaction network are illustrated on Figure 10.

Figure 10 Betweenness, degree centrality and group density of the liking network, (a) with the moderator (b) without the moderator (see online version for colours)

(a)　　　　　　　　　　　　　(b)

Notes: ---- Betweenness Centrality ······· Degree Centrality —— Group Density
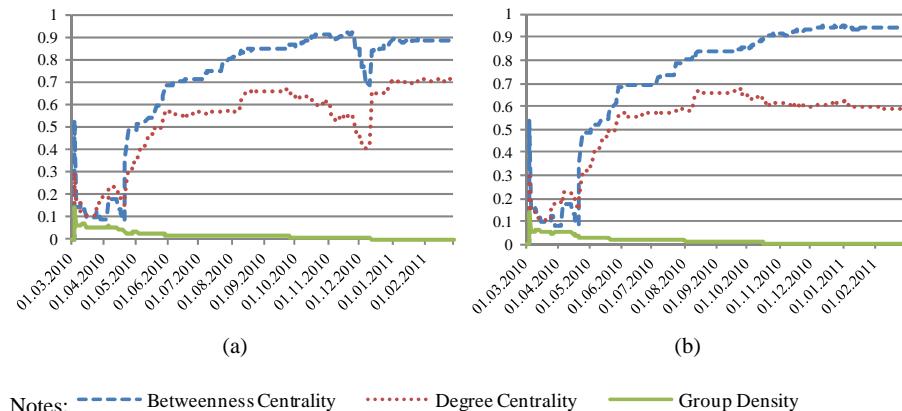
While the graph density presented on Figure 10(a) shows similar behaviour, the distribution of centrality measures differs. On overall level, an increase is present again, with a negative peak in December 2010. This period corresponds with the period when a mobile phone line was launched which had some problems at the beginning, causing the flow of *criticism and complaints* posts as already mentioned in the content analysis section (see Figure 8). After the problems were solved, the communication returned to its 'normal' form.

To compare if the same difference between the interaction dynamics can be observed when the moderator is removed from the network, we have performed the same filtering as before. The resulting network dynamics is shown on Figure 10(b).

It can be seen that in the case of liking, the behaviour of users resembles the behaviour of the network in whole. This means that while some fans frequently like the content shared by other fans, others only occasionally or never engage.

Further, if we look at the 'critical' period in December 2010, we can see that the same effect is not visible. This clearly indicates that the previously observed difference originates from the actions undertaken by the moderator, who was unlikely to 'like' the posts containing the complaints.

6   Implications for SMM

In order to successfully run a Facebook brand page as a part of the SMM approach, companies need to understand their users by learning how and why they interact on the brand pages. We have addressed this problem from two perspectives,

1   the content shared by the fans

2   the interactions between the fans.

In the continuation we present the implications of the obtained results.

Posts shared on a Facebook brand page represent a valuable source of knowledge for companies. The results of the topic analysis indicate that posts reveal:

1   perception of the brand

2   acceptance of a new product

3   most favoured products and features

4   required products and features

5   locations with great volume of sales. In addition, by listening to the page conversation companies could identify

6   existing problems

7   perceived competitors.

Further, with 27% of the posts belonging to the *requests and suggestions* category, Facebook brand pages provide the possibility for generation of ideas about new products and services.

Organisationally, the topics of conversation can be used to understand what different sources of information should be available to moderators to successfully run a Facebook brand page, i.e., members of the support board behind the moderator that can be addressed when a specific question is posed. Looking at the main topic groups as well as those referred within the Information Inquiry category, this study indicates the need for following experts:

1   sales

2   logistics

3   company/brand information (producer, founder, history, etc.)

4   product information

5   environmental issues.

Our study shows that Facebook is a suitable platform for SMM which supports the SMM goals for

1   building brand awareness

2   gathering insights and knowledge for future steps

3   community engagement in dialog.

This arguing is based on the results showing that *product*, *sales* and *brand* are the three most discussed topics, while *requests and suggestions*, *expressing affect* and *sharing* are the most common intentions for participation.

Furthermore, we show that the topics and categories are interconnected and propose the topic-category matrix, as a tool for practitioners that enables measurement of success of SMM utilisation over time. Our results show that on the ok.- brand page *product – requests and suggestion* (24%) and *expressing affect* towards the *products* (20%) were the most common combinations which are in line with the company's SMM goals.

In terms of user interaction, Facebook provides a possibility for user classification based on the activities available to the users. This classification is important for the companies in order to understand how their customers use SM, to develop appropriate communication strategy and prepare themselves for the possible public display of negativity. Continuous monitoring of the activities based on the proposed classification enables deeper understanding of the level of interactions and its evolution over time. Our results show that majority of the brand page users (98%) are passive users, i.e., *lurkers* and that this group grows with the growth of the community. An appropriate approach should be undertaken by the company to address different classes of users in order to lead them towards the higher level of interaction.

Further, we show that active users prefer not exposing themselves to possible reactions from other members of the community, thus embracing the 'safest' option for interaction, i.e., 'liking'. Still, previous findings showed that daily users exhibit significantly more interest in brand profiles (Li, 2007), and that triggering the user interaction could result in optimisation of the marketing investment (Sterne, 2010). Therefore, improving the level of user's activity is a worthy goal for the companies. This could be achieved through a stronger moderation, e.g., encouraging posters and preventing or discouraging aggressive posts and comments.

The proposed approach also provides the possibility to identify influencers, i.e., the 'superfans'. Companies should devise a plan to address them directly, since these are the "customers who are so positive about a brand that they do much of its marketing and sales themselves – and for free" [Harris and Rae, (2009), p.31].

Finally, the DSNA has shown that Facebook brand pages exhibit similar interaction evolution as already recognised in the sociology research, i.e., as the community grows, the percentage of active users reduces and the interaction becomes uniform. Furthermore, the moderator becomes a central figure in the network which influences the overall network structure. Since by definition a community assumes existence of a group of people interacting together, our results show that the brand communities on Facebook do not comply with the expected behaviour. While users interact with the moderator, the interaction between users themselves is relatively limited. This is an indication for the company that the increase of the size of the community requires more active moderation in order to encourage creation of relations between the fans. This might be achieved by organising activities such as competitions, polls, discussion threads, etc.

7   Conclusions and future work

This study presents an analysis of the UGC and the interactions between the users on a Facebook brand page. We have identified the topics and categories of posts and have explained their interconnections. Further, we have proposed a user classification scheme based on the interaction patterns and have analysed the interactions in relation to the community size. Based on the obtained results we have drawn implications SMM.

We are aware of the limitations of our study in the sense of a relatively small dataset extracted from only one Facebook page. Still, we believe that the long period of time for data collection offers more insights. In terms of the content analysis we were able to observe greater variety of addressed topics, since certain subtopics occurred only within a limited number of posts (e.g., *environment*) and during a limited period of time (e.g., the subtopic *technical details* occurs for the first time on December 4th, 2010). Furthermore, we provide a detailed analysis for the specific domain of a Facebook brand page, managed by the company offering FMCG. Finally, the longitudinal approach has provided us with the possibility to observe the effect of community size over the interaction level and characteristics.

The analysis presented in this paper can be used by marketing practitioners as a measure for successful SMM utilisation over time. Automation of the proposed methods would provide the possibility for real-time monitoring and timely reaction. The next steps would comprise integration of the script used for data collection with Condor or some other SNA tool. In addition, the manual process of content analysis will be replaced with a system that performs automatic topic classification and sentiment analysis. Further optimisation of the existing opinion mining techniques, addressing in particular the specific language structure and multilingualism would contribute in this direction.

In our future work we would like to compare the obtained results with those of other brand pages and over different categories of Facebook pages, e.g., celebrities' pages.

References

Agozzino, A.L. (2010) *Millennial Students Relationship with 2008 Top 10 Social Media Brands via Social Media Tools*, Unpublished PhD thesis, Bowling Green State University, Ohio, USA.

Asur, S. and Huberman, B.A. (2010) 'Predicting the future with social media', *IEEE/WIC/ACM'10: Proceedings of International Conference on Web Intelligence and Intelligent Agent Technology, IEEE*, pp.492–499.

Bolotaeva, V. and Cata, T. (2010) 'Marketing opportunities with social networks', *Journal of Internet Social Networking and Virtual Communities* [online] http://www.ibimapublishing. com/journals/JISNVC/2009/109111.pdf (accessed 17 November 2011).

Boyd, D.M. and Ellison, N.B. (2007) 'Social network sites: definition, history, and scholarship', *Journal of Computer-Mediated Communication*, Vol. 13, No. 1 [online] http://jcmc.indiana.edu/vol13/issue1/boyd.ellison.html (accessed 17 November 2011).

Brandt, K.S. (2008) 'You should be on YouTube', *ABA Bank Marketing*, Vol. 40, No. 6 [online] http://findarticles.com/p/articles/mi_hb4838/is_6_40/ai_n29454087/ (accessed 17 November 2011).

Brown, J., Broderick, A.J. and Lee, N. (2007) 'Word of mouth communication within online communities: conceptualizing the online social network', *Journal of Interactive Marketing*, Vol. 21, No. 3 [online] http://onlinelibrary.wiley.com/doi/10.1002/dir.20082/pdf (accessed 17 November 2011).

Cassidy, J. (2006) *Me Media: How Hanging Out on the Internet Became Big Business*, The New Yorker [online] http://www.newyorker.com/archive/2006/05/15/060515fa_fact_cassidy (accessed 17 November 2011).

Clemons, E.K., Barnett, S. and Appadurai, A. (2007) 'The future of advertising and the value of social network websites: some preliminary examinations', *ICEC'07: Proceedings of 9th International Conference on Electronic Commerce*, pp.267–276, ACM, New York, NY, USA.

Condor [online] http://wiki.soberit.hut.fi/virtualbrownbag/tiki-index.php?page=GettingStartedWithCondor (accessed 17 November 2011).

Coon, M. (2010) *Social Media Marketing: Successful Case Studies of Businesses Using Facebook and YouTube with an In-Depth Look into the Business Use of Twitter*, Unpublished M.A. Thesis, Stanford University, Stanford, CA.

Das, S.R. and Chen, M.Y. (2007) 'Yahoo! for Amazon: sentiment extraction from small talk on the web', *Management Science*, Vol. 53, No. 9, pp.1375–1388.

Dave, K., Lawrence, S. and Pennock, D.M. (2003) 'Mining the peanut gallery: opinion extraction and semantic classification of product reviews', *WWW '03: Proceedings of 12th International Conference on World Wide Web*, pp.519–528, ACM, New York, NY, USA.

De Choudhury, M., Sundaram, H., John, A. and Seligmann, D.D. (2008) 'Can blog communication dynamics be correlated with stock market activity?', *HT'08: Proceedings of 19th Conference on Hypertext and Hypermedia*, pp.55–60, ACM, New York, NY, USA.

Facebook Graph API [online] http://developers.facebook.com/docs/reference/api/ (accessed 17 November 2011).

Facebook Insights [online] http://developers.facebook.com/docs/reference/api/ (accessed 17 November 2011).

Facebook Statistics [online] http://www.facebook.com/press/info.php?statistics (accessed 17 November 2011).

Freeman, L.C. (1979) 'Centrality in social networks: conceptual clarification', *Social Networks*, Vol. 1, pp.215–239.

Glaser, B. and Strauss, A. (1967) *The Discovery of Grounded Theory: Strategies for Qualitative Research*, Aldine, Chicago.

Gloor, P., Laubacher, R., Zhao, Y. and Dynes, S. (2004) 'Temporal visualization and analysis of social networks', *NAACSOS'04: North American Association for Computational Social and Organizational Science Conference* [online] http://www.ickn.org/documents/CKN4NAACSOS.pdf (accessed 17 November 2011).

Gordhamer, S. (2009) '4 ways social media is changing business', *Mashable Business* [online] http://mashable.com/2009/09/22/social-media-business (accessed 17 November 2011).

Harris, L. and Rae, A. (2009) 'Social networks: the future of marketing for small business', *Journal of Business Strategy*, Vol. 30, No. 5, pp.24–31.

Jansen, B.J., Zhang, M., Sobel, K. and Chowdury, A. (2009) 'Twitter power: tweets as electronic word of mouth', *Journal of American Society for Information Science and Technology*, Vol. 60, No. 11, pp.2169–2188.

Javitch, D.G. (2008) *Entrepreneurs Need Social Networking' Entrepreneur* [online] http://www.entrepreneur.com/humanresources/employeemanagementcolumnistdavidjavitch/article198178.html (accessed 17 November 2011).

Joinson, A.N. (2008) 'Looking at, looking up or keeping up with people?: Motives and use of Facebook', *CHI'08: Proceeding of 26th Annual SIGCHI Conference on Human Factors in Computing Systems*, pp.1027–1036, ACM New York, NY, USA.

Kozinets, R.V., de Valck, K., Wojnicki, A.C. and Wilner, S.J.S. (2010) 'Networked narratives: understanding word-of-mouth marketing in online communities', *Journal of Marketing*, Vol. 50, No. 3, pp.140–189.

Lampe, C., Ellison, N. and Steinfield, C. (2006) 'A Face(book) in the crowd: social searching vs. social browsing', *CSCW'06: Proceedings of 20th Conference on Computer Supported Cooperative Work*, pp.167–170, ACM, New York, NY, USA.

Li, C. (2007) *How Consumers Use Social Networks*, Forrester Research Paper [online] http://www.eranium.at/blog/upload/consumers_socialmedia.pdf (accessed 17 November 2011).

Liu, Y., Huang, X., An, A. and Yu, X. (2007) 'ARSA: a sentiment-aware model for predicting sales performance using blogs', *SIGIR'07: Proceedings of 30th International Conference on Research and Development in Information Retrieval*, pp.607–614, ACM, New York, NY, USA.

Nasukawa, T. and Yi, J. (2003) 'Sentiment analysis: capturing favourability using natural language processing', *K-CAP'03: Proceedings of 2nd International Conference on Knowledge Capture*, pp.70–77, ACM, New York, NY, USA.

Nazir, A., Raza, S. and Chuah, C.N. (2008) 'Unveiling Facebook: a measurement study of social network based applications', *IMC'08: Proceedings of 8th SIGCOMM Conference on Internet Measurement*, pp.43–56, ACM New York, NY, USA.

Nonnecke, B. and Preece, J. (2000) 'Lurker demographics: counting the silent', *CHI'00: Proceedings of SIGCHI Conference on Human Factors in Computing Systems*, pp.73–80, ACM, New York, NY, USA.

Palmer, A. and Koenig-Lewis, N. (2009) 'An experiential, social network-based approach to direct marketing', *Direct Marketing: An International Journal*, Vol. 3, No. 3 [online] http://www.emeraldinsight.com/journals.htm?issn=1750-5933&volume=3&issue=3&articleid=1806140&show=html (accessed 17 November 2011).

Richter, D., Riemer, K. and Von Brocke, J. (2011) 'Internet social networking: research state of the art and implications for enterprise 2.0', *Business & Information Systems Engineering*, Vol. 3, No. 2, pp.89–101.

Simm, W., Ferrario, M-A., Piao, S., Whittle, J. and Rayson, P. (2010) 'Classification of short text comments by sentiment and actionability for voice your view', *SocialCom'10: Proceedings of International Conference on Social Computing, IEEE*, pp.552–557.

Simmel, G. and Wolff, K. (1950) *The Sociology of Georg Simmel*, Free Press, Glencoe, IL.

Sterne, J. (2010) *Social Media Metrics. How to Measure and Optimize your Marketing Investment*, Wiley, Hoboken, NJ, USA.

Wasserman, S. and Faust, K. (1994) *Social Network Analysis: Methods and Applications*, Cambridge University Press, Cambridge.

Weston, R. (2008) *'7 Social Networking Strategies' Entrepreneur* [online] http://www.entrepreneur.com/technology/bmighty/article191312.html (accessed 17 November 2011).

Zhang, M. and Jansen, B.J. (2008) 'Using action-object pairs as a conceptual framework for transaction log analysis', in Jansen, B.J., Spink, A. and Taksa, I. (Eds.): *Handbook of Research on Web Log Analysis, Information Science Reference*, pp.414–433, Hershey, New York, NY, USA.