

A Survey on Analysis of user generated content on Facebook

Vaishali Nilesh Gavandi* and Umesh Kulkarni**

ABSTRACT

The social networks such as face book plays important role to collect the information of user perceptions. This information helps to bridge the gaps between service provider and user. The objective of this paper is to search the target users who are interested or sharing their likes and un likes on category such as Entertainment, Education, Travelling, Thinking etc. The service provider will collect this information of the target group and will help to improve their services. The insights of the face book user will be collected using decision tree learning algorithm with the help of ETL. The Standard 3-Tier Architecture is used which helps to handle the data not only meticulously but also be available in ready state for further usage. The entire API is developed using Python .To search the target user the category input is provided and the users who were interested in that field will be displayed.

Index Terms: Social Networking Sites; Graph search; Decision Tree Algorithm

1. INTRODUCTION

In a month billions of active users of similar interest sharing information such as messages, pictures videos etc. The graph search was launched to upgrade face book for the searching capability. The queries were phrased in simple English language and give the result from search. The personal information of friends, strangers used for threatening [1].For huge collection of data from media sharing sites, social network analysis techniques was used. To filter the data form noisy, unstructured nature,mining was performed by various SNA methods and tools to extract information useful for business, consumers, and users. This had the hidden attributes with higher degree of nodes[2].Yung-Tzu Joyce Lin *et al* [3] proposed three-layer model for the customers of the Facebook Fan Page of a fashion brand. Data clustering, correlation coefficient and linear regression methodology was proposed for data mining [3].

The LatentDirichlet Allocation (LDA) was used to retrieve the insights of the face book users such as likes andcomments [4].The decisionson collected data from web is the challenging task. The transitive voting systems was also introduced. This has limitation due to lack of transparency i.e. Credibilitydecisions were difficult to adopt and analyze [5]. Reality Miningwas based on sensed data from digital devices such as mobiles which were used to measure human physical and social activity. The Netviz, Gephi and Weka tools have been used to convert and analyze the Facebook [6].The social advertising uses the consumers perceptions with the brand and product and collection and analysis on the data would help to increase the advertisement efficiency [7].

Decisions are one of the most important elements required in making one's business to the acme of success. But decision cannot be solely taken on the core logic of an individual. The attributes that aggregate in forming the base of any important decision is information or processed data.The online social network platforms (e.g. face book, LinkedIn, Flickr, and Instant Messenger etc.) Contributes huge social network data. This data, if used and modelled into proper schemas can constitute in benefitting one's business by

* M. E. Pursuing, Computer Engineering, Armiet, Asangaon, *Email: vaishali.jadhav7@gmail.com*

** Assistant Professor, Department of Computer Engineering, Vidyalankar Institute of Technology, Mumbai, *Email: umesh.kulkarni@vit.edu.in*

giving him ability to make correct decision. But time and again many methods have been created, evolved nevertheless improvised but no viable and tacit mechanism exists. Thus the objective of this paper to provide concrete resolution with a paragon prototype. In the proposed system data is processed scrupulously with assiduous data mining algorithms such as to satiate the ever increasing needs of any business. The section 2 covers system architecture section 3 modus-operandi, section 4 covers result and discussion and section 5 describes conclusion followed by references.

2. SYSTEM ARCHITECTURE FOR USER GENERATED CONTENT ON FACE BOOK

Fig 1. Shows the system architecture to collect the user content from face book which consist of Input module, API, Graph search, ETL, decision Tree and as described below.

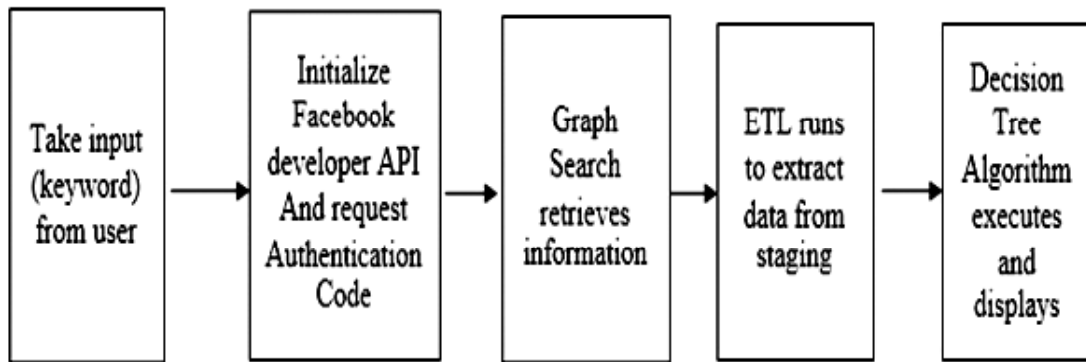


Figure 1: System architecture of proposed system.

Input Module Input module takes input as keyword or one category such as Entertainment, Education, Travelling, and Thinking.

Facebook developer API Using Facebook Developer API which Connect to Facebook developer account using normal Facebook account to get privileges. The authentication code is used to established connectivity with profile access.

Graph Search The collected information is pushed into staging table and will be stored the retrieved information by graph search algorithm into staging table.

ETL Information which is extracted from graph search is organised by ETL. ETL works likes a trigger and distinguishes the data from staging table and store into the database according to the database schema.

Decision tree algorithm Decision Tree Learning algorithm classifies and displays information.

This API is developed using Python.

3. METHODOLOGY

Flow chart 1 describes the step used in analysis as described below

3.1. Initialize Facebook Developer API

Connect to Facebook developer account using normal Facebook account to get the privileges.

3.2. Take authentication code to connect to profile

Get authentication code from the Facebook developer account and pass-in to the program for profile access. Fetch data from one's profile and retrieve using graph search algorithm. Using the authentication code, connect to the profile and by graph search algorithm search and retrieve all the information of one's profile.

Push the retrieved data into staging table. Store the retrieved information by graph search algorithm into staging table.

3.3. ETL runs to organize and extracts information

ETL works like a trigger and distinguishes the data from staging table and stores into the database according to the database schema. Decision Tree Learning algorithm classifies and displays information.

3.4. Business Model

After the analytical model is generated provide keyword to show the probability or possibility of success, and if yes then with which users. The tool is efficient both in terms of memory as well usage.

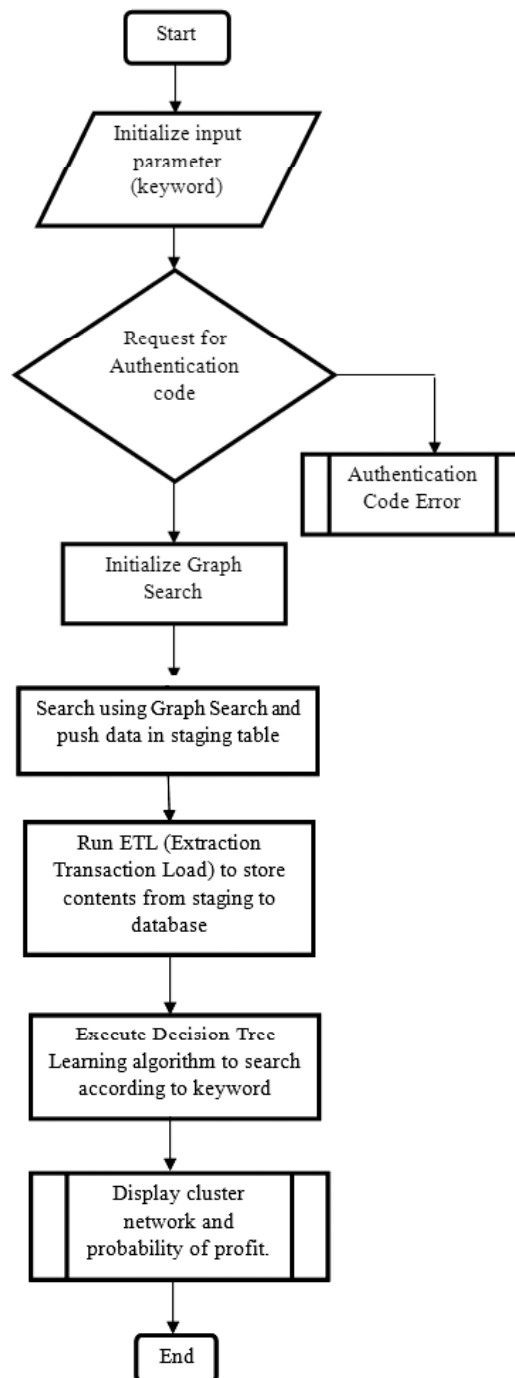


Figure 2: System flowchart of proposed system.

4. RESULT AND DISCUSSION

Fig. 2 shows the first stage of output which select user category. This tool is developed for four category to demonstrate the result to display selected category and related friend list .Initially first category entertainment is selected .The connectivity is established with port via 8080 local host. The generated output contents friend list with like, share and comment for that selected category post which helps any business client to post or share advertises with target group. Similar for the second ,Third and Fourth category output is as shown, which gives selected category with friend list who like, share or comment for the related category post on Face book.

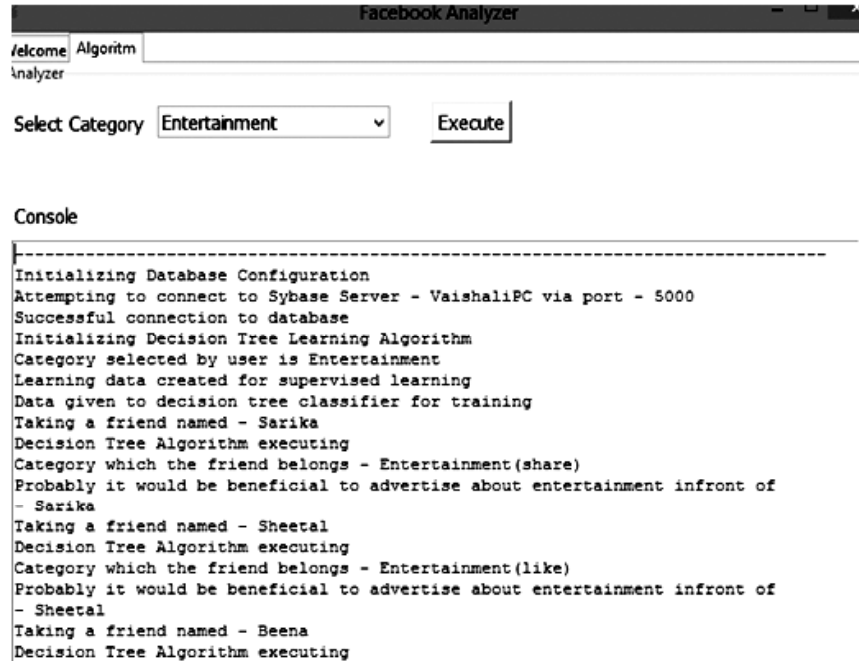


Figure 3(a): Input request Processing on category Entertainment

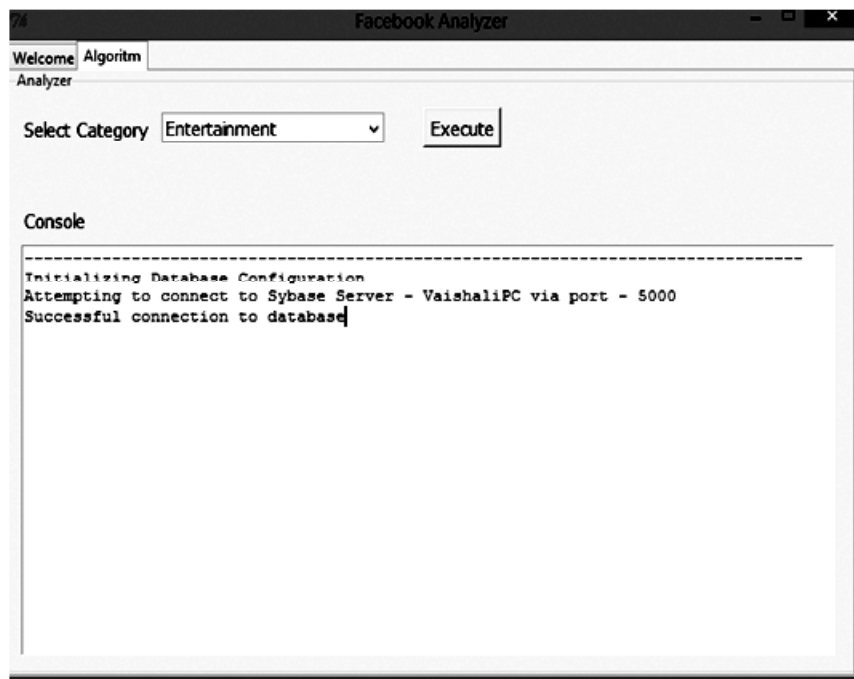


Figure 3(b): Input request Initialization category Entertainment

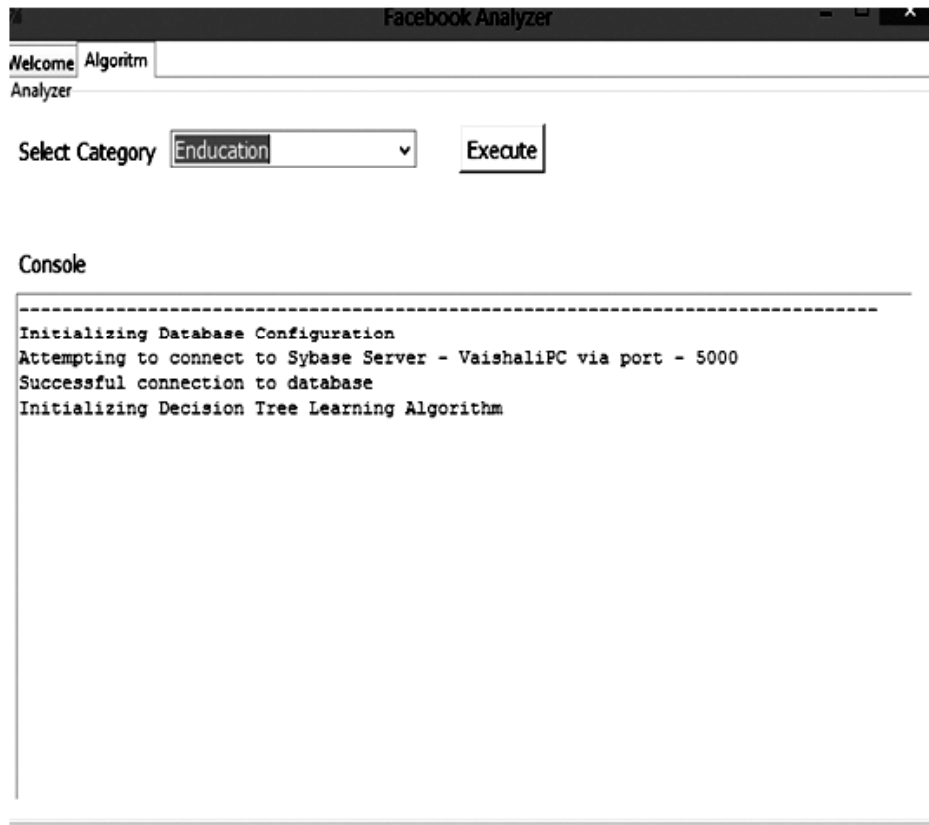


Figure 3(c): Input request Initialization category Entertainment

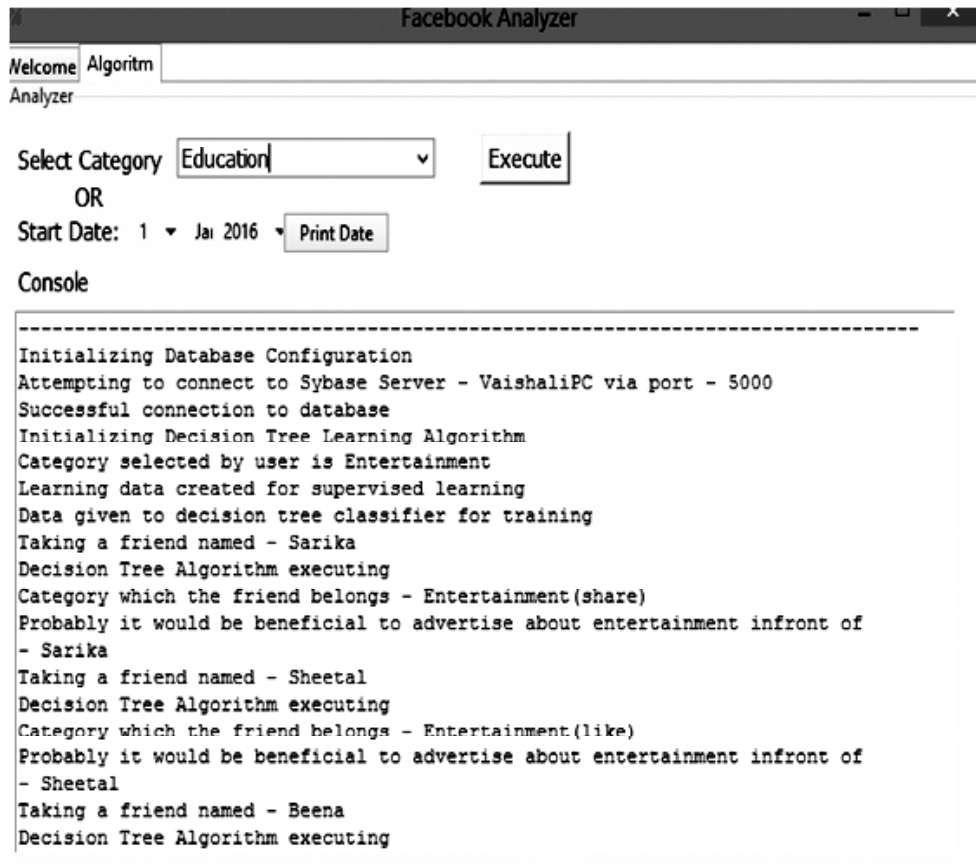


Figure 3(d): Input request Processing on category Entertainment

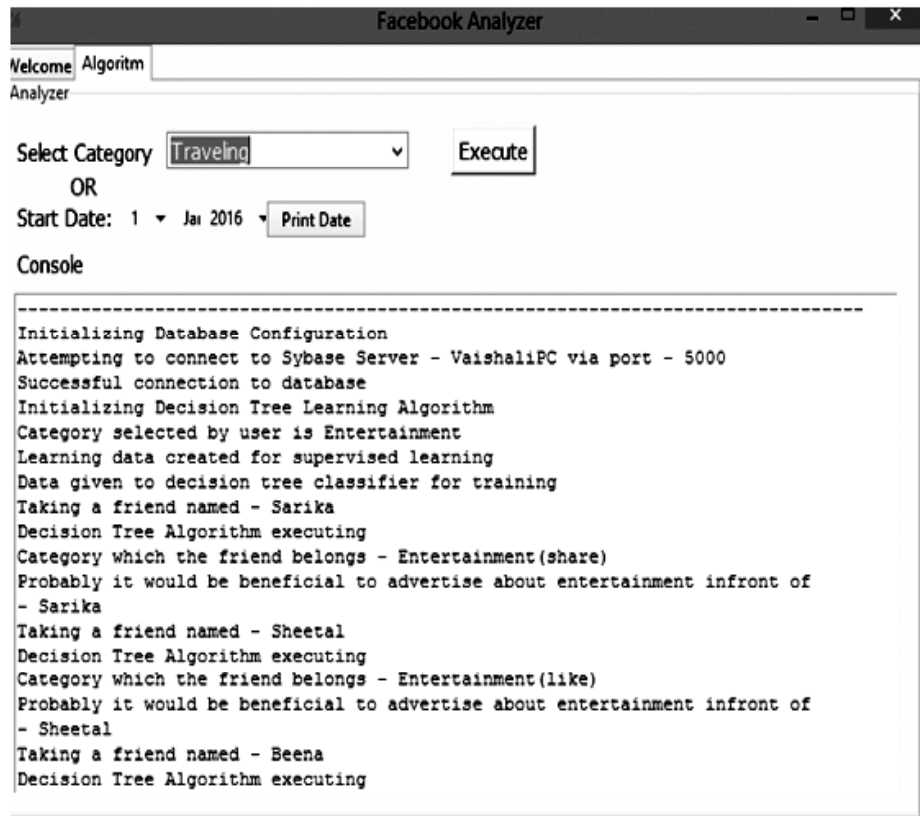


Figure 3(e): Input request Initialization category Travelling

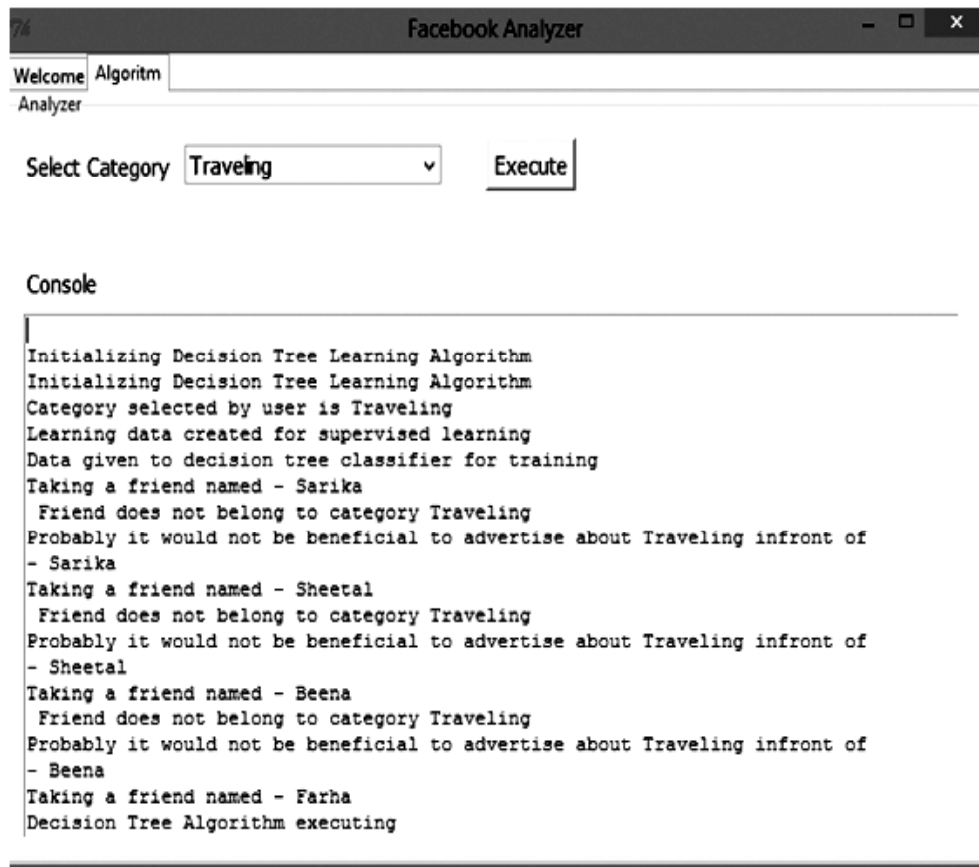


Figure 3(f): Input request Processing on category Travelling

5. CONCLUSION

This system gives decision about people who are interested in which kind categories so that any business client can use this data to post his business information to related interest of people. Thus, Face book is beyond just likes, shares, comments, pokes, opinion share blog, etc. and this can be proved completely true with our proposed system/prototype which will take the raw data, processes it and will give a hallmark model which can be used to make profitable decision in one's business in coming times. The API is developed on open source system using python.

REFERENCES

- [1] Zubeida Casmod Khan, ThulaniMashiane"An analysis of Facebook's graph search" Information Security for South Africa (ISSA), 2014, 978-1-4799-3384-6/14/\$31.00 ©2014 IEEE.
- [2] Nadeem Akhtar, Hira Javed, GeetanjaliSengar, "Analysis of Facebook social network"- 2013 5th International Conference on Computational Intelligence and Communication Networks, 978-0-7695-5069-5/13 \$26.00 © 2013 IEEE.
- [3] Yung-Tzu Joyce Lin, Meng-Yen Tom Lin, Kuo-Chen Li, "Consumer Involvement Model of Fan Page: Mining from Facebook Data of a Real Celebrity Fashion Brand"- Service Systems and Service Management (ICSSSM), 2015 12th International Conference978-1-4799-8327-8©2015 IEEE.
- [4] Yi-Chia Wang, Moira Burke, Robert Kraut, "Gender, Topic, and Audience Response: An Analysis of User-Generated Content on Facebook"- CHI '13 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 978-1-4503-1899-0© 2013 ACM.
- [5] Jeff Pasternack, Dan Roth, "Latent Credibility Analysis"- WWW '13: Proceedings of the 22nd international conference on World Wide Web, 978-1-4503-2035-1/13/05©2013 ACM.
- [6] Jyoti Sunil More, Chelapa Lingam, "Reality Mining based on Social Network Analysis"- 2015 International Conference on Communication, Information & Computing Technology (ICCICT), 978-1-4799-5522-0/15/\$31.00 ©2015 IEEE.
- [7] EytanBakshy, Dean Eckles, Rong Yan, ItamarRosenn, "Social Influence in Social Advertising: Evidence from Field Experiments"- EC'12, June 4–8, 2012, Valencia, Spain, 978-1-4503-1415-2/12/06...\$10.00©2012 ACM.

