Identification of Influential Users in Online Communities of Customers: Towards a Social Knowledge Management Approach

Hosniyeh S.Arian*
Department of Computer Engineering and Information Technology, Islamic Azad University
Qazvin, Iran
hosniyeh.safiarian@gmail.com

Omid R.B. Speily
Department of Computer Engineering and Information Technology , Urmia University of Technology
Urmia, Iran
speily@uut.ac.ir

Received: 2017/06/07 Revised: 2018/01/06 Accepted: 2018/03/14

***Abstract*—.Social media has made major changes in various e-commerce areas. One of these marketing cases is in e-commerce systems. The relationship between customers and business is very much appreciated by marketers. The use of social media by customers has given marketers the opportunity to get more information from customer feedback. Recently, in social media, marketers look for customers who have the most impact on other customers. They can influence the ideas of other customers with their opinions about a new product. In addition, influential users can have the greatest impact on specific domains. This domain may be in the domain of a product or service. Therefore, in this article influential users on social media have been studied in terms of impact in different areas. The proposed approach is for influential users using the social knowledge management approach. The knowledge cycle consists of knowledge organization, storage, retrieval, and knowledge discovery and knowledge management, where all explicit and implicit knowledge has been tried to accurately disclose affected users. In this paper, firstly, the problem was adapted to the knowledge management cycle, and in the steps of this cycle, artificial intelligence techniques such as Baysian networks were used to classify and identify influential users .In order to investigate the proposed method, various scenarios based on a variety of data sets are used for evaluation and the results of these studies show the high accuracy of the proposed method in identifying influential users.**

Keywords*— *influential users, publishing, knowledge management, social networks, e-shops, domains

1. INTRODUCTION

Recently, businesses have modernized their production, services, and tools to meet customer requirements. Businesses need to take innovative strategies for marketing, advertising and serving customers to overcome continuous changes in the environment and market requirements. One of the strategies is e-commerce stores.

In the years 1999 and 2000, when the Internet broke out, online shopping became known an Amazon became available. The first electronic store on the Internet was founded by Jeff Bezos [1]. At first, only books were sold online on this website. Following the success of Amazon, other booksellers physically present, sold their books online. Then portals like yahoo and msn began to build channels for selling items in addition to books for people [2]. Now, stores such as eBay and Amazon are the most famous and reputable online stores.E-shopping has many benefits, including no geographical constraints, no difference with the physical stores, board sales, wider and less costly advertising, fast track purchasing process, more accurate and convenient shopping planning [3]. Types of business in e-shops are divided into four categories: B2B e-commerce, B2C e-commerce, C2C e-commerce and C2G e-commerce [4].

One of the most powerful tools for marketing and advertising in electronic stores is social media. The speed of publishing news in such medium is so high that in the least possible time, they inform each other aware of the latest post published. The daily growth of online contributions and interactions suggests an increase in the number of participants (members) and an increase in the volume of documents on social networks [5]. This great volume of online stored data provides an exceptional opportunity for marketers and e-mailers to explore the behavior of customers. Through these interactions, researchers can now carefully review issues such as customer preferences for products, the quality of share of knowledge and information among members of the groups. Another benefit of advertising on the social network platform can be reduced costs and targeted targeting of customers.

One of the main challenges of advertising on social networks is increasing the effectiveness and spread of advertising on the level of social networks. Today, researchers have focused on the role of knowledge in gaining competitive advantage for the advent of a knowledge perspective in marketing. [6] Knowledge has replaced mechanical devices, capital, and raw materials to create important manufacturing resources for organizations. In addition to making profits for organizations by enriching it continuously and adapting its content to knowledge bases, knowledge increases its own value. On the other hand, increasing competition among companies has made knowledge management and market sensitivity important [7]. Discussion on knowledge management improves information about the importance of knowledge (new knowledge economy) and the impact on the competitive advantage of organizations. Knowledge management is known as a process, through which the organization generates, acquires, captures and deploys knowledge to enhance the productivity of the organization [8]. Knowledge management involves maintaining, sharing, improving and imparting knowledge in an organizational environment [9]. The objectives of knowledge management (KM) are based on three levels of tactics, developmental policies, short-term goals related to executive and operational needs, trends, internal rules, and everyday tasks [10]. KM requires that the organization organizes its products and interacts with customers, such as purchasers and suppliers of raw materials, not based on pre-defined public specifications, but based on customer preferences. While organizations move forward with the goals of customer relationship management, marketing functions play a key role in the interaction with the consumer and supplier [11].

The application of KM in marketing has empowered businesses to create a new type of needs more in line with today's consumer needs [12]. KM is a business strategy that helps corporations in the face of changes in the market environment, improving competitiveness and creating a coordinated internal organization structure. Companies tend to convince that they have a great understanding of their customers and market position; this misunderstanding makes managers make mistakes. In Ducker's view, only the customer perceives the market properly, so companies must conduct comprehensive market research and analyze customer behavior in the market, and must turn into active from the inactive market by predicting future customer needs [13].

This article seeks to find a way to determine the impact of content on social media users in order to maximize its propagation. To do this, one must identify the influential nodes in terms of the content of the advertisement. Influential users are the users who will maximize the range of advertisements in social media when advertisement is done by them. In other words, this is a matter of following users. If the advertisement is published, it is also desirable for the users who follow them, which will increase the publication of the post recursively.

KM cycle is used to explore the influential users in this paper, and we will begin to explore influential users based on KM cycle using the knowledge gained from the activities published by users in the social networking platform.

In the remainder of this article, we will present the literature in the second part, the proposed method will be presented in part three, we will review the proposed model in the fourth section and, finally, we will outline the results.

2. EASE OF USE

Here is a quick overview of the studies related to exploration of users. In [14], knowledge sharing activities on Yahoo's responses and favorite user profiles was studied. They combined the user attributes and the response specifications to predict a specific answer that would provide the best answer for users raised questions on social networking pages. In [15], the available methods for evaluating the quality of the content were discussed in a Q & A forum based on a strong bilateral relationship between user reputation and content generated by them. They improved a semi-regulatory approach to detecting high-quality content. In most cases, they examined studies that require users to vote for publications useful. For this purpose, they started to define criteria for determining the usefulness of the publication. However, most publications did not receive votes on the line, which did not mean that they were not useful. They chose three types of journalistic features (fundamental, linguistic, and semantic) and then using text mining and logical regressions, they assessed the impact of these three characteristics on the relevance of the publications. In [16], they examined prediction of the popularity of tweets on the Twitter community using the measure of tweeting propagation. They mapped the issue to a problem arranged using various features such as tweets, temporary information, metadata and users, and the characteristics of the social graph structure of users.

.in [17] a sentiment analysis method proposed that takes into account the social network information as well. they pay attention on the effect of influential users on the sentiment polarity of a topic based microblogging.in[18] two methods are proposed that can estimate the expected influence and diffusion time of a seed set in an efficient manner. Getting the set of all potentially optimal solutions helps a decision maker evaluate the trade-offs between the two objectives, i.e., the number of influenced users and diffusion time. In [19], the features of social marketing messages contribute to different levels of reputation. They reviewed the messages posted by the restaurants on their Facebook pages, and measured the importance of a message according to the number of likes. They also looked at message features by using keywords in content and focusing on media, links, videos, or photos. In [20], a method was proposed called H-INDEX-FAMILIY, which identified influential bloggers based on the function of the Index Hedgehog and its types. In [21], a method called MIIB was developed to explore the influential bloggers, which was examined using TUAW Dataset. They considered three features, including blogger reputation (number of links and comments received), blogger activity (number of content generated on the blog), and blog reputation (popularity of the blog among users). In [22], a method that examined the power of impact on user interactions and their impact on each other was examined. The model, which was tested on the Twitter and Facebook datasets, released interesting results. The content and sharing of activities do not have enough motivation to influence, and important events and exciting interactions lead to express more feelings among users, and large companies can take customers with these events and constructive relationships. In [23], an exploratory method was introduced to influence a particular product or service. In this method, the first time cut-off takes into account the users who turn into a commodity purchase. In the second time cut, for impact provided by each user, a threshold is determined and the impact that users place on each other is calculated. In the third time cut, with the discovery of influential users, our customers have discovered the extent of influence of influential customers on purchasing a product. The Summarized of literature is showed in Table 1.

TABLE I. SUMMARIZED OF LITERATURE

Reference	Model	Dataset
[14]	*Knowledge sharing activities and favorite user profiles was studied.*	*Yahoo's responses*
[15]	*a semi-regulatory approach to detecting high-quality content improved*	*Q & A forum*
[16]	*They mapped the issue to a problem arranged using various features such as tweets, temporary information, metadata and users, and the characteristics of the social graph structure of users.*	*twitter*
[17]	*A novel sentiment analysis method is proposed. Their approach extends the classical sentiment analysis.*	*microblogging*
[18]	*two methods which can estimate the expected influence and diffusion time of a seed set are proposed*	*microblogging*
[19]	*measured the importance of a message according to the number of likes*	*Face book*
[20]	*a method was proposed called H-INDEX-FAMILIY, which identified influential bloggers based on the function of the Index Hedgehog and its types*	*bloggers*
[21]	*method called MIIB was developed to explore the influential bloggers*	*TUAW*
[22]	*A method that examined the power of impact on user interactions and their impact on each other was examined.*	*Twitter and Facebook*
[23]	*An exploratory method was introduced to influence a particular product or service.*	*bloggers*

3. SOCIAL KNOWLEDGE MANAGEMENT METHOD

Social media has not only changed personal communication and interaction between them, but also transformed the way people work. By using social media, organizations are able to apply knowledge management optimally. It is possible to distribute and share better and access knowledge to the components of the organization using social media. This feature is very important in the world of today's business, given the increasing speed and complexity of the work environment. Social Knowledge Management is one of the social media applications of a business that can be used to customer sentiment analysis, social learning, or social collaboration among members.

According to the definition of social knowledge management, the use of social media in the field of knowledge management is to identify, share, document, transfer, develop, use and evaluate knowledge.Influential users have different influences in a variety of topics; in some areas, they show stronger influences, and in some cases they provide a weaker one. This paper presented a method called ECSN based on KM cycle to explore influential users based on permeability in a variety of topics, which is described in detail.

3.1. KM cycle

KM cycle is shown in Figure 1. Influential users will be explored based on this cycle. Subsequent steps will be presented in detail.

1) Knowledge maintenance

Knowledge management is defined as an interdependent set of functions that focuses on the recognition of individual knowledge and organizational knowledge, and most importantly the interaction between them that creates a cycle. Knowledge management is an attempt to reveal hidden assets in the minds of members and transform these hidden assets into an organizational asset to all the employees of the organization to access this asset.KM system tries to extract knowledge from users' personal accounts and their activities in various social networks. Knowledge is divided into two categories of explicit and implicit knowledge. Figure 2 shows the knowledge classification.

2) Knowledge organization

According to [24], knowledge is divided into structured and unstructured groups.

Structured Knowledge: Knowledge extracted from structured data and information resources that are structured in their own right. Structured data sources include databases, knowledge bases, data warehouses, and markets and datacenter databases.

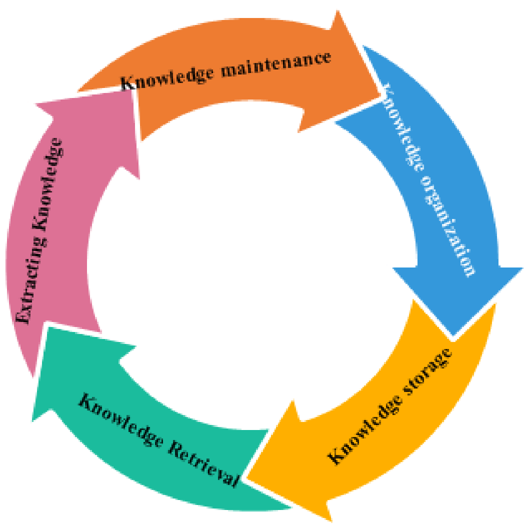


Fig. 1. KM cycle

Unstructured Knowledge: this is the knowledge from unstructured resources such as text documents, graphical indexes, audiovisual profiles, e-mail, and all tacit knowledge of employees.

a) Preparation of structured data

Structured data has a high degree of organization. To prepare structured data, four main steps take place as shown in Figure 3.

b) Preparation of semi-structured data

Semi-structured data is a form of structured data that does not conform to the formal structure of tables and data models dependent on interfacing databases, but include labels or symbols and indices that separate semantic elements from each other, and hierarchical descriptions of fields and fields between the data. Non-structured data refers to data that does not follow any predefined model, an example of which is data of heavy text. Figure 4 shows the preparation of semi-structured data.

3) Knowledge storage

Different conceptual models are presented for storing knowledge, one of which is the concept of the use of Bayesian network [25]. In many cases, we seek to find the best assumption in the space of H assumptions, with the possession of D training data. One way to express the best assumption is to say that we are looking for the most probable assumption, having D data plus the first knowledge of the possible possibilities of the H assumptions. Bayes' theorem provides a direct method for calculating these probabilities.

Some notation is needed to define Bayes' theorem. We use P (h) to express the initial probability that h is correct before we see the training data. P (h) is generally referred to as the former possibility, and represents any prior knowledge that speaks of the probability of correctness of the assumption h. If we have no prior knowledge of assumptions, we can assign the same probability to the entire space of H assumptions. Similarly, we use P (D) to express the previous possibility that D data are observed (in other words, the probability of observing D if no knowledge is available about the correctness of the assumptions). We also use P (D | h) to express the probability of D in a world where the assumption h is true. In machine learning, we seek P (h | D), that is, the probability of the correctness of assumption h provided viewing the training data D. P (h | D) is the probable latency h, because it expresses our assurance of the assumption h after observation D data. Bayes' theorem is the main foundation for Bayesian learning, since it provides a method for calculating the probability P (h | D) from the former P (h) with P (D) and P (D | h).

$$P(h|D) = \frac{P(D|h)P(h)}{P(D)} \qquad (1)$$

As expected, P (h | D) increases with increasing P (h) as well as P (D | h). Similarly, it seems logical that P (D | h) decreases with increasing P (D) because with increasing

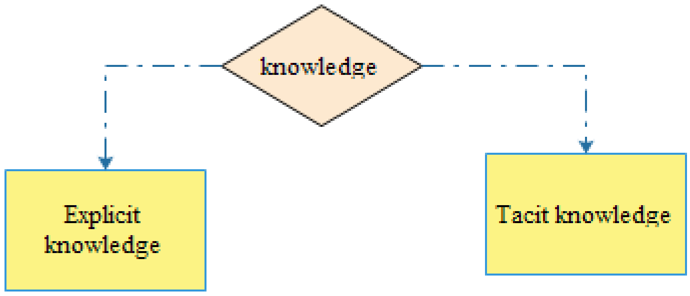


Fig. 2. Knowledge classification

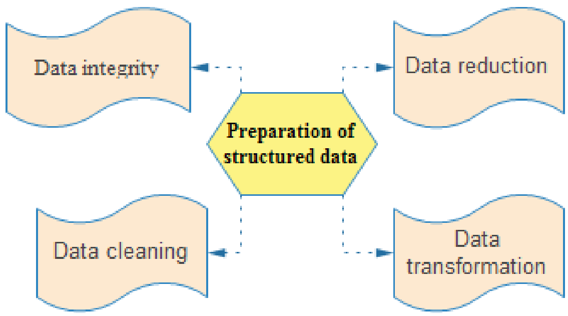


Fig. 3. Preparation of structured data

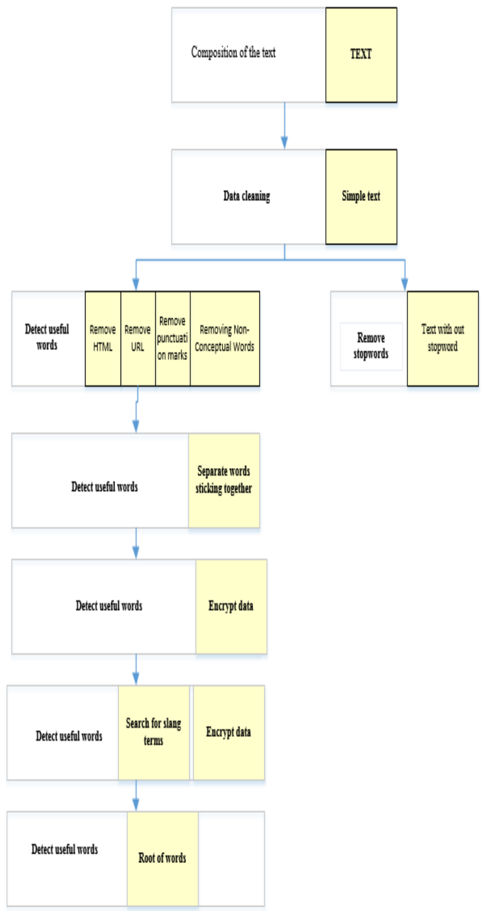


Fig. 4. Preparation of semi-structured data

probability P (D) occurring independent of h, there is less evidence for D in support of h.

4) Knowledge Retrieval

In this section, information is searched for, based on specific keywords in the text, diverse metadata and documents. One of the search methods is TF-IDF method. Depending on the content generated by the user, one begins to explore the keywords. Keywords are a collection of important vocabulary in a document that provides descriptions of documentary content that speeds up comprehension of text concepts. The method used to find keywords in this thesis is TF-IDF. This method gives words a weight based on the frequency of the document. In fact, this weighting system shows how much a word is important for a document. This method is a combination of two TF and IDF methods and is calculated from equation (2) [26].

$$W_{ki} = TFIDF(t_k, d_i) = tf(t_k, d_i) * idf(t_k, d_i) \quad (2)$$

5) Extracting Knowledge

The extraction phase is divided into two sections: detection of the important activities of users on the social networking platform and users classification by content, presented in more detail below.

a) Activities affecting users in propagation

Multiple regression models are used to explore the effective features of the proposed method. In multiple regressions, the value of a dependent variable (goal) is obtained from the values of independent (predictive) variables, which is achieved by constructing a linear relation (3):

$$y = b_0 + b_1(x_1) + b_2(x_2) + \cdots + b_p(x_p) \quad (3)$$

In relation (3) y is the dependent variable and x is independent variable. The regression coefficient is constant. B0 is constant regression coefficients. Parameters from b1 to bp are called partial regression and are calculated from Relation (4). These coefficients show the increase of the dependent variable increment for an incremental independent variable [27]:

$$\left(\frac{r_{y,x_p} - r_{y x_{p+1}} r_{x_p x_{p+1}}}{1 - (r_{x_p x_{p+1}})^2}\right)\left(\frac{SD_x}{SD_y}\right) \quad (4)$$

SDx is the standard deviation of the independent variable x, SDy is the standard deviation of the dependent variable y and, r is the correlation coefficient between the dependent and the independent variables.

b) Classification of Influential users based on Subject

To classify influential users by subject, the correlation between the keywords and the subject matter is first obtained. Since the variables are nominal, Kramer correlation coefficient is used. This correlation coefficient is between zero and one displayed with V2. As the correlation coefficient is closer to one, it indicates the strength between two variables. This correlation coefficient is obtained according to relations (5) and (6) [28]:

$$v = \frac{\sqrt{x^2}}{\sqrt{n * (k-1)(l-1)}} \quad (5)$$

$$x^2 = \frac{(F_O - F_e)}{F_e} \quad (6)$$

In relation (5), L is the number of columns, K is the number of rows and n is the number of samples. X2 is calculated from relation (6), F0 is the observed frequencies (actual frequencies that exist in houses), and Fe is the expected frequencies.

To explore the influential users in terms of trust and the activities they do, we used tree regression in their interests. For example, an activity should affect a certain number (>μ) of users. An activity must affect the other users in a given amount of time (> π), it must be repeated in a certain number (> α) of the activities performed by the leader. As shown in Figure 5, the rightmost leaf of the number of followers has an approximate size of 370 people a reissue of its is less than 40 minutes, and the operation is iterated over 60 times by the node, so its prediction in influence is close to 0.89, and can be considered the leader node in the future.

4. Findings

4.1. Twitter dataset

To collect Twitter search API data, all general data from the arts, sports and politics fields were extracted from July 10 and November 24, 2013. The data obtained consisted of three thousand one hundred twenty eight modules. Table 2 summarizes the data of the Twitter data set.

4.2. Performance Analysis

The purpose of this section is to compare the performance of ECSN model with models [29] and [19] that we will discuss.

Model [29] (Lee et al. 2010): This method is proposed for exploration of influential users. Data mining models are used, and parameters such as the number of generated

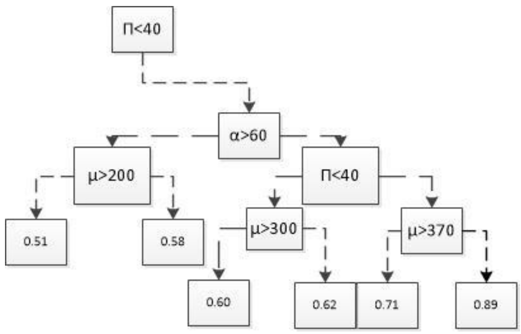


Fig. 5. An Example of Discovering Influential Users by Tree Regression

TABLE II. TWITTER DATASET

Data set	number
Number of words	*140*
Number of content	*5942*
Number of like	*1210*
Number of retweet	*518*
Number of comment	*358*
Number of mention	*498*
Number of users	*283*

content, the number of responses that users have in relation to the content, and the end of the number Content Generation Frequencies. This model is very suitable for viral marketing.

Model [19] (Yu et al. 2011): This model was proposed to explore influential users. The features of this model are generated content, as well as the type of media it is used for.

As shown in Figure 6, users groups begin to generate content in 6 ad groups as follows and shown in Figure 6 [30]:

1. Direct-Response Ad: This kind of advertising can be from any media, but its message varies by promoting the retail products, which is just the beginning to encourage consumers and customers to buy that product or service.

2. Public-service ad: public service advertising focuses on messages with good intentions such as not driving in drunkenness.

3. Interactive ad: In this ad, users share their opinion about a product or service and use it to make purchases or use of that service or merchandise.

4. Political ad: This kind of advertising is used by politicians, so that the people vote for them.

5. Cultural ad: Cultural advertising includes all types of advertising related to the culture of a nation.

6. Institutional ad: this is related to creating an identity and identifying a brand for customers.

As seen in the figure 6, interactive advertising has the most (re)posted ads on Twitter. Given the social nature of Twitter, it was predictable that interactive advertising would have the most advertising. This kind of advertising is very attractive for users and has a high number of publishers. Two types of cultural and political ads have the least (re)posting in this network. This can be due to the wide variety of users geographically cultural and political.

a) Influential users in the Twitter dataset for direct-response advertising

The number of influential users in the Twitter dataset for the direct-response ad is six. As shown in Fig. 7, ECSN model line-threshold limit exceeds models [29] and model [19].

As shown in the diagram, the direct-response advertising for users identified by the proposed method is more than other methods. The proposed method introduces the greater amount of direct-response advertising by employing the knowledge management cycle and more accurate ifluential users identifying method.

b) Influential users in Twitter dataset for public service advertising

The number of influential users in Twitter dataset is ten for public service advertising. ECSN propagation level is higher than the linear threshold of models [29] and [19], as shown in Figure 8, and on the other hand, in this propagation, for the number of users of one to two people, propagation for each model [29] and ECSN model is not equal.

Users who have been selected for public service ads as influential users in this dataset are 10 users. The release of these types of ads by these users is different as seen in the Figure 8. The propagation of this type of advertising through effective users identified by the proposed method is significantly higher than other methods.

c) Influential users in Twitter dataset for interactive ad

The number of influential users in Twitter dataset is fourteen for public service advertising. The models [29] and [19] have the same propagation rates, and ECSN propagation rate of both [29] and [19] is greater as shown in Fig. 9.

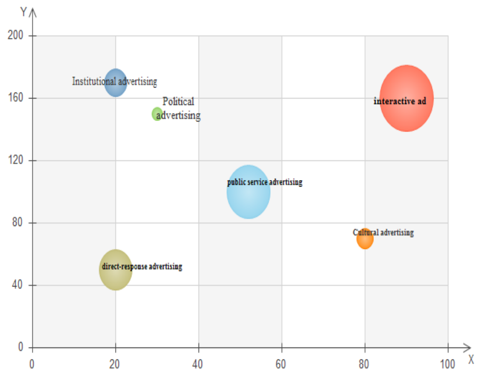


Fig. 6. Topics discussed by Twitter users

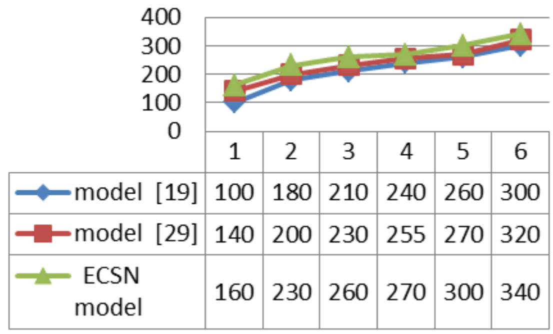
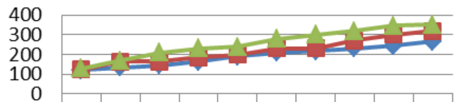
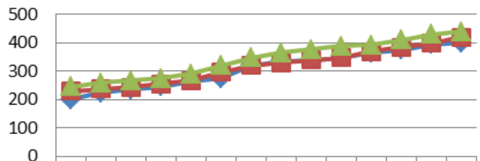


Fig. 7. The extent of publication of influential users in Twitter dataset in the subject matter of direct-response advertising



	1	2	3	4	5	6	7	8	9	10
model [19]	120	135	145	165	190	210	218	230	245	265
model [29]	125	165	165	185	200	230	230	270	300	320
ECSN model	128	169	210	230	240	280	300	320	345	350

Fig. 8. The rate of propagation of influential users in Twitter data set in the public service advertising



	1	2	3	4	5	6	7	8	9	10	11	12	13	14
model [19]	200	225	235	245	266	275	320	335	340	348	365	375	395	400
model [29]	230	238	243	256	268	297	320	330	339	348	370	385	398	420
ECSN model	245	260	268	275	290	320	348	365	378	389	392	410	430	440

Fig. 9. Propagation of influential users in Twitter dataset in the interactive ad

As shown in Figure 9, in interactive advertising, different methods have shown similar levels of propagation. The reason for this is the nature of this type of advertising. This type of advertising has a high potential for appealing content in terms of content and various methods for selecting influential users do not have a significant difference in the propagation of these types of ads.

4.3. Performance Evaluation

To evaluate the proposed model, Three well-known measures, namely “precision”, “recall”, and “F measure”, are used (Formulas 7 to 9 show these measures) . precision (also called positive predictive value) is the fraction of True influential users (IU) among the predicted IU’s, while recall (also known as sensitivity) is the fraction of True influential users that have been predicted over the total amount of IU’s. This relation shows all the things that must be met by a category in order to be able to achieve the desired performance. The results are obtained by a 10-fold cross-validation. In this type of validation, data is subdivided into 10 subsets. Of these 10 subsets, each time, one is used for validation and the other 9 for training. This is iterated 9 times, and all data is used once for training and once for validation purposes. Finally, the average result of this nine-time validation is selected as a final estimate. The performance of the proposed model is compared with models [19] and [29]. The results are shown in Table 3. The precision, recall and F1-measuer of the proposed model are higher than those of [19] and [29]. ECSN model has better accuracy and prediction than [19] and [29].

$$Precision = \frac{|\{\text{Predicted IU}\} \cap \{\text{True IU}\}|}{|\{\text{Predicted IU}\}|} \quad (7)$$

TABLE III. COMPARISON OF THE PROPOSED MODEL PERFORMANCE

Evaluation	F measure	Recall	Precision
Proposed Repost Prediction	***0.643***	***0.541***	***0.793***
Model [19]	***0.554***	***0.477***	***0.662***
Model [29]	***0.51***	***0.42***	***0.61***

$$\text{Recall} = \frac{|\{\text{Predicted IU}\} \cap \{\text{True IU}\}|}{|\{\text{True IU}\}|} \quad (8)$$

$$F = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (9)$$

The proposed method works better than other methods in all three criteria. The selection of influential users based on the type of advertising and the use of the social knowledge management cycle are the reasons for this improvement in identifying influential users in ads propagation.

5. DISCUSSION

This paper identifies influential users in the ads propagation process based on a KM-based approach. The most important limitation of this research was the lack of a diverse set of data along with metadata needed to extract relationships. It's not easy to collect data about ads within user posts. The innovations of the proposed method are relevant to the type of advertising in the selection of influential users. In the proposed method, influential users are selected based on the type of advertising, and the results of the surveys indicate that the different types of advertising are disseminated more widely by proposed method. One of the positive features of this article is the use of the knowledge management cycle in the advertising and marketing process of social networks. Today, a large amount of business advertising in cyberspace is using social networks. This article is one of the first researches in the field of social knowledge management in business marketing. The positive results of this research against the research conducted without considering knowledge management show the effect of this approach. The accuracy of identifying influential users in propagating ads is significantly higher than previous methods. As future work, review the content of the advertisement based on criteria such as the negativity of the message, the duration of impact and the impact of events outside the social network environment on the impact of advertising. For this purpose, valid data sets must be collected at different times. Improving the knowledge management cycle and improving the predictability of influential users is another important future work.

6. . CONCLUSION

Influential users in social networks will maximize the dissemination of information. Companies tend to identify these users on social networking platforms. Various methods have been proposed to identify these users in various researches. A method is proposed in this paper called ECSN - exploring these users is based on a subject-based social networking platform in the KM cycle. Users will have a great impact on one subject and not on other areas. KM cycle includes knowledge maintenance, knowledge retrieval, and knowledge exploration. At each stage, attempts are made to explore influential users. The proposed model was compared with models [29] and [19]. The propagation rate of the proposed model is higher than the two models [29] and [19]. In the future, efforts should be made to consider parameters such as the release time and the degree to which users trust each other to select a product.

References

[1] Wang, C., Jin, X.L., Zhou, Z., Fang, Y., Lee, M.K.O., Hua, Z. 2015. The effect of perceived media capability on status update in microblogs. Electronic Commerce Research and Applications, this issue.

[2] Phang, C.W., Tan, C.H., Sutanto, J., Magagna, F., Lu, X. 2014a. Leveraging O2O commerce for product promotion: an empirical investigation in Mainland China. IEEE Transactions on Engineering Management, 61(4), 623-632.

[3] Tan, C.H., Sutanto, J., Phang, C.W., Gasimov, A. 2014. Using personal communication technologies for commercial communications: a cross-country investigation of email and SMS. Information Systems Research, 25(2), 307-327.

[4] Hayat, Tsahi Zack, Ofrit Lesser, and Tal Samuel-Azran. "Gendered discourse patterns on online social networks: A social network analysis perspective." Computers in Human Behavior 77 (2017): 132-139.

[5] O'Connor, Kathleen M., and Eric Gladstone. "Beauty and social capital: Being attractive shapes social networks." Social Networks 52 (2018): 42-47.

[6] Tarbush, Bassel, and Alexander Teytelboym. "Social groups and social network formation." Games and Economic Behavior 103 (2017): 286-312.

[7] Kim, Soyeon, Jay Kandampully, and Anil Bilgihan. "The influence of eWOM communications: An application of online social network framework." Computers in Human Behavior 80 (2018): 243-254.

[8] Bai., M. 2012. Exploring the dynamics of rumors on social media in the Chinese context. Unpublished master's thesis, Department of Informatics and Media, Uppsala University, Uppsala, Sweden.

[9] Trainor, K. J., J. Andzulis, A. Rapp, and R. Agnihotri, 2014. Social media technology usage and customer relationship performance: A capabilities-based examination of social CRM.. J. Bus. Res., 67 (6): 1201–1208

[10] Vermeulen, B. (2007). Knowledge based method for solving complexity in design problems (Doctoral dissertation, TU Delft, Delft University of Technology).

[11] Milton, N. R. (2008). Knowledge technologies (Vol. 3). Polimetrica sas.

[12] Choudhury, M. and P. Harrigan, 2014. CRM to social CRM: the integration of new technologies into customer relationship management, J. Strateg. Mark., 22 (2): 149–176, 2014.

[13] Wymbs, C. (2011). Digital marketing: The time for a new "academic major" has arrived. Journal of Marketing Education, 0273475310392544.

[14] Adamic, L. A., Zhang, J., Bakshy, E., & Ackerman, M. S. (2008, April). Knowledge sharing and yahoo answers: everyone knows something. In Proceedings of the 17th international conference on World Wide Web (pp. 665-674). ACM.

[15] Stieglitz, S. and L. Dang-Xuan, 2013. Emotions and Information Diffusion in Social

[16] Cao, Q., Duan, W., & Gan, Q. (2011). Exploring determinants of voting for the "helpfulness" of online user reviews: A text mining approach. Decision Support Systems, 50(2), 511-521.

[17]]Eliacik, Alpaslan Burak, and Nadia Erdogan. "Influential user weighted sentiment analysis on topic based microblogging community." Expert Systems with Applications 92 (2018): 403-418.

[18] Mohammadi, Azadeh, and Mohamad Saraee. "Finding influential users for different time bounds in social networks using multi-objective optimization." Swarm and Evolutionary Computation (2018).

[19] Yu, B., Chen, M., & Kwok, L. (2011, March). Toward predicting popularity of social marketing messages. In International Conference on Social Computing, Behavioral-Cultural Modeling, and Prediction (pp. 317-324). Springer Berlin Heidelberg.

[20] Bui, D.-L., Nguyen, T.-T., & Ha, Q.-T. (2014). Measuring the influence of bloggers intheir community based on the H-index family. In 2nd international conference on computer science, applied mathematics and applications (ICCSAMA), Budapest

[21] Khan, H. U., Daud, A., & Malik, T. A. (2015). MIIB: A Metric to identify top influential bloggers in a community. PloS one, 10(9), e0138359.

[22] Wakefield, R., & Wakefield, K. (2016). Social media network behavior: A study of user passion and affect. The Journal of Strategic Information Systems, 25(2), 140-156.

[23] Zhu, Zhiguo, et al. "Exploring factors of user's peer-influence behavior in social media on purchase intention: Evidence from QQ." Computers in Human Behavior 63 (2016): 980-987.

[24] Bebensee, T., Helms, R., & Spruit, M. (2012). Exploring Web 2.0 applications as a mean of bolstering up knowledge management. Leading issues in social knowledge management, 1, 22.

[25] Kardan, Ahmad A., Omid RB Speily, and Yosra Bahrani. "Modelling the Effectiveness of Curriculum in Educational Systems Using Bayesian Networks." arXiv preprint arXiv:1506.02794 (2015).

[26] Salton, G., & Yang, C. S. (1973). On the specification of term values in automatic indexing. Journal of documentation, 29(4), 351-372

[27] Bian, J., Liu, Y., Zhou, D., Agichtein, E., & Zha, H. (2009, April). Learning to recognize reliable users and content in social media with coupled mutual reinforcement. In Proceedings of the 18th international conference on World wide web (pp. 51-60). ACM.

[28] Kemper, A. (2010). Reconsideration of Valuation in Software Markets. In Valuation of Network Effects in Software Markets (pp. 43-63). Physica-Verlag HD.

[29] Li, Y. M., Lin, C. H., & Lai, C. Y. (2010). Identifying influential reviewers for word-of-mouth marketing. Electronic Commerce Research and Applications, 9(4), 294-304.

[30]] Chinchanachokchai, S., Duff, B. R., & Sar, S. (2015). The effect of multitasking on time perception, enjoyment, and ad evaluation. Computers in Human Behavior, 45, 185–191

Hosniyeh Safiarian received the B.Sc. degree in Computer Engineering from Pnu University, the M.Sc. degrees in Information Technology from the Qazvin Islamic Azad Univeristy. She is currently the Ph.D student at Science and Research Branch Islamic Azad Universiity of technology. Since 2014 She works as a researcher at Saman Electronic payment (SEP). her research interest includes the data mining, graph theory, intelligent systemand social media.

Omid Reza Bolouki Speily received the B.Sc. degree in Computer Engineering from Urmia University, the M.Sc. degrees in Information Technology from the AmirKabir University of Technology. He is currently the Ph.D candidate at Amirkabir Universiity of technology. He worked as a researcher at the Iran Telecommunication Research Center (ITRC). Since 2009 he joined the Urmia University of Technology as a faculty member of Information Technology & Computer Engineering Department. His research interest includes the dynamics complex networks, graph theory, intelligent system, e-Services.