

[Home](#)[Table of Contents](#)[Titles & Subject Index](#)[Authors Index](#)

Content marketing through data mining on Facebook social network

Saman Forouzandeh

Department of Computer Engineering, Kurdistan Science and Research Branch, Islamic Azad University, Sanandaj, Iran. E-mail: saman.forouzandeh (at) gmail.com

Heirsh Soltanpanah

PhD, Department of Industrial Engineering, Sanandaj Branch, Islamic Azad University, Sanandaj, Iran. E-mail: heresh (at) iausdj.ac.ir

Amir Sheikhhahmadi

PhD, Department of Computer Engineering, Sanandaj Branch, Islamic Azad University, Sanandaj, Iran. E-mail: amir.sheikhhahmadi (at) gmail.com

Received January 20, 2014; Accepted June 25, 2014

Abstract

As the Internet has found its place within different sectors of society, the use of various social networks is increasing and that is why they are utilized for business purposes. Most commercial companies need to do marketing in social networks in order to introduce their goods and/or services. In marketing, appropriate and accurate advertisement of goods based on needs of users of these social networks is required to sell the goods of different companies and manufacturing units in effective ways. In most cases, however, advertising and marketing methods in social networks are not correct and the resulted advertisements are boring and tiring to users and they might ignore them without noticing. Such advertisements will be considered as spam and will be annoying to network users. This is taken into consideration in the present study where users' interests, attitudes, and behavior on Facebook are specified through data mining techniques, based on which content marketing is conducted. The present study is conducted on the social network of Facebook, where content marketing, a new form of marketing, is utilized and instead of introducing the goods, the contents of different goods are presented. The data utilized in the study are actual and related to Facebook users, and therefore, the results can be generalized to other networks.

Keywords

Social networks; Marketing; Advertisement; Data mining; Decision tree

Introduction

Online marketing can fail due to several different reasons. One of the main reasons for such failure is that users remain unaware of the advertised goods and their advantages. Another considerable issue that causes marketing failure is the lack of information about users and their tastes; that is, they are offered to purchase goods without analyzing them and their behavior in the network. In this regard, competition among different manufacturers and service companies makes the development of new kinds of electronic marketing a necessity. One of these new types of marketing is marketing in social networks. An online social network is a community on the Internet where individuals usually keep in touch through a profile that introduces them (Donath & Boyd, 2004).

Individuals' communication with each other is also significant in marketing on social networks and these individuals can be used as marketers so that they can distribute the information. Content marketing is defined as a marketing process of creating and properly distributing the content in order to attract, make communication with, and understand other people so that they can be motivated to do beneficial activities (Pulizzi & Yoegel, 2012). Pulizzi and Barrett (2008) states that, "*The reason for choosing content marketing is that it seems to be an understandable term in professional marketing. Content marketing is a combination of content and marketing in which the content shows the customers' behavior.*"

Facebook is a social network that is increasingly developing and at least 100 million users visit it every day (Karimzadehgan et al., 2009). That is why the present study is conducted on Facebook. In 2002, Facebook officials claimed that it had one billion users per month (Rieder, 2013). Statistics have showed that the number of Facebook users had a remarkable increase between 2004 through 2011 (Zarella & Zarella, 2011). In order to develop content marketing, Facebook users' behavior needed to be analyzed. In so doing, Netvizz application is utilized to construct the friendship graph that is analyzed through Gephi software. Then, users' behavior is observed through data mining techniques, and based on their interests, content presentation is privately accomplished.

Literature Review

There is no doubt that websites of social networks have been special events in the field of the Web and Internet marketing. Publication of large number of articles on marketing through social networks in scientific journals in recent years is an evidence for such claim. In this section, some of these articles are reviewed.

Content marketing is professionally being studied by Content Marketing Association (CMA). Different scholars are also introducing and teaching it through their books and research studies.

Content marketing has recently been paid prior attention in social networks. It has been introduced as an appropriate marketing trend. Some authors have also written several how-to books and proposed methods of conducting such marketing. For instance, Sonia Simon (2014) in her book, "*Content Marketing*" has introduced content marketing and its different methods. CMA has trained content marketing in social networks and published its results according to conducted surveys and its comparison with other marketing methods.

There are also different studies on content marketing in social networks. For example, Saravanakumar and SuganthaLakshmi (2012) investigated marketing in social networks. They discussed the advantages of Internet marketing in these networks and argued that social networks have more users compared to other media. They also mentioned that the social networks of Facebook and Twitter have been introduced as the most popular ones. SMM (Social Media Marketing) is a method to promote a website, a brand, or a business through creating communication with customer in order to absorb their current benefits and predict their future needs through social networks. Hartline et al. (2008) investigated various strategies of marketing and discussed smart sale and viral marketing. It is stated that a customer purchases an item under the impact made by other customers. This strategy is called viral marketing and it is based on finding influential individuals in social networks, absorbing them, and conducting marketing through them.

Jothi et al. (2011) discussed communication and advertisement in social networks. It investigates the effectiveness of using brand in advertisement available in communications of social networks. Qiao (2008) studied electronic commerce (e-commerce) in social networks and its positive and negative aspects. Qiao stated that social networks could change and revolutionize e-commerce and its limitations can be overcome to a high extent through social networks. What is commonly argued in all of these studies is that marketing in social networks is based on communication among the users of these networks. That is, the broader and stronger the communication among the users is, the higher the level of sale and service provision will be. Another significant issue in social networks is level of understanding the users. The higher the level of understanding, the higher the chance of selling goods will be, because knowing the users according to their behavior and needs causes better prediction about what services should be provided next.

Proposed Method

In the present study, content marketing is proposed as the method for marketing in social networks. In the first phase, information on Facebook users' behavior is studied and a number of users are selected based on their behavior in the past and effectiveness of their posts among other users in order to distribute the information through. Then without introducing a special brand, the target contents are presented to other users through them. In the second phase, by utilizing data

mining techniques, the users' individual and collective behavior is analyzed according to the likes given on posts, based on which the users are provided with special contents.

Method Introduction

To conduct the current study, a Facebook account is created, and to analyze the users' behavior, Netvizz application is utilized. This application enables researchers to extract needed data from different services of Facebook and save or view the output files in different formats (Rieder, 2013). Netvizz application elicits fields from users as presented in Table 1 below.

Table 1. Fields analyzed by Netvizz application

	Elicits fields	Fields operation
1	Sex	user specified sex
2	Locale	user selected interface language
3	Agerank	accounts ranked by creation date where 1 is youngest
4	Like_Count	number of user likes
5	Post_Count	number of user posts
6	Post_Like_Count	number of likes on user's posts
7	Post_Commet_Count	number of comments on user's posts
8	Post_Enguagement_Count	post_comment_count + post_like_count

In the present study, fields of 5 and 8 are taken into consideration and it is aimed at finding users who have sent a lot of posts and whose posts have received a lot of comments and likes. After the users' data are extracted, the output file is launched in Gephi software and the graph constructed by the users is observed and analyzed. Gephi (www.gephi.org) is an open-source software that is utilized for analyzing networks and graphs. Its flexible and multipurpose architecture provides new features for dealing with complicated data and creating valuable image outcomes (Bastian et al., 2009). After the influential individuals are identified, a business page is created and these individuals are used to distribute information and the presented contents are provided to them and because of their behavior of having a lot of shares they are asked to distribute the information among other users. Through friendship networks, this method of marketing caused the contents to be shared among a large number of Facebook users and the users to become customers. In the next phase, the provided content is analyzed through data mining techniques and individuals' interest level in the contents is measured whereby it is specified which content is more liked by the users.

Data Preparation

In this section, data consisting of Facebook users' extracted information are presented. Intended users are found and target posts are sent to them and they provided other users with the posts by sharing them, this is how content marketing took place. In this study, a Facebook account that had 889 friends is studied because individuals who are highly active and influential among other users are selected to distribute the contents. Figure 1 shows the file gained from Facebook users using Netvizz application in Gephi software.

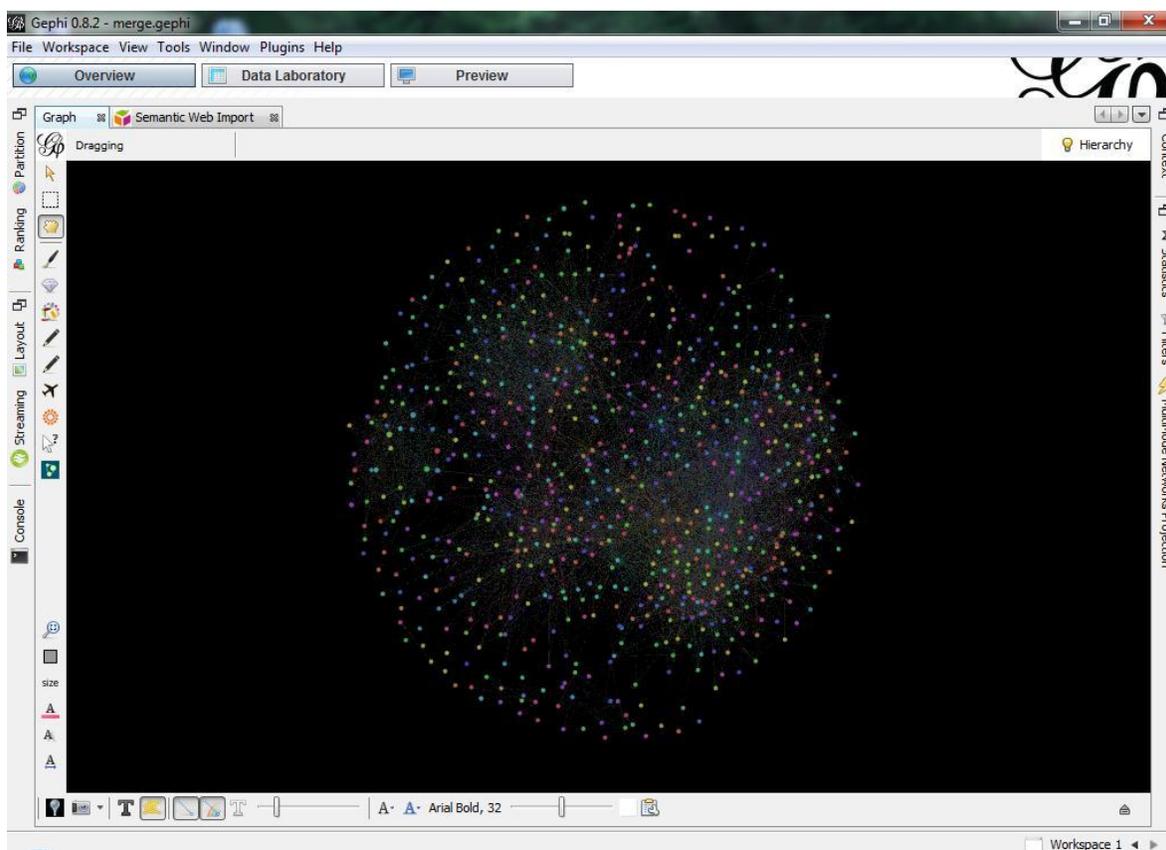


Figure 1. Graph display in Gephi software based on a specific field related to the users' performance

Gradation of nodes is performed in such a way that a node that has most mutual friend with the studied account received the highest grade in this graph. The graph can be filtered and showed based on different fields. Selecting users is based on related fields so that they can be utilized as marketer and their influence in network is used as a criterion to provide them with different contents. This way other users welcomed the content provided by those influential users, which would raise the sales and also its variety in the social network of Facebook. It is notable that in the current study it is supposed that individuals who make comments or give likes have read the content and then decided to send their comments or likes. The output data are saved in Excel format. In order to enhance the individuals' choice accuracy, a program is designed and the

Excel database achieved from them is analyzed and individuals who are within fields 5 and 8 are chosen. Figure 2 illustrates the output for the abovementioned individuals.

Row	Lable	sex	agerank	like_count	post_count	post_like_count
1	Sor	male	139	1000	419	4996
2	Kata	female	94	1000	410	32001
3	Ham	male	41	1000	381	743
4	Haj	male	133	1000	369	1218
5	Kay	male	137	1000	351	5707
6	Sam	male	31	1000	47	2565
7	Sol	male	138	1000	124	357
8	Xab	male	130	1000	377	7284
9	Hed	male	71	927	332	7526
10	Far	male	142	844	421	5261
11	Dav	male	101	672	339	1693
12	Hazh	male	134	657	398	5517
13	Kho	male	108	650	138	686
14	Res	male	119	610	94	1699
15	Pey	male	121	597	293	629
16	Moha	male	122	578	471	1344
17	Xos	male	126	558	299	8490
18	Sad	male	34	528	472	8558
19	Ras	male	199	500	446	813
20	Aya	male	356	500	462	5150
21	Ars	male	114	500	439	1065
22	Dar	male	352	500	470	9196
23	Agi	male	241	500	431	6869
24	Vor	male	242	500	490	1392
25	Sir	male	289	500	449	688
26	Sha	female	336	500	33	99

Row	Lable	post_count	influence
1	Sha	4	796
2	a3	67	358
3	Haz	1	265
4	Per	78	235
5	Eln	82	203
6	Neg	34	200
7	Khos	13	198
8	a15	46	193
9	Sor	1	180
10	a4	197	178
11	Raf	147	175
12	Ley	170	174
13	Moh	2	158
14	Amir	3	138
15	Meñ	30	132
16	a8	254	131
17	Son	19	128
18	Arez	12	128
19	San	206	124
20	Hass	47	122
21	Han	288	120
22	Osm	77	117
23	Sain	75	115
24	Saha	23	113
25	Sara	131	113
26	Meh	1	112
27	Sae	35	108

Figure 2. Individuals who received the highest number of posts, comments, and likes for each post

Individuals with these characteristics are showed in the right part of the figure above. Measuring the rate of influence per post sent by the users is important because it shows the ration between sent posts and given likes/comments. According to the analysis conducted on the Facebook users' behavior and results of Netvizz application, users can be categorized into three different groups:

1. Users who like to give different likes. These users usually have fewer likes compared to other uses and prefer that other users make comments about their posts (Ricci et al., 2011).
2. Users who have few posts but give a lot of likes.
3. Users who are popular, reliable, and influential among other users.

In this study, users of group 1 and 3 are selected. Based on the type of Facebook users' behavior, their performance in the future can be predicted whereby they can be provided with services and their behavior can be utilized to carry out the marketing task.

Content Marketing through Data Mining Techniques

In order to provide the users with the contents and establish a sense of trust through content marketing, data mining techniques are utilized. The contents to be provided to the users are classified and for each group, decision trees are drawn and utilized based on the type of content and their characteristics. The user collects information about his/her favorite topics and gains the required knowledge and awareness. In order to classify the presented materials, a software is designed and the content is assigned into 6 groups (A-F) and a decision tree is drawn for each group. In research done by Forouzandeh et al. (2012), a similar method is utilized to predict success/failure percentage of software projects. After the contents are presented and distributed by the individuals who have been chosen in previous phase, the number of likes received for each post is inserted, whereby number of likes given to each group is specified. Figure 3 that is drawn through the designed software, number of likes given to each group related to the contents is illustrated.

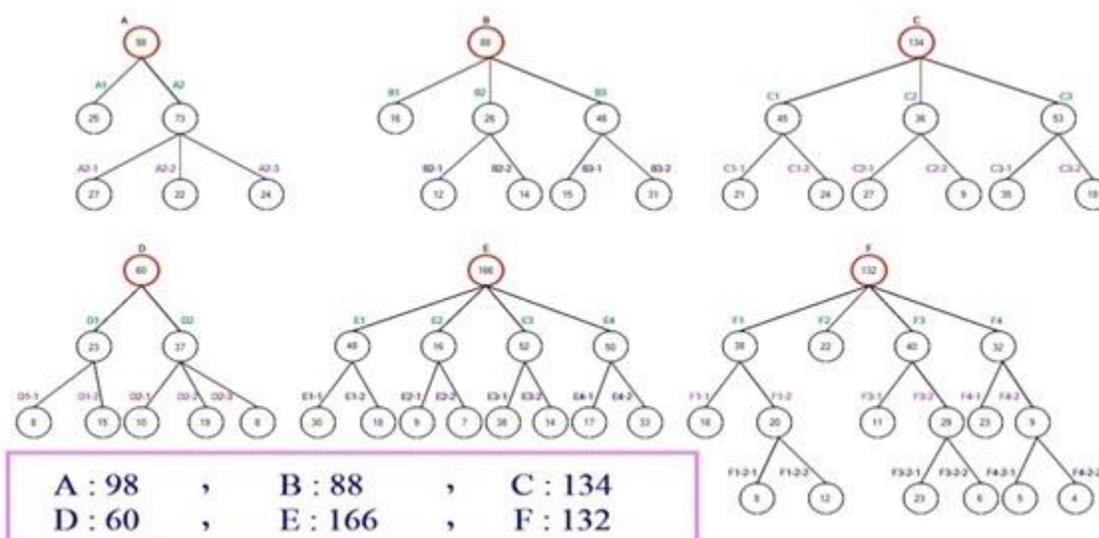


Figure 3. Classification of the presented contents based on decision trees and calculation of each groups' likes

As seen above, number of likes received for each group is specified and based on the output it can be determined that on which content higher investment should be made and which one is preferred more by the users. After the contents are presented to the users and content marketing is carried out and also likes for each group is specified, the users' behavior in the social network

of Facebook and their given likes are analyzed. In Figure 4, one of the users' likes is recorded and their path is tracked in decision trees.

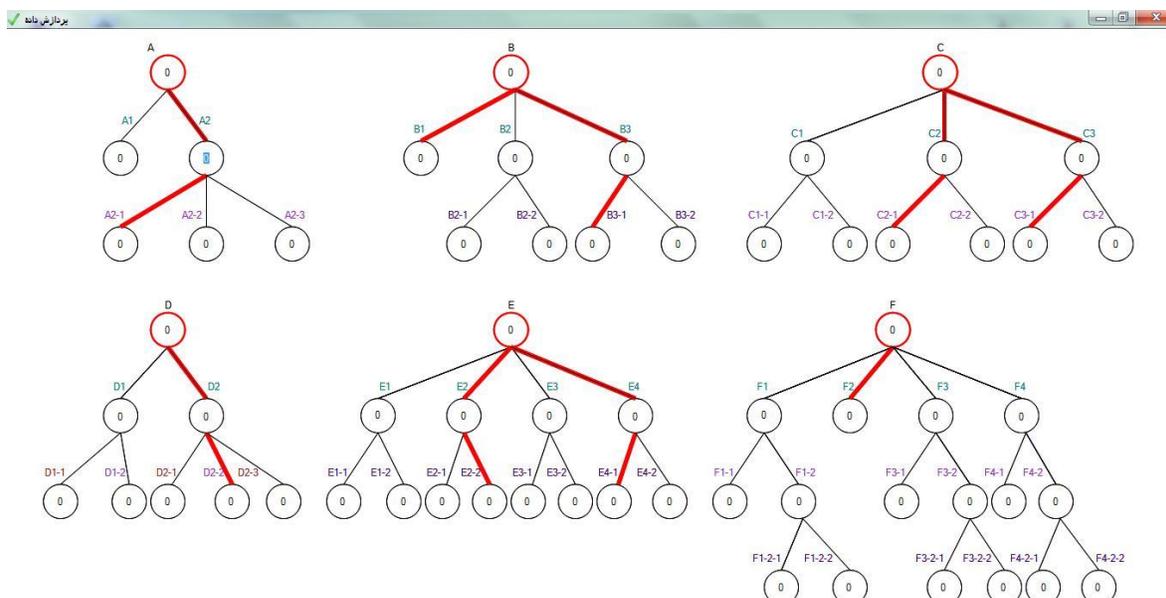


Figure 4. Illustration of the user's likes through decision trees on Facebook

As presented in Fig 4 above, the path of the user's likes is specified in each group and it can be simply determined in which content the user is interested and about which one he/she acquired required knowledge. Therefore, through the proposed method the users' both general and personal interests can be identified and their preferences can be specified. And in next phase, he/she can be suggested to purchase items that are related to the content liked by him/her because he/she already has needed knowledge about the item and the reason for using it.

Evaluation of the Proposed Method

Two aspects can be taken into consideration in evaluation of the current study. First, the results of the study can be compared with those of previous studies. Second, a survey can be conducted on the users' view about the proposed method. Content marketing can be combined with other types of marketing; and it can be utilized as a complementary method to other types of marketing. For instance in papers of (Hartline et al., 2008) and (Jianming & Wesley, 2010), viral marketing is utilized and influential individuals in the network are found, who receive the services for free. Here, it is supposed that other users will follow them and they will purchase accordingly. Likewise in this study, influential individuals are selected and investigated on. The purpose, however, is to gain their trust and distribute the contents among other users through them. Therefore, content marketing can be conducted in a viral way whereby strengths of both methods can be used. In article (Hartline et al., 2008). It is suggested that in order to absorb

influential individuals, they should be provided with services and goods for free and also users are not much aware of these services or goods. In content marketing; however, expenses are lower and customers have enough knowledge about the services and goods before purchase.

Since the present study is conducted on Facebook, a survey is conducted to evaluate the proposed method. The survey contained two questions each having three options about the content marketing. The survey is distributed among 500 Facebook users and following results are achieved.

In the first question, 415 users (83%) preferred that before introducing an item with a specific brand, required knowledge about goods and their purposes should be provided. Provision of knowledge and awareness before introducing the items is the basis of content marketing, whereby a sense of trust will form in users. The second question is about the users' purchase method, where 411 users (82%) preferred to know enough about an item and the reason for buying it before a specific brand is introduced to them. According to the results, utilizing this type of marketing is welcomed by the users and they prefer this method for their future purchases. In Figure 5, the users' method of responding to each option is presented. It should be noted that the second option of the question is about non-use of content marketing and the third option is about whether or not there is a significant difference between typical purchase and purchase through content marketing.

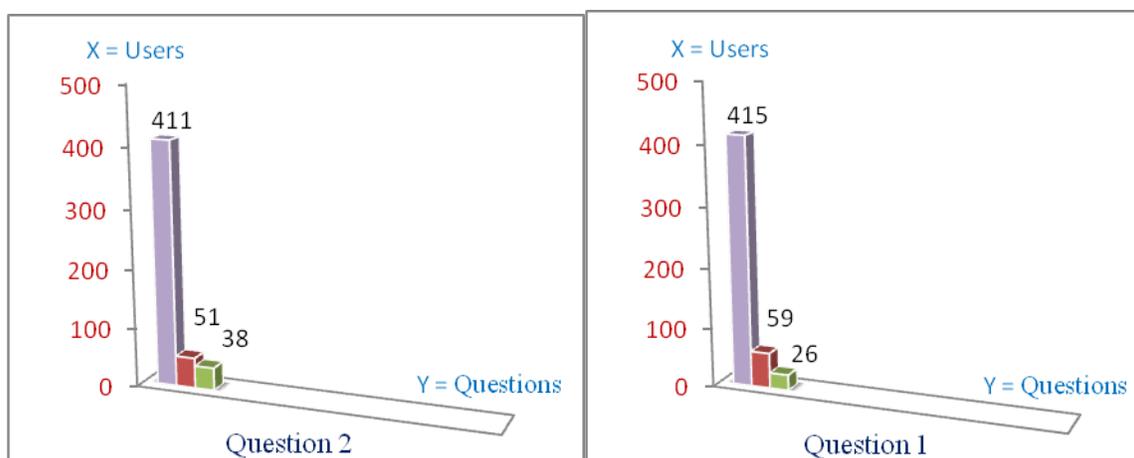


Figure 5. The results of the 2-question survey taken by the Facebook users

Conclusion

Compared to other types of marketing in social networks, content marketing has an important advantage. In this type of marketing, the users are not presented with a direct suggestion; instead they are involved with another issue and provided with information that is certainly useful and not considered commercial by them, which finally causes them to trust the content provider.

Establishing a sense of trust in its turn creates loyalty on the users' side. The social network of Facebook is a suitable basis for content marketing, where different contents can be easily presented to and rapidly distributed among the users, and it is in a quite agreement with the proposed method. Content marketing can be combined with other forms of marketing. According to the conducted evaluation and the survey taken by the users, the results show that users welcomed this type of marketing and they liked its performance. By developing and promoting knowledge and information among the users, content marketing creates a sense of trust in them to buy goods through the presented content. Moreover, this trust creation among the users raises goods sale. In fact, content marketing is the knowledge of absorbing customers indirectly. The results of the study prove this claim that content marketing is beneficial, whereby different goods can be introduced to the users.

Further Studies

According to the results of the study, a recommender system can be designed based on investigating the users' behavior and likes given to different contents, and their preferences can be compared with their profiles and then based on the similarity between their profiles and liked contents, they can be suggested to purchase related goods. Designing this recommender system is significant because users are provided with suggestions about which they have enough knowledge and know the reasons for buying a suggested item.

References

- Bastian, M., Heymann, S., & Jacomy, M. (2009). Gephi: An open source software for exploring and manipulating networks. *Proceedings of the Third International ICWSM Conference*, pp 361-362.
- Donath, J. & Boyd, D. (2004). Public displays of connection. *BT Technology Journal*, 22(4), 71-72.
- Forouzandeh, S., Sheikhhahmadi, A., & Forouzandeh, S. (2012). Using data mining techniques to determine the percentage of success and failure in software projects. *The sixth Iran Data Mining Conference (IDMC) – Tehran*, Article 143.
- Hartline, J., Mirrokni, V., & Sundararajan, M. (2008). Optimal marketing strategies over social networks. *Refereed Track: Internet Monetization - Recommendation & Security, ACM 2/08/2004*. Beijing, China.
- Jianming, H., & Wesley, W. (2010). *A social network- based recommender system (SNRS)*. Springer US, Vol. 12, pp. 47-74.
- Jothi, P.S., Neelamalar, M., & Prasad, R.S. (2011). Analysis of social networking sites: A study on effective communication strategy in developing brand communication. *Journal of Media and Communication Studies*, 3(7), 234-242.

- Karimzadehgan, M., Agrawal, M., & Zhai, C. (2009). Towards advertising on social networks. *Proceedings of the ACM SIGIR workshop on Information Retrieval and Advertising (IRA)*, Boston, MA, USA.
- Qiao, Y. (2008). Social networks and e-commerce. *Seminar on Internetworking*, TKK T-110.5190.
- Ricci, F., Rokach, L., Shapira, B., & Kantor, P.B. (2011). *Recommender Systems Handbook*. 1st Edition. New York: Springer.
- Rieder, B. (2013). Studying Facebook via data extraction: The Netvizz application. *WebSci '13 Proceedings of the 5th Annual ACM Web Science Conference* (pp. 346-355), New York: ACM.
- Saravanakumar, M., & SuganthaLakshmi, T. (2012). Social media marketing. *Life Science Journal*, 9(4), 4446-4451.
- Simon, S. (2014). *Content marketing: How to build an audience that builds your business*. Retrieved May 10, 2014, from <http://my.copyblogger.com/basic/content-marketing-building>
- Pulizzi, J., & Barret, N. (2008). *Get content, get customers: How to turn prospects into customers with content marketing*. Retrieved May 10, 2014, from http://getcontentgetcustomers.com/get_content_get_customers_ebookhtml/comment-page-1
- Pulizzi, J., & Yoegel, R. (2012). *Six useful content marketing definitions*. Retrieved May 10, 2014, from <http://contentmarketinginstitute.com/2012/06/content-marketing-definition>
- Zarrella, D., & Zarrella, A. (2011). *The Facebook Marketing Book*. Canada: O'Reilly Media, Inc.

Bibliographic information of this paper for citing:

- Forouzandeh, Saman, Soltanpanah, Heirsh, & Sheikahmadi, Amir (2014). "Content marketing through data mining on Facebook social network." *Webology*, 11(1), Article 118. Available at: <http://www.webology.org/2014/v11n1/a118.pdf>

Copyright © 2014, Saman Forouzandeh, Heirsh Soltanpanah, & Amir Sheikahmadi.