Marketing based on user behavior on Facebook social network through recommender system design

Bahareh Shams Zamenjani
Department of Computer Engineering & Information Technology
Amirkabir University of Technology
Tehran, Iran
B_shams_1364@yahoo.com

Abstract— **the influence of social networks among people and at the same time inevitable spread of commercial use of them. Accordingly, in order to sell products, recommender systems designed based on user behavior on social networks, providing a variety of commercial offers tailored to the user. The accuracy of recommender systems that make recommendations to users, and how many of the proposals are accepted by the users is important. In this paper, a recommender system is designed based on user behavior in social network Facebook in two acts and suggests that users purchase their favorite products. The first step is to examine user behavior based on user interests will be given an offer to buy products. In the second stage recommender system uses data mining techniques and suggestions to the user that is associated with their previous purchases. This is real data and the real results of it and it is valid, as well as the results show a high level of accuracy recommender system is designed to offer suggestions to users.**

Keywords: marketing, social network, profiles, recommender system, user modeling.

I. INTRODUCTION

Simultaneously with Internet penetration among different segments of society, the use of social networks has been expanded and this creates a variety of commercial uses of them. Different commercial enterprises in order to introduce their products or services need social media marketing.

A. Problem Definition

In marketing what is necessary to sell the products of companies and manufacturers is the appropriate and correct advertising of products based on the needs of users in social networks, but in most cases the advertising and marketing methods on social networks was not correct and are boring and tedious for users often without looking and paying attention to them act to remove them and actually this marketing becomes spam and may cause torture to users.

In this regard, competition among manufacturers and services firms also causes the need for a special type of e-marketing and social media marketing is one of the new types of marketing.

II. RELATED WORKS

An online social networking is a society on the Internet, where people interact with each other often through a profile represent and introduce them. People relations with others also are of importance in social media marketing and they can be used as marketers and to disseminate the information to others. Social media is one of the most popular parts of the virtual space for public use and diverse needs of users every day expand the scope of social media. Social media have various types, one of which is social networks. The social network is a social structure and is formed from nodes (which are generally individuals or organizations) that by one or several specific types of dependencies are connected to each other. Many types of edges can exist between nodes. Using people's social position and their influence among other members and the so-called their influence in relevant virtual community, can be used for commercial purposes or publication of material and ideas by them and also in some cases their position has been used in marketing a variety of products and services.

III. INNOVATION

This paper has been done on Facebook because Facebook is one of the fastest-growing social networks and more than 100 million users log on to Facebook at least once each day. In 2012, Facebook announced that it has 1 billion active users a month. According to the statistics show between 2004 and 2011, the number of Facebook users has been increased dramatically each year. In order to do marketing we require analysis of user behavior on Facebook and Netvizz application has been used to this and friendship graph has been obtained using this application, Gephi application has also been used to analyze graphs. Next, data mining techniques are used to track the users' behavior; delivering content is also done based on their preferences and in private. Generally, when people become members of social networking sites, they set up a profile of their information. The profile is a description of people’s interests and activities. When the profile is completed, people seek to other users who have common interests with each other or invite other members to join the network. Since users have an

inner desire to communicate with each other, this network provides a unique opportunity for targeted marketing to people. User profiles on Facebook includes gender, image, user interests, hobbies, place of education and academic degrees, affiliations and interests of working, occupation and place of work and place of residence. Another important topic of social network is the amount of our knowledge of target users, more knowledge can enhance the chance to sell the products to them, because using user recognition we can anticipate user's needs and behavior and thereby act to provide services to them. In this study, one account on Facebook has studied, Netvizz application has been run to investigate the behavior of users that using studied account establishes a friendship. This application allows researchers to extract their needed data from different services of Facebook and attempt to save and view the output data in a variety of standard formats. Gephi application can also be used to analyze and observe the obtained graph. Fields extracted by Netvizz application from users are as follows in Table 1:

TABLE I. FIEIDS EXTRACTED BY NETVIZZ APPLICATION

1	Sex	User's gender
2	Locale	User's selected language to use the interface of Facebook
3	AgeRank	Ranking of user account creation date and account youngest with No. 1
4	Like_Count	The number of user's like
5	Post_Count	The number of user submitted posts
6	Post_Like_Count	The number of likes on user submitted posts
7	Post_Comment_Count	The number of comments on user submitted posts
8	Post_Enguagement_Count	Total of sixth and seventh state

What is considered in this research is related to the fields 5 and 8, and aims to find users who have a large number of posts and that posts on their behalf received a large number of comments and likes. After extraction of user data, the output file is implemented in Gephi application and then the graph formed by users can be seen and investigate. The output of implemented application is in .gdf format. Using this marketing method through takes advantage of established friendships on Facebook, causes the content can be quickly provide for many users and users covered by marketing become costumer and be able to fine people whose behavior indicates that they release ads one network and those who thorough like or leave comments on them are affected by it and try and act to design a recommender system be able to fine their interests and recommendations offer in accordance with the user's taste.

IV. EVALUATION OF RESULTS

In this section user's extracted data on Facebook is displayed. They find desired users and send considered post to them, users by sharing these posts make available them to other users. We must consider that the content first be shared then this content receive like from users, each like is equivalent to the login of one new user with its own specifications and doing marketing operation based on its profile. In this research one account on Facebook has been studied, this account has 899 friends; the aim is to select people to publish the contents who are more active and have much influence among other users. Fig. 1, below, shows the graph related to studied account on Facebook using Gephi application.

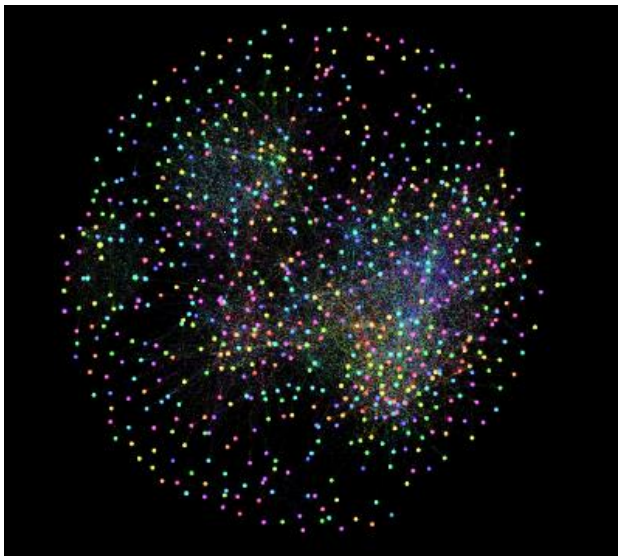

Fig. 1. Graph related to studied account on Facebook.

The graph drawn on figure above is related to people who have established friendship with the studied account. The ranking method of nodes is such that node with the highest grade in this account has most common friend. We can filter and displayed graphically based on different fields. For example, we consider the field No.8 from table 1 that is the sum of Likes and comments received for user-submitted posts and then the nodes will represented based on this field. Such that, each user with more like and received comments will have bigger node on graph and also it will has distinctive color than the rest of the other nodes and is easily visible. Fig. 2 displays a graph based on the above-mentioned fields

Users are selected based on relative field to use them as marketers, the amount of their influence on network has been used to introduce and provide different types of contents. When other users see the content provided will be made more welcoming, and this can also increase sales and its variety on the Facebook social network. It should be noted that in this study we assume that people who leave comments or like the content, they first read the relevant posts and then act to like or leave the comment. In order to increase the accuracy in selecting these individuals, we design an application and the considered operation has been done on obtained Excel

database of these users and then we attempt to extract users whose behavioral characteristics include both field 5 and 8.

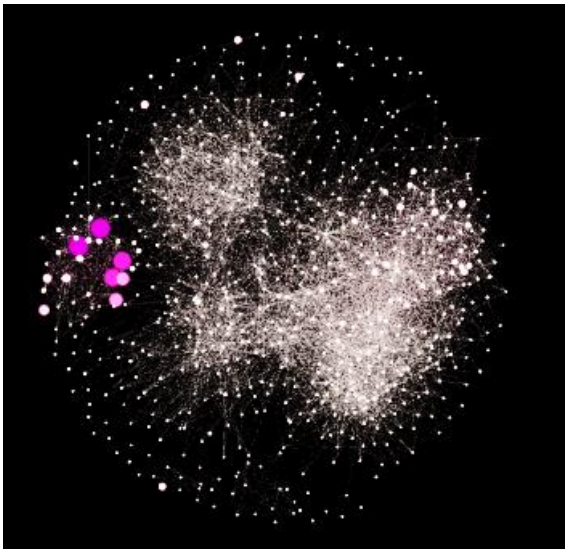

Fig. 2. Graph based on the specified fields.

In Fig. 3, the output obtained from users with the above conditions can be observed.

Row	Lable	sex	agerank	like_count	post_count	post_like_count
1	Soran Mobaraki	male	139	1000	419	4996
2	Katayoon Tabrizi	female	94	1000	410	32001
3	Hama Sharif	male	41	1000	381	743
4	Hajir Bokan	male	133	1000	369	1218
5	Kayvan Abbassi	male	137	1000	351	5707
6	Saman Baradari	male	31	1000	47	2565
7	Soleiman Mob...	male	138	1000	124	357
8	Xabat Foruzan...	male	130	1000	377	7284
9	Hedi Bukani	male	71	927	332	7526
10	Farid Esmaili	male	142	844	421	5261
11	Davoud Shafei	male	101	672	339	1693
12	Hazhir Rahmani	male	134	657	398	5517
13	Khosrow Agha...	male	108	650	138	686
14	Reshad Fateh	male	119	610	94	1699
15	Peyman Tavasoli	male	121	597	293	629
16	Mohammad Re...	male	122	578	471	1344
17	Xosro Mardanb...	male	126	558	299	8490
18	Sadegh Farha...	male	34	528	472	8558
19	Rasul Sheykh ...	male	199	500	446	813
20	Ayat Ruhani	male	356	500	462	5150
21	Arsalan Jamalz...	male	114	500	439	1065
22	Dara Ghoraishi	male	352	500	470	9196
23	Agiri Esmailnaz...	male	241	500	431	6869
24	Vorya Ranjbari	male	242	500	490	1392
25	Sirwan Bayazidi	male	289	500	449	688
26	Shawbo Shalavi	female	336	500	33	99

Row	Lable	post_count	influence
1	Sha Hed	4	796
2	a3	67	358
3	Hazhir Rahman...	1	265
4	Perwin Bokani	78	235
5	Elnaz Mokhtari	82	203
6	Negar Rahmani	34	200
7	Khosrow Mamle	13	198
8	a15	46	193
9	Soran Mahmo...	1	180
10	a4	197	178
11	Rafiq Hossen P...	147	175
12	Leyla Ghazi	170	174
13	Mohsen Zarei	2	158
14	Amir Sheikh A...	3	138
15	Mehran Ebrahimi	30	132
16	a8	254	131
17	Sonia Maroufi	19	128
18	Arezoo Moha...	12	128
19	Sanaz Shokry	206	124
20	Hassan Darzi	47	122
21	Hana Horame	288	120
22	Osman Khalili	77	117
23	Saina Mahabadi	75	115
24	Sahar Shekarchi	23	113
25	Sara Ahmadzadh	131	113
26	Mehran Beygz...	1	112
27	Saeid Najjari Aso	35	108

Fig. 3. List of users who have the characteristics of fields 5 and 8.

People who have the considered conditions can be seen in the right box (Fig. 3). Obtaining the amount of this influence to user-submitted posts is important because that this number indicate the ratio of user-submitted posts to receive like and comment. In studies on user's behavior on Facebook considering the obtained results by Netvizz application, users can be categorized into three different groups that include:

1- Users who are interested to send different posts, and these users receive usually low likes than others and they most prefer that other users leave comment on their submitted posts.

2- Users with low submitted posts but receive more likes.

3- Users who are more popular among others and so-called have high influence.

In this study, users who have the first and third group condition are considered. Considering user's behavior on Facebook we can predict their future action and on this basis act to provide services to them and take advantage of their behavior type of marketing. After selecting the desired users through making use of Gephi application, on the next step we attempt to design the relative recommender system. During this phase, we analyze users like using their profile. 5 fields have been considered from users profile on Facebook include: gender, marital status, age, education, place of residence. Five fields marked terms of products are considered for each of products that will be introduced to users. This means that each product will be adjusted with user profiles and products offered to the user should be most adapted to user profiles. Terms of products are based on user profile fields and matching is done accordingly. When referring to the fields the considered fields are shown. Comparison of user profile with these fields is peer to peer and the result of the comparison is shown numerically. Coding procedure of the fields is as shown in table 2:

TABLE II. USER PROFILE CODING BASED ON 0 AND 1

	Fields	Mode 1	Code	Mode 2	Code
1	Gender	Female	0	man	1
2	marital status	Single	0	Married	1
3	Age	Less than 25	0	More than 25	1
4	Education	Non- College education	0	College education	1
5	Living place	Unavailability of product	0	Access to product	1

About the fifth field that is considered as living place it should be said that this field coding is done with regard to desired product, such that if the product was near the user then it coded as zero (that is, offer to go to a concert or a restaurant that is near the user's living place) otherwise its code will be 1. Then differentiation matrix is used to compare user profiles with the terms of the product. This matrix determine distance or dissimilarity between two objects and is usually as n·n. d (i, j) is the considered amount to show differentiation and dissimilarity between objects i, j. distances are good criteria to similarity or dissimilarity between objects. To calculate the difference between the variables, their distance should be calculated, in this study Manhattan distance has been used is as follows:

$$d(i,j) = |Xi1 - Xj1| + |Xi2 - Xj2| + |Xi3 - Xj3| + \ldots + |Xip - Xjp| \quad (1)$$

Then we compare the obtained matrix. The obtained result is such that: more distance, the matrix are away from each

other and are no more similar to each other and if the value obtained is equal to zero, two matrices are equal. The created matrix is initialized with codes in table 1. 5 fields from user profile are also taking into consideration which is as single row matrix by 5 columns. The created matrix from user profile is compared to the matrix related to terms of product as field to field and their distance is obtained, if they were similar their distance will be zero and otherwise distance equal 1 will be considered for each field. And finally, a matrix compared to the user profile matrix has minimum distance that service or product is the best option for suggestions to users. For example, in the following chart we show the matrix corresponding to the five products, each of these products were compared to user profiles that were extracted in the previous step and determine which products is best match to user profiles and thus can offered to the user.

TABLE III. CODING OF PRODUCTS OFFERED TO USERS

	Living place	education	age	marital	gender
A	1	0	1	0	1
B	0	1	1	1	1
C	0	1	0	1	1
D	1	0	0	1	1
E	1	1	0	0	0

To do this, the extracted user profile which can be seen on figure 3, are coded in Excel format and then compare with each of products and their similarity is obtained. It should be said that the total number of users who have the characteristics of both fields 5 and 8 from Table 1 is 95. For example, in Fig. 4 can be seen for the first product.

	A	B	C	D	E
1	Sex	Marriage	Age	Education	Living Place
2	1	0	1	0	1
3	1	0	0	1	0
4	0	1	0	0	1
5	1	0	0	1	0
6	0	1	0	0	1
7	1	1	1	0	0
8	1	0	1	1	1
9	0	0	0	0	0
10	1	1	0	1	0
11	0	0	0	0	1
12	1	0	0	0	1
13	1	1	0	1	0
14	0	0	1	1	1
15	1	0	1	1	1
16	0	1	1	1	1
17	0	0	1	0	0
18	1	0	1	1	1
19	0	1	0	1	1
20	1	0	1	1	1
21	0	0	1	0	0
22	0	0	1	1	1
23	1	0	1	1	1
24	0	0	1	0	1
25	1	0	1	1	0

Sheet1

Ready

Fig. 4. Comparison of the user product coding with product's code.

As can be seen, the proposed product code marked in red font in the second row and the row will be compared with the user's profile on a field-by-field basis and this will be repeated for all 5 products. Then the result will provide as input file to MATLAB software and the results are displayed in a graph. The results for each of the products can be seen in Fig. 5.

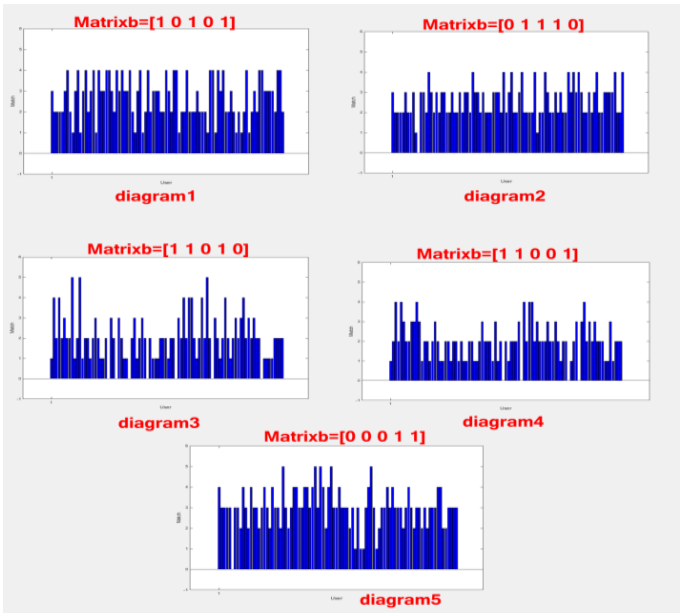

Fig. 5. Comparison of proposed product and user profile for Matlab software.

As can be seen in Fig. 5, each of the products has been compared separately with user profiles, the graph X includes members who are 95 users, and Y is number of compared fields which are 5 fields, on each number it indicates the user comparative field with proposed product. For example, if the chart was on the number 4 means that the product code and user profiles are similar in 4 fields and the user can be a good option to offer the product.

V. CONCLUSION

In this research during the first phase one account has been considered the behavior of users who have friendship with studied account was investigated, then users who have more posts and their posts receive high like and comments were specified. We assume that if we be aware of these user's interests, then considering their behavior type, they will release their favorite advertisements or products and put it on Facebook and will affect other users and by this way marketing is done by the users themselves. In the second phase, we should determine favorites and preferences of these users. To do this, user's profile, which are 95 users, is coded as 0 and 1 and is saved in Excel format. Then we coded the services or products that we want to offer to them such that we determine terms of use for them, means that these products are suitable for which users and this definition of terms of use is based on user profiles. Then using MATLAB application we compare products with users and identify users who have more similarity between their profile and product terms of use. Considering the conducted research we assume that using the proposed method in this research, we can specify the preferences of the users and product or

service provided is in accordance with their desire and thus they will purchase it, also considering their behavior in the past, they will share advertising related to favorite product or services and because that they have high influence and other users follow them and marketing is done by this way, the possibility of selling products or services to other users will increase.

REFERENCES

[1] Maver, John, Cappy Popp; Essential Facebook Development; Addision-Wesely: page 12; 2010.

[2] Y. A. Kim and J. Srivastava; Impact of social influence in e-commerce decision making; Technical report, ICEC 07, vol.258. ACM; New York 2007.

[3] Jiawei Han and Micheline Kamber; Data Mining: Concepts and Techniques; Second Edition, Diane Cerra; Elsevier Inc - 13: 978-1-55860-901-3; 2006.

[4] Lorence, Jessica; Word of Mouth Rules; Brand Strategy - pages 40 – 41; 2007.

[5] Nong Ye (Edited) ; The Handbook Of Data Mining ; Books published by Lawrence Erlbaum Associates are printed on acid-free paper - Printed in the United States of America ; 2003.

[6] Andrew W. Moore; Decision Trees; Professor School of Computer Science Carnegie Mellon University http://www.autonlab.org/tutorials/dtree18.pdf.

[7] B.K. Bharadwaj and S. Pal. ; Data Mining: A prediction for performance improvement using classification ; International Journal of Computer Science and Information Security (IJCSIS), Vol. 9,No.4,pp.136-140 ; 2011.