ADVERTISING RECOMMENDATION SYSTEM BASED ON DYNAMIC DATA ANALYSIS ON TURKISH SPEAKING TWITTER USERS

Onur Sevli, Ecir Uğur Küçüksille

Original scientific paper

Online environments and especially social networks have become a great alternative to advertisement publishing. In order to accomplish effective advertising it is important that the contents coincide with the expectations of the target audience. Considering that expectations may change over time, it is required to identify the orientation of the users in real time and dynamically. In this study, the messages shared by Turkish Twitter users were analysed in real time and the instant expectations of the users have been identified. To perform this work, a web service was designed which analyses the user's profile and presents the advertisements that suit best to expectations. A method called Heuristic Pruning Method (HPM) has been revealed in order to filter the most appropriate advertising content. The developed system has been tested on a voluntary participant group who actively uses Twitter, and the effectiveness of the system is demonstrated by the received feedback.

Keywords: advertising recommendation; content filtering; map reduce; social network marketing

Preporučeni sustav oglašavanja zasnovan na dinamičkoj analizi podataka turskih korisnika Twittera

Izvorni znanstveni članak

Online okruženja, a posebno društvene mreže postala su snažna alternative objavljivanju oglasa. Za učinkovito oglašavanje važno je da se sadržaj poistovjećuje s očekivanjima ciljane publike. Uzimajući u obzir da se očekivanja mogu s vremenom promijeniti, potrebno je u realnom vremenu i dinamički prepoznati orijentaciju korisnika. U ovom su se radu u realnom vremenu analizirale poruke turskih korisnika Twittera i identificirala njihova trenutna očekivanja. U tu je svrhu dizajnirana web usluga koja analizira profil korisnika i daje oglase koji najbolje odgovaraju očekivanjima. Za filtriranje odgovarajućeg sadržaja oglašavanja korištena je metoda nazvana heuristička metoda odstranjivanja suvišnog (Heuristic Pruning Method - HPM). Razvijeni sustav je testiran na grupi volontera, aktivnih korisnika Twittera, a učinkovitost sustava se pokazala dobivenom povratnom informacijom-feedbackom.

Ključne riječi: filtriranje sadržaja; marketing društvene mreže; preporuka oglašavanja; redukcija mape

1 Introduction

Advertising is the announcement that is exhibited by purchasing space or time from communication devices, for the purpose of trying to attract peoples' attention to a product or service [1]. It is necessary to use media such as television, radio and internet for the broadcasting of ads. In ad publishing, one of the ideal ways to meet the expectations of increasing sales with a low cost of advertising is to use online media publishing. Although traditional media such as television and radio are still widely used, in terms of cost and immediate access to a wide audience they have fallen behind the internet. Since the beginning of this century, the use of internet and the time users spend on the internet has increased rapidly. The fact that internet allows reaching a wide audience regardless of time and space and in an affordable way, makes it an ideal medium in terms of today's internet advertising.

Online media is one of the most advantageous ways to increase brand recognition and to create awareness. Realizing the power of Internet, the companies have turned to the use of online resources and social networks, especially in the marketing field. Data on social networks are quickly spread to a wide audience and offer the ideal solution in terms of cost-effectiveness. In addition, the social network platforms which contain personal information and geographical location data of users, facilitate reaching the right audience based on location as well.

Among the social networks, Twitter has performed a tremendous growth recently [2]. Although it is quite new as an advertising medium, it has a huge potential for advertising due to the increase in its number of users and

due to its infrastructure. With 500 million registered users and 271 million active instant users as of 2014, more than 340 million messages are shared daily on Twitter. As a result of the analysis of these data, we are able to correctly identify users' instant expectations.

In social networks, users' interests can vary over time, so identification of instant expectations is extremely important in order to present the most relevant content. Ads that cannot adapt to users' instant orientations are ineffective for the user, as well as causing unnecessary expenses for the company promoting a product or service. Therefore, marketing researchers have turned to marketing efforts that target the presentation of the most relevant content by analysing the users' expectations. Thus, the user is prevented from coming face to face with unnecessary content, contents are reached to the right target, and marketing strategies become more efficient.

In this study an advertising recommendation system has been developed which facilitates the delivery of correct content to users, based on the Twitter platform. For this work, shares of a group of users who make sharing in Turkish have been processed with the natural language processing and big data analysis techniques in order to make them structural and semantic. Personal interest areas have been identified by categorizing the most commonly used word patterns in the sharing. A web service has been designed which presents to the users the contents, marked by a variety of categories and keywords on advertising database and suited best to the user expectations. A content filtering method called Heuristic Pruning Method (HPM) has been revealed in order to filter the most advertising content that suits best the expectations. This method gives scores to ad contents with weight values calculated based on the users'

identified fields of interest and frequency of use of repeating words, and it obtains the optimal solution set by eliminating content with low iterative suitability.

The developed system analyses the shares on users' timeline by using the screen names as parameter. It serves ad recommendations produced as a result of the analysis, to the requesting applications in a way independent from the platform. A prototype system has been developed to test the application. This application has been tested on a voluntary group of users actively using Twitter. Based on the feedback received from this voluntary group the system has been found to be successful at the rate of 88%.

2 Literature review

Recently, the web environment has become more public and real-time [3]. Especially the popularity of social networks is rapidly increasing in information sharing [4]. Platforms such as Twitter, Facebook, and Flickr are the communities that bring together numerous individuals online or offline, and Twitter has recently grown rapidly. Kwak et al. have examined Twitter platform and they revealed its structural properties [5]. Jones has emphasized on how the Twitter platform affects interaction between individuals [6].

The amount of content generated on social platforms is growing rapidly, users are often faced with a huge amount of social data [7]. Many studies conducted in recent years have focused on filtering the content that appeals to the user in this great collection of data [8]. Content filtering is performed according to the users' orientations that are detected based on their activity on social networks [9]. Bosch et al. have developed a system that conducts real-time filtering in order to acquire the contents that appeal to the customer out of the share clusters the user is exposed to [10]. Based on the shares that are made, Cataldi et al. have tried to make inferences about the issues on agenda [11]. On the other hand, Kang et al. sought to determine the connection between the events in addition to the event inferences [12]. Yang and Rim, through the topical analysis model they developed, sought to identify interesting issues in the tweets [13]. Chorley et al., by getting clues about the Twitter users' feelings from their shares, conducted a study for the purpose of filtering the contents in the timeline [14]. On the other hand, Dai et al. conducted a study on the feeling analysis and propagation characteristics on audible social networking sites [15].

In social networks many users remain unaware of other users and subjects that can be followed [16] therefore social networking recommendation mechanisms come to the forefront. Armentano et al. sought to identify users associated with each other based on the correlation factor [17]. Li et al. proposed a graphics-based model for user recommendation in social networks [18]. Sudo et al. developed a recommendation system which would encourage users to interact with new users through words and follower relations [19]. Jamil et al. developed an application that recommended to the user similar users in the nearby locations through the location information obtained from the user's profile and through the filtering technique they developed [20]. Zhou et al. developed a system that recommended related users to each other by tagging the users through the analysis they conduct on the Twitter and Netease sample data sets [21]. Kim and Shim

developed an algorithm that produced probability-based advice on establishing a connection between popular users and tweets [22]. Islam et al. put forward a strategy that made recommendations in line with the data they obtained about Twitter users' past friendships [23]. Lee et al. developed a system that made news recommendations to Twitter users based on the keywords they obtained from tweets, retweets and hashtags [24]. Hashtags are an important way of accessing the messages associated with each other in social networks. Lu and Lee developed a model that produced hashtag recommendations [25]. Jonnalagedda and Gaucher designed a hybrid application which identified the popular news based on commonly used words on Twitter, and from among these words which recommended the words that complied with the profile data self-defined by the user [26].

Social platforms have great importance in the commercial field. It has been found that companies using these platforms exhibit a more rapid growth than non-users. In these platforms there are comments about companies, products or services including advices or complaints. The companies that take into account these comments exhibit a positive development. Vos and Verbeke examined these discussions performed on companies in social networks in their studies [27]. D'Avanzo and Pilato on the other hand focused on the opinions given by the customers on social networks, and emphasized on how they gave direction to shopping activities [28]. Lee et al. developed a framework that detected data on product campaigns in the shares [29]. Spina et al. conducted a study that distinguished whether some words referring to specific brands and also used in daily life are in fact brands or ordinary words [30]. Ghiassi et al. carried out a study based on artificial neural network over the tweets, intended at analysing the users' views on any brand [31]. Nettelhorst et al. in their study conducted on students, studied the effects of user expectations on the advertising choices [32]. In a study by Burkhalter et al. it was found that word of the mouth is extremely effective among the social network marketing methods [33].

3 Material and methods

In this study, data belonging to users who make shares in Turkish on Twitter are analysed with the natural language processing and map reduce method, and a service is developed which identifies the user's interests and makes ad content recommendations accordingly. The word patterns in the contents shared by users are identified and categorized, and the user's areas of interest are found based on the frequently used word patterns and specified categories. The ad contents that are included in the advertisement database of the developed system, and which are marked with specific categories and keywords, are filtered and served in accordance with the user's identified areas of interest.

3.1 Turkish natural language processing

Turkish language structure consists of the roots and affixes. There are three basic types of word roots according to the meaning and duty, which are names, verbs and prepositions. Name refers to living, non-living things, abstract or concrete concepts, emotions, thoughts

and situations. In the name type, the following sub-types are included: adjectives, pronouns, adverbs.

Adjectives are the words that come before a name and affect its meaning. Pronouns are the words that represent beings and which are used temporarily in lieu of names. On the other hand, adverb limits the meaning of an adjective or adverb in terms of location, direction, time, measurement, and question.

Verb refers to a deed or action carried out depending on time. Prepositions are the words that don't have a meaning alone, but establish relations between the other words in the sentence.

Affixes are added to Turkish word's roots in order to specify their roles in the sentence or in order to form new meanings. They have two types, which are the derivational affix and inflectional suffix. Derivational affixes are used for deriving a new and different word from the original word. On the other hand, inflectional suffixes are used for conjugating the same word for different locations and tasks.

In this study, a Turkish natural language processing library was used. This library has a data repository containing the roots and suffixes of Turkish words. The pool contains approximately 30 000 words. In the language processing, firstly it must be determined whether the word is Turkish or not, and its spell check must be done. Morphological analyses are carried out to accomplish this task. Out of the pool containing the word roots, the candidates which can be the root of the word in question are identified. After this, the possible affixes are added to the word root and there the word sought is tried to be found. If the word is found, then it means that the appropriate root and affixes are found. In order to identify the word root candidates, a search is conducted on the roots placed in double tree structure (Fig. 1). This tree structure allows the root candidates to be identified quickly.
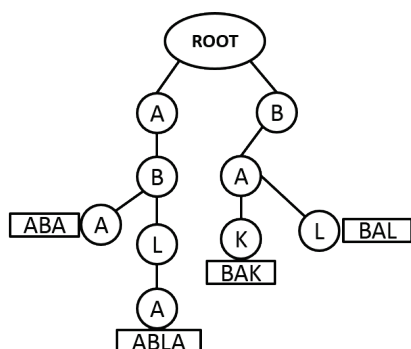

Figure 1 Tree structure containing the word roots

3.2 Map reduce function

MapReduce is the process in which large amount of data is divided into small blocks, and mapping and reducing the processes is carried out. It consists of two functions, "Map" and "Reduce". The data divided into individual parts are converted into the key-value pairs by the Map function. Then, this data is organized and transmitted to the Reduce function. Reduce function generates groups and a summary result whose data depend on the keys.

The MapReduce function consists of five basic steps:

- Preparation of input data
- Running the Map function
- Arrangement of outputs and transfer to Reduce
- Running the Reduce function
- Production of final output.

In the present study, MapReduce technique is used to determine the frequently repeated phrases in the content which is shared by the user, and to categorize words and identify areas of interest.

4 Advertising recommendation system based on dynamic data analysis on Twitter

Advertising recommendation system based on dynamic data analysis on Twitter is a web service that analyses the users' shares by using the natural language processing and MapReduce method, and which determines in an exploratory way the most relevant content contained in an advertising cluster in line with the findings acquired from the conducted analysis, and which presents this contents to the user. Twitter user's data is obtained by using the unique screen name used by Twitter while tagging the user. Out of these data, the commonly used word patterns are found and categorized, allowing the detection of the user's real-time area of interest. Online and offline current Turkish dictionary databases are used for the categorization of words.
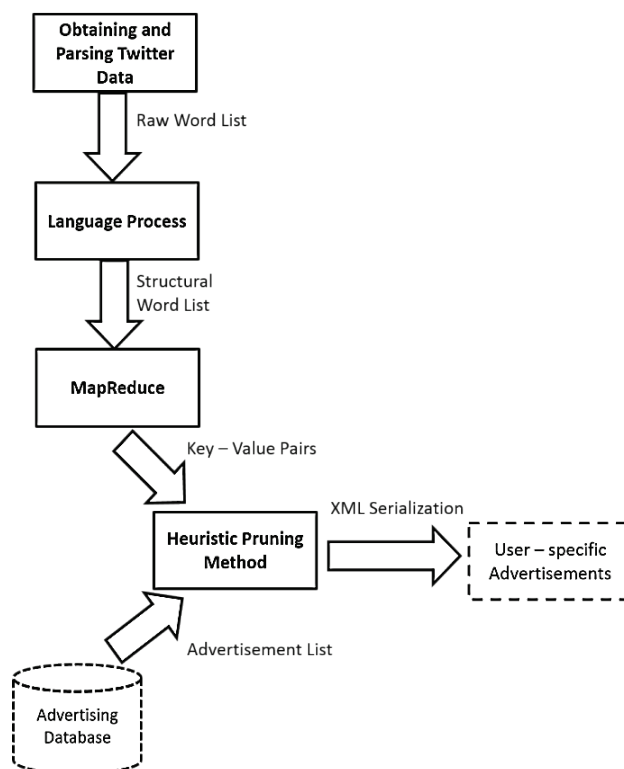

Figure 2 Project work flow chart

The functioning of the system consists of four basic processes (Fig. 2). These are: obtaining and separation of the Twitter data, language processing procedure, MapReduce process, and the filtering of the appropriate ads with HPM. The advertising content identified as a result of this process, is presented as XML output in a way that suits the user's interests best.

4.1 Obtaining and separating Twitter data

Twitter offers Application Programming Interface (API) which can be integrated into applications, allowing access to user data via software. Through the Twitter API, access to a user's shares made with his screen name is possible. For access to shares, the user's profile must be open or there must be access permission for the system that uses API. For ethical purposes, the present study was conducted on the users who granted permission to access to their profiles.

In the Tweets, the following are present: colloquial phrases, references to other users, retweet expressions and link URLs. As a result of the extraction and analysis of data in relation to spoken language, information about the user's expectations and emotions can be obtained.

In the first stage of data analysis the shares are parsed into sentences which are the most basic unit of the spoken language having semantic integrity. After this, the following are cleaned which cannot be subject to analysis: URLs, retweet (RT) statements, references, punctuation marks, exclamation and meaningless phrases. Subsequently, the sentences are divided into words and word patterns are determined by comparing them to pre-defined word groups.

After the parsing process, word groups are passed to language processing to perform semantic analysis. Through this process, words in the form of raw text attain a structured form.

4.2 Language processing

In this process, the morphological and semantic analysis of partial words is carried out. In the first stage, the obtained words are firstly subjected to spelling check. Then each word is marked according to its root type. After identifying the current word's root and suffixes, the type of root and the positive / negative status of the suffix is obtained by querying from the data repository in the word processing library. It is identified whether the sentence is negative or not, by analysing the negative words in the sentence and the types of these words. In general Turkish sentences consist of three parts: subject, complement, predicate. Subject is the doer, object is the part which completes the sentence, and verb is the main predicate expressed in the sentence. A sentence which has a negative verb and which does not contain any other negative words might be thought to be negative in terms of meaning. On the other hand, in case a negative action is complemented by a negative verb, a positive situation in terms of semantics may emerge.

A Turkish word contained in a sentence may have a meaning on its own, or it may lack meaning when it is independent. The words with name roots generally point out to a being or concept in real life, and they have a meaning on their own. However, words such as adverbs and prepositions do not carry a meaning on their own but they add additional meanings to other words. In the present study, words that do not carry a meaning on their own are not included in the analysis process, but the words with name roots have been interpreted.

In order to find into which area of interest the acquired meaningful words fall, queries are made on online and offline current dictionary database that include word-category couples.

At the end of these stages, the raw word in the beginning is transformed into a structure that contains properties in relation to root, type, negative/positive status, and category information. After the language processing procedure, the word groups are transmitted to the MapReduce process in order to be grouped and summarized in accordance with their common properties.

4.3 MapReduce process

At this stage, the words and category data obtained from the user profile are processed. This process divides large amounts of data into processable pieces, and forms a cluster by summarizing them according to specific properties. It consists of a combination of two different functions called "Map" and "Reduce". Firstly, the input data is separated into blocks. The Map function processes each word and category and transforms them into key-value form. The key phrase is the root information belonging to the word, and the value is the word's frequency data. In a similar way, the category data is the key, and the frequency belonging to it is the value. After the creation of the key-value couples, similar words and categories are brought together. After this, obtained result is transmitted to the Reduce function. Reduce function groups the word phrases in organized form. In the output summary, the word root or the category phrase is the key, whereas the total value of the frequencies makes up the value.

The pseudo code representing the MapReduce function is as follows:

function map(**string** dataBlock)
for each word w **in** dataBlock
    keyvaluepair(w,1)
function reduce(keyvaluepairs)
group keyvaluepairs

Based on the values in the summary data, the most commonly found words and categories are taken into consideration. With these data, the users' weighted areas of interest and instantaneous orientations are detected. After this process, the selection process begins to identify the ads in the ad set that suits most to expectations.

4.4 Heuristic pruning method (HPM)

HPM is an elimination approach intended to determine in an exploratory way the most suitable candidates out of the contents marked with one or several tags and which fulfil some specific criteria. At the beginning stage all elements of the data mass are inside the solution population, whereas throughout the iterative process the individuals that do not fulfil the criteria are eliminated.

In the advertising database that is used, contents are located under certain categories. There can be multiple categories that cover an ad. In addition to category tags, each ad is marked with one or several additional keywords. These keywords allow for each process to be distinguished independently after the categorical elimination.

HPM is carried out in two stages, which are the categorical elimination, and the keyword-based elimination. Firstly, the ads that are located under the correct category are found, and then the suitable contents are identified based on keywords. After these two stages, the ultimate solution is achieved (Fig. 3).

When we analyse the shares of Twitter users, one or several areas of interest are found for the user. These areas of interest match with the keywords that categorize the ads. The percentage weights are calculated based on the repetition frequency of the words that specify the user's area of interest. Out of the contents in the advertisement set each one similar to the users' area of interest is scored at the rate of the percentage weight of each area of interest. The requested number of solutions which have the highest score are placed into the "probable solutions" set; the individuals that are ranked below the specified score limit are eliminated; whereas the ones that remain in the "probable solutions" and the eliminated individuals are placed into the "semi-probable solutions" set. Due to the reason that an ad can fall within several categories, an ad that is placed to the "semi-probable solutions" set can be included to the "probable solutions" set. This procedure is re-conducted for each category, and individuals in the quantity determined in each stage are transferred to the next step.
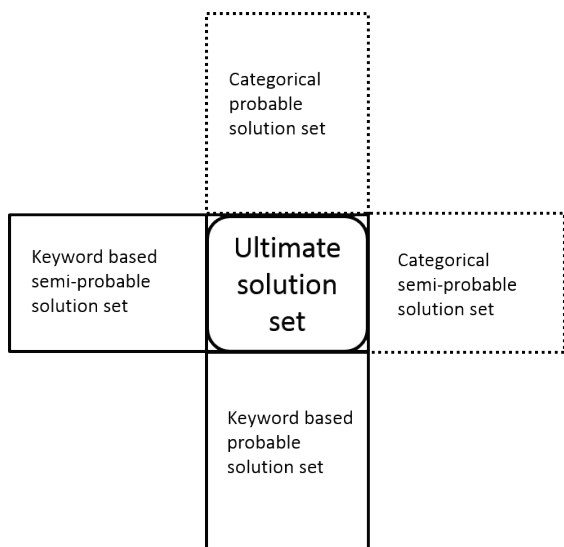

Figure 3 HPM result set

Pseudo code showing the categorical elimination process is as follows:

```
for each category w in categoryList
    for each advertisement in advertisementList
    if advertisement is in category
        addcategoryScoretoadvertisementScore
    orderByAdvertisementScore
    add Top N toprobableSolutionList
    eliminate under specified scpore
select Top N Advertisement
```

Other stage is the process where the keyword-based selection procedure is carried out. Percentage weights are calculated in accordance with the frequencies of the words in the shares of the analysis' subject user. The ad contents are scored by analysing the similarity between the word list attained as a result of the analysis, and the

keyword marked by the ad content. Specific number of individuals who have the highest score are placed to the "probable solutions" set, the ones who fall below a certain score are eliminated, and the ones that fall in the middle are placed to the "semi-probable solutions" set. The process is repeated for each word.

After this two-stage elimination process, the individuals with the highest scores make up the ultimate solution set. Out of this set, requested number of ads are presented as output.

After the categorical elimination and keyword based elimination procedures, the ad's score value, which is calculated by the overlapping of the category and word scores, and which is carried by each ad content, makes up the ad's impact factor (Eq. (1)).

$$ImpactFactor = CategoryScore \cap WordScore \qquad (1)$$

With regard to the ads to be presented to the user, the impact factor is used for determining any ad's display priority or frequency compared to the mentioned ads.

5   Sample application of the system

The developed advertising recommendation system has been tested and integrated with a developed web application (Fig. 4).This application can analyse a Twitter user's profile with his/her screen name.
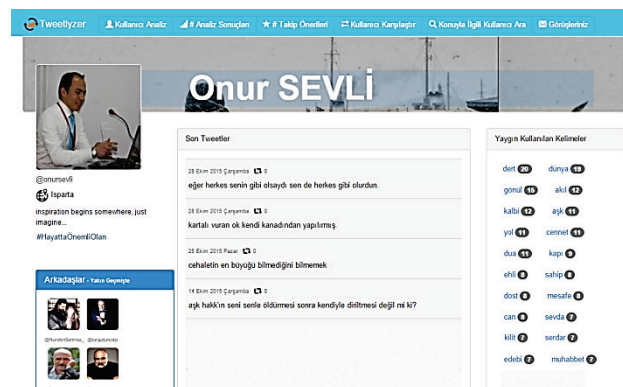

Figure 4 Sample web application

42,65 % emotion

30,88 % music

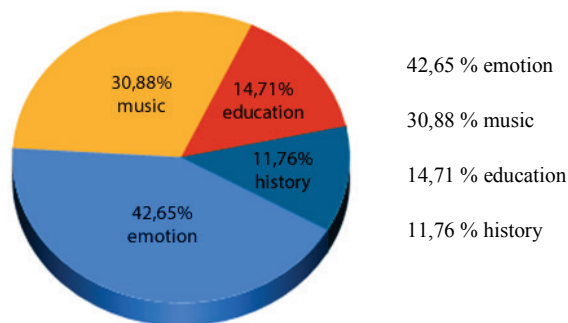14,71 % education

11,76 % history

Figure 5 Percentage distribution graph of user's area of interest

Web application presents the screen name information as a parameter to the web service which makes the data analysis. The analysing service obtains the relevant user's data from Twitter with the screen name parameter. The percentage distribution graph in relation to the identified interest areas is shown in Fig. 5.

The ads that are found to be appropriate for the user as a result of the filtering performed on the ad database, based on the frequently used words and identified areas of interest, are presented to the user. Impact factor values are obtained as a result of the categorical and keyword-based scoring of the determined ads. At the last stage, the data in relation to the ads are serialized in XML form and they are presented to the application as XML output. Sample output is as follows:

```
<ads>
<advertisement>
<description>YeniBirAşk (A New Love), Ferzende Mora, 20 TL</description>
<multimediaurl>someurl.com/img/yenibirask.png</multimediaurl>
<impactfactor>91</impactfactor>
</advertisement>
<advertisement>
<description>GünaydınNağmeleri (Morning Melodies), SezginAtılay, 13 TL</description>
<multimediaurl>someurl.com/img/gunaydinnagme.png</multimediaurl>
<impactfactor>54</impactfactor>
</advertisement>
<advertisement>
<description>Hayat BoyuÖğrenelim (Lifelong Learning), ErdemDizdar, 18 TL</description>
<multimediaurl>someurl.com/img/hbogrenelim.png</multimediaurl>
<impactfactor>38</impactfactor>
</advertisement>
<advertisement>
<description>GeçmişteBirFincanKahve (A Cup of Coffee In The Past), Tarık Efgan, 35 TL</description><multimediaurl>someurl.com/img/gecmistekahve.png</multimediaurl>
<impactfactor>29</impactfactor>
</advertisement>
</ads>
```

The ad's display frequency or order can be determined according to the impact factor.

With regard to the findings of the application based on user analysis, and with regard to the ads presented to the user, feedbacks are received from users online and through face to face meetings in order to evaluate the system's success.

6 Evaluation of the system

The developed system was tested with 563 active Twitter users for a period of five months. The feedbacks from these users were received on a regular basis and the users were asked to rate the system online.

The success of the system has been questioned on three main criteria. These are the success in identifying the user field of interest, the ad's suitability for the user's field of interest, and the success in responding to the sudden changes of user's field of interest. The average success values of the system which were obtained in accordance with the user feedbacks during the midterm and at the end of the implementation period are given in Tab. 1 below.

Table 1 System performance values

	Success percentage (First)	Success percentage (Last)	The Average
Success in identifying the user's area of interest	85,30	89,20	87,25
Ads suitability to area of interest	81,80	83,60	82,70
The success in responding to sudden changes	93,40	96,70	95,05
The Average	86,80	89,80	88,30

Although the system shows high success about conformity of the advertisements to the field of interest, by the expansion of advertising database, the marking of the advertisements with key words and advertisement being under the right category it is observed that the success rate can be increased. With the feedbacks received from users, enhancements on the database have been made. As a consequence, the success in the delivery of the advertisements in line with the user's field of interest has increased.

With regard to the success of the system in responding to sudden changes due to the reason that the changes in the user field of interest at the first stage are not at the sufficient level, the ratio obtained is lower than the final evaluation. In the final evaluation the users have been able to see more clearly that the system can respond to the changes in the user field of interest in the course of time.

One of the factors that negatively affect the system's success is due to the fact that Turkish words have more than one meaning, thus falling into more than one category. As a solution to this problem a holistic meaning analysis of the sentence should be carried out. However, this analysis can become a complicated task because of the structure of Turkish sentence. Due to the reason that such an analysis does not exist in the developed system, although being rare, the problem in which a word is matched with more than one category is encountered.

Today, in social networking problems may occur due to abbreviations being widely used or because of the wrong spellings due to differences of expression. The developed system has the ability to produce recommendations for incorrect spelling of words, and it has the ability to make corrections, but sometimes suggestions cannot be found. In some cases, numerous recommendations may be present but it cannot be determined which one of these is appropriate. If the spelling of the word which is crucial in identifying the user's field of interest is wrong, it may adversely affect the success of the system. This situation is reflected to the other stages.

The difficulties encountered in the system's operation stem from the structure of the language. Due to the reason that a meaning analysis for Turkish language has not been fully performed yet in computer environment, the desired success was not achieved in the results obtained from the language processing procedure.

In line with user feedback, the average success of the system in the initial evaluation is 86,8 %, whereas in the final evaluation the value increased by 89,8 %. The

improvements made to the system have a significant impact on this increase, as well as the development of ad and word-category database. The system's average success is 88,3 %, and it is expected to increase over time.

7 Conclusion

Social networks are the ideal environments for advertising publishing in terms of fast and economic access to many users. If the right ads are presented to the right user, the expected increase in sales can be achieved. Currently, the social networks try to find out the users areas of interest through the personal information they request in certain intervals. In this way, they present content recommendation through the static profiles they have drawn for the users. However, it must be taken into consideration that user expectations can change instantly. Therefore, users' expectations must be determined dynamically and accurately. Social network shares contain information regarding users' expectations and orientations. The identification of this information facilitates the presentation of ads that suit to the expectations of users.

Twitter is the fastest growing social network of recent times; however, it is new in terms of becoming an advertising medium. Due to this reason it is an environment open for study.

In this study, an advertising recommendation system has been developed. Shares of a group of user who make sharing in Turkish have been processed in order to make them structural and semantic. With this system based on dynamic data analysis, the word patterns in user's shares are identified and sorted under categories. In line with these categories, the user's areas of interest are identified. Based on the identified areas of interest, the ads addressing that area of interest are presented to the relevant user. A content filtering method called Heuristic Pruning Method (HPM) has been revealed in order to filter the most advertising content that suits best to expectations.

A prototype system has been developed to test the application. This application has been tested on a voluntary group of users actively using Twitter. Based on the feedback received from the users, it has been concluded that the system is successful at the rate of 88%.

The presentation of ads in line with the expectations, by correctly analysing the users' instantaneous expectations, makes the marketing strategy more effective. In this way, users' exposure to unnecessary contents is avoided and it is ensured that the content is delivered to the right audience. It is obvious that this method will yield to increases in marketing activities.

Acknowledgements

8 References

[1] Gürsoy, T. Reklam Terimlerive Kavramları Sözlüğü. Adam Yayınları, İstanbul, 1999.

[2] Martha, V.; Zhao, W.; Xu, X. A study on Twitter user-follower network a network based analysis. // IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM) / Canada, 2013, pp. 1405-1409.

[3] Hannon, J.; Bennett, M.; Smyth, B. Recommending twitter users to follow using content and collaborative filtering approaches. // The fourth ACM conference on Recommender systems / Barcelona, 2010, pp. 199-206.

[4] Yu, Z.; Wang, C.; Bu, J.; Wang, X.; Wu, Y.; Chen, C. Friend recommendation with content spread enhancement in social networks. // Information Sciences. 309, (2015), pp. 102-118. DOI: 10.1016/j.ins.2015.03.012

[5] Kwak, H.; Lee, C.; Park, H.; Moon, S. What is twitter, a social network or a news media? // 19th international conference on World Wide Web / USA, 2010, pp. 591-600.

[6] Jones, J. Switching in Twitter's Hashtagged Exchanges. // Journal of Business and Technical Communication. 28, 1(2014), pp. 83-108. DOI: 10.1177/1050651913502358

[7] Vu, X. T.; Abel, M. H.; Morizet-Mahoudeaux, P. A user-centered and group-based approach for social data filtering and sharing. // Computers in Human Behaviour. 51, (2015), pp. 1012-1023. DOI: 10.1016/j.chb.2014.11.079

[8] Wu, L.; Luo, N. Social streams recommendation in sina microblog with relation of user and interest. // IEEE International Conference on Information Science and Technology (ICIST) / Shenzhen, 2014, pp. 480-483. DOI: 10.1109/icist.2014.6920521

[9] Puglisi, S.; Parra-Arnau, J.; Jordi, F.; Rebollo-Monedero, D. On content-based recommendation and user privacy in social-tagging systems. // Computer Standards & Interfaces. 41(2015), pp. 17-27. DOI: 10.1016/j.csi.2015.01.004

[10] Bosch, H.; Thom, D.; Heimerl, F.; Puttmann, E.; Koch, S.; Kruger, R.; Woerner, M.; Ertl, T. ScatterBlogs2: Real-Time Monitoring of Microblog Messages Through User-Guided Filtering. // IEEE Transactions on Visualizatıon and Computer Graphics. 19, 12(2013), pp. 2022-2031. DOI: 10.1109/TVCG.2013.186

[11] Cataldi, M.; Ballatore, A.; Tiddi, I.; Aufaure, M. A. Good location, terrible food: detecting feature sentiment in user-generated reviews. // Social Network Analysis and Mining. 3, 4(2013), pp. 1149-1163. DOI: 10.1007/s13278-013-0119-7

[12] Kang, D.; Han, D; Park, N.; Kim, S.; Kang, U.; Lee, S. Eventera: Real-Time Event Recommendation System from Massive Heterogeneous Online Media. // IEEE International Conference on Data Mining Workshop / Shenzen, 2014, pp. 1211-1214.

[13] Yang, M. C.; Rim, H. C. Identifying interesting Twitter contents using topical analysis. // Expert Systems with Applications. 41, 9(2014), pp. 4330-4336. DOI: 10.1016/j.eswa.2013.12.051

[14] Chorley, M. J.; Colombo, G. B.; Allen, S. M.; Whitaker, R. M. Human content filtering in Twitter: The influence of metadata. // Int. J. Human-Computer Studies. 74, (2015), pp. 32-40. DOI: 10.1016/j.ijhcs.2014.10.001

[15] Dai, W.; Han, D.; Dai, Y.; Xu, D. Emotion Recognition and Affective Computing on Vocal Social Media. // Information & Management. 52, 7(2015), pp. 777-788. DOI: 10.1016/j.im.2015.02.003

[16] Liang, B.; Liu, Y.; Zhang, M.; Ma, S.; Ru, L.; Zhang, K. Searching for people to follow in social networks. // Expert Systems with Applications. 41, 16(2014), pp. 7455-7465. DOI: 10.1016/j.eswa.2014.05.050

[17] Armentano, M.; Godoy, D.; Amandi, A. Topology-based recommendation of users in micro-blogging communities. // Journal of Computer Science and Technology. 27, 3(2012), pp. 624-634. DOI: 10.1007/s11390-012-1249-5

[18] Li, J.; Ma, S.; Hong, S. Recommendation on social network based on graph model. // 31st Chinese Control Conference (CCC) / Hefei, 2012, pp. 7548-7551.

[19] Sudo, K.; Nagasaka, S.; Kobayashi, K.; Taniguchi, T.; Takano, T. Encouraging User Interaction of Social Network through Tweet Recommendation Using Community Structure. // Technologies and Applications of Artificial Intelligence (TAAI) / Taipei, 2013, pp. 300-305.

[20] Jamil, N.; Alhadi, A. C.; Noah, S. A. A collaborative names recommendation in the Twitter environment based on location. // International Conference on Semantic Technology and Information Retrieval (STAIR) / Putrajaya, 2011, pp. 119-124. DOI: 10.1109/stair.2011.5995775

[21] Zhou, X.; Wu, S.; Chen, C.; Chen, G.; Ying, S. Real-time recommendation for microblogs. // Information Sciences. 279, (2014), pp. 301-325. DOI: 10.1016/j.ins.2014.03.121

[22] Kim, Y.; Shim, K. TWILITE: A recommendation system for Twitter using a Probabilistic model based on latent Dirichlet allocation. // Information Systems, 42(2014), pp. 59-77. DOI: 10.1016/j.is.2013.11.003

[23] Islam, M.; Ding, C.; Chi, C. Personalized Recommender System on Whom to Follow in Twitter. // IEEE Fourth International Conference on Big Data and Cloud Computing (BdCloud) / Sydney, 2014, pp. 326-333. DOI: 10.1109/bdcloud.2014.84

[24] Lee, W. J.; Oh, K. J.; Lim, C. G.; Choi, H. J. User profile extraction from Twitter for personalized news recommendation. // 16th International Conference on Advanced Communication Technology (ICACT) / Pyeongchang, 2014, pp. 779-783. DOI: 10.1109/ICACT.2014.6779068

[25] Lu, H.; Lee, C. A Twitter Hashtag Recommendation Model that Accommodates for Temporal Clustering Effects. // Intelligent Systems. 30, 3(2015), pp. 18-25. DOI: 10.1109/MIS.2015.20

[26] Jonnalagedda, N.; Gauch, S. Personalized News Recommendation Using Twitter. // IEEE/WIC/ACM International Joint Conferences on Web Intelligence (WI) and Intelligent Agent Technologies (IAT) / Atlanta, 2013, pp. 21-25.

[27] Vos, I. F.; Verbeke, S. Sense and nonsense of social media in failure analysis. // Engineerıng Failure Analysis. 35, (2013), pp. 413-417. DOI: 10.1016/j.engfailanal.2013.03.023

[28] D'Avanzo, E.; Pilato, G. Mining social network users opinions' to aid buyers' shopping decisions. // Computers in Human Behavior. 51, (2015), pp. 1284-1294. DOI: 10.1016/j.chb.2014.11.081

[29] Lee, K.; Caverlee, J.; Cheng, Z. Y.; Sui, D. Z. Campaign Extraction from Social Media. // Acm Transactions On Intelligent Systems And Technology (TIST). 5, 1(2013). DOI: 10.1145/2542182.2542191

[30] Spina, D.; Gonzalo, J.; Amigo, E. Discovering filter keywords for company name disambiguation in twitter. // Expert Systems with Applications. 40, 12(2013), pp. 4986-5003. DOI: 10.1016/j.eswa.2013.03.001

[31] Ghiassi, M.; Skinner, J.; Zimbra, D. Twitter brand sentiment analysis: A hybrid system using n-gram analysis and dynamic artificial neural network. // Expert Systems with Applications. 40, 16(2013), pp. 6266-6282. DOI: 10.1016/j.eswa.2013.05.057

[32] Nettelhorst, S. C.; Jeter, W. K.; Brannon, L. A. Be careful what you wish for: The impact of advertisement choice on viewers' expectations. // Computers in Human Behavior. 41, (2014), pp. 313-318. DOI: 10.1016/j.chb.2014.10.009

[33] Burkhalter, J. N.; Wood, N. T.; Tryce, S. A. Clear, conspicuous, and concise: Disclosures and Twitter word-of-mouth. // Business Horizons. 57, 3(2014), pp. 319-328. DOI: 10.1016/j.bushor.2014.02.001

Authors' addresses

Onur Sevli
Mehmet Akif Ersoy University,
15030 Burdur, Turkey
Phone: +902482134130
E-mail: onursevli@mehmetakif.edu.tr

Ecir Uğur Küçüksille
Süleyman Demirel University,
32260 Isparta, Turkey
Phone: +902462111478
E-mail: ecirkucuksille@sdu.edu.tr