

Profiled M-advertising Framework for Consumer Advertising

G.L.P.N. Gangabada^{1*} and S. C. Premaratne²

^{1,2}*Faculty of Information Technology, University of Moratuwa, Sri Lanka*

Email: ¹glprageethnuwan@outlook.com, ²samindap@uom.lk

Abstract

Mobile advertising and profiled advertising has become major discussion topics recently. With rapid development of mobile technology in the last decade, many e-business organizations are trying to combine the latest mobile technologies with customer profiling in order to increase the reachability of their product information among the target market. This paper presents a potential mechanism to enhance such mobile advertising by improving existing user profile methods. The main goal of this research is to introduce a novel user profiling mechanism that can be used to improve the productivity of a mobile advertising system. The proposed system is capable of providing additional information such as video, audio, etc. about advertisements to the user while giving profile-based and proximity suggestions.

Keywords: Profiled m-marketing, collaborative filtering, user profiling

1. INTRODUCTION

Mobile advertising is one of the major focuses of today's consumer advertising. The main purpose of mobile advertising is to provide a more user friendly and efficient message delivery system for consumers digitally without any limitations on access location and the time. Mobile advertising includes the use of multimedia such as text, graphics, animation, audio and video. Although several approaches have been introduced to adopt mobile technologies in the field of advertising, most of them provide limited functionality with basic requirements such as linking an advertisement with a video hosted online [9].

Even though many advertising media are available, paper-based advertising is still remains as the media through which a majority of the public can be addressed as they spend more time with printed media such as newspapers, magazines, etc., than other digital media. Even with advertising with printed media, there are limitations such as the amount of information that can be displayed, descriptiveness of the product or service, attractiveness, etc.

For an example, descriptive information about a property sale may not be present in a printed advertisement. People who are interested in such an advertisement would have to visit the place and examine the property. Sometimes, people may need to know more information about a particular advertised product or service, but they may not tend to log in to a computer and browse the Internet for information. As paper based advertisements are posted targeting a general customer segment, individual user profiling will not take place.

This paper focuses on such an attempt to use data mining and mobile technologies to bring a hybrid solution for printed media-based advertising to make it more efficient and effective. The ability to provide user-based and proximity-based suggestions will efficiently

increase the usability and effectiveness of the system thereby, ultimately reducing the searching time and effort spent by the user.

2. OVERVIEW

Mobile advertising mechanisms that are presently available can be divided into two segments considering the type of the mobile devices being targeted. The common mechanism is the use of pre-recorded voice messages and short message service. Service providers frequently use these mechanisms to advertise their new features, promotions and services. In the past decade, with rapid popularization of smart mobile devices, mobile advertising introduced new features into smart phones such as the ability to connect to the Internet, location detection, camera etc. [8-10].

One of the most common usages of mobile advertising along with printed media is the quick response (QR) code which is a type matrix barcode that can be used to embed information in an image [1]. When a particular advertisement is printed using a QR code, it encodes a URL to a product or server's web site or an online video clip which that provides more descriptive information about the particular product or service [9]. The user is able to visit the site by scanning the QR code embedded in the advertisement provided that a QR decoding application is installed in the device where in most cases, such software is pre-installed or is available for free.

In recent literature, concepts such as augmented reality have been used in advertising [6] where users can scan a particular object in the advertisement and view more information about that particular product or service via augmented reality based applications.

As described above, the methods available simply use the mobile devices as a tool for providing access to an online resource where in most cases, is a video or a web site. As current implementations do not include more than one resource of information a user may still need to search for the product or service to grasp a better understanding [2]. Moreover, the inability to find out which users are interested in the advertised product or service is a major drawback. Many online e-business applications such as e-bay and Amazon used to track user activities in order to provide a more effective service to their user base [5, 7].

User specific suggestions about products and services are not supported in existing implementations of mobile advertising [9, 12]. Therefore, user specific advertising will certainly result in more customer loyalty and will increase the effectiveness of both the advertiser and the consumer.

3. PROFILED M-ADVERTISING FRAMEWORK

The introduction of a novel hybrid solution which allows users to obtain more information about a particular paper-based advertisement is considered important when carrying out intended message to the public in a more efficient way. The proposed solution is a hybrid framework that integrates mobile technology and customer profiling to provide a more user friendly mechanism to communicate messages to the public.

By tracking the behavior of individual user activities and utilizing a data mining algorithm and collaborative filtering, a unique profile can be provided for the each person (a hybrid data mining algorithm needs to be generated that matches the present context using existing algorithms). Furthermore, users are allowed to rate and comment on a particular product or service that is being viewed where the ratings and comments are used to enhance effectiveness of the customer profiling algorithm.

4. SYSTEM ARCHITECTURE

The proposed system consists of several sub modules that fulfill different functionalities expected by the entire system. The main components of the proposed architecture are mobile client application, communication service, application server, file server, database server and recommender system.

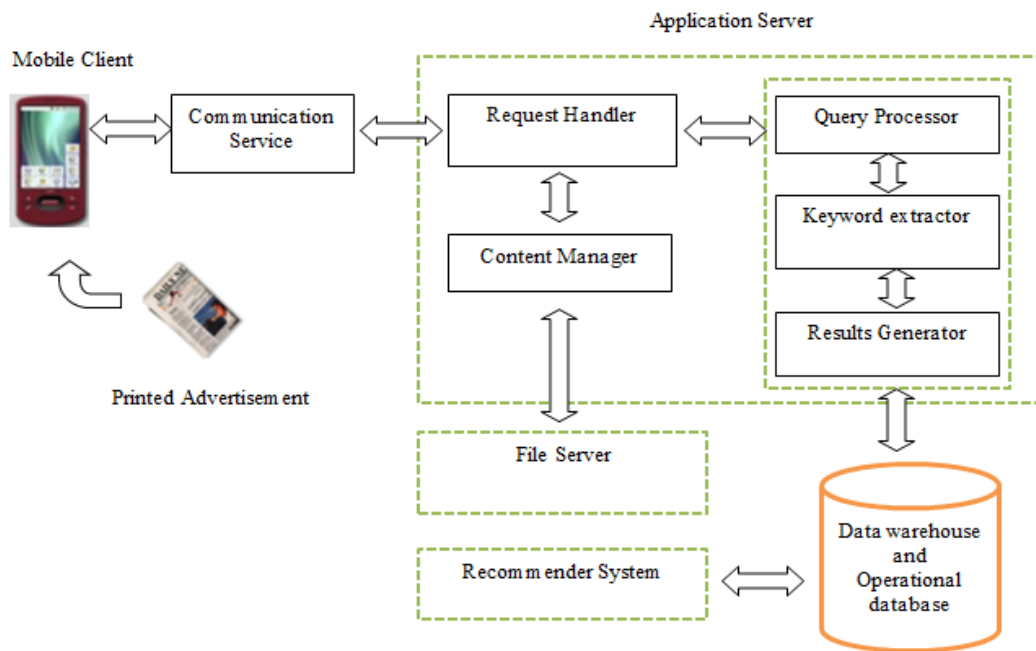


FIGURE 1: High level architecture of the proposed system.

Fig. 1 shows the high level architecture of the proposed system. A consumer can either scan a QR code embedded in the advertisement or enter a search query as the input via a mobile application which is transmitted to the server via a communication service. The application service includes different modules to perform different activities. The request handler processes the request sent by the mobile device and initiates a content manager and query handler where the content manager connects with a database server to retrieve the location information of the particular multimedia object for the advertisement requested. The query processor initiates the process of retrieving suggestions for the user where suggestions are generated by the recommender system with the aid of the data warehouse associated with the system. This allows both the multimedia content and the suggestions to be transmitted back to the user.

5. RECOMMENDER ALGORITHM

The proposed algorithm is a hybrid mechanism of item based recommendation and user based recommendation. The goal of the recommender algorithm is to suggest new items or predict the utility of a certain item for a particular user based on the user's previous preferences and the opinions of other like-minded users.

In a typical scenario, there is a list of m users $U = \{u_1, u_2, \dots, u_m\}$ and a list of n items $I = \{i_1, i_2, \dots, i_n\}$. Each user u_i has a list of items I_{ui} , to which the user has expressed his/her opinion about. In this scenario, an opinion is derived from analyzing user view records. Note that $I_{ui} \subseteq I$ and it is possible for I_{ui} to be a null-set. A distinguished user $u_a \in U$ is referred to as the active user for whom the task of the algorithm is to find an item likeliness that can be of two forms as follows:

- **Prediction** is a numerical value, $P_{a,j}$, expressing the predicted likeliness of item $i_j \notin I_{ua}$ for the active user u_a . This predicted value is within the same scale.
- **Recommendation** is a list of N items, $I_r \subset I$, that the active user prefers the most. Note that the recommended list must be on items not already viewed by the active user [4].

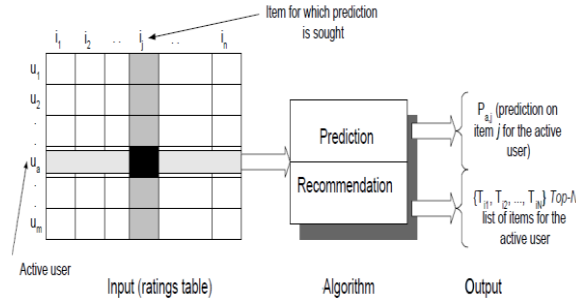


FIGURE 2: Suggestion generating process.

Fig. 2 shows the schematic diagram of suggestion generating process. This process represents the entire $m \times n$ user-item data as a ratings matrix. Each entry in the matrix represents the preference score of the i^{th} user on j^{th} item. Each individual score is within a numerical scale and it can as well be 0 indicating that the user has not yet viewed that item.

This principle process will be used in both user-based recommendation generation and the item-based recommendation. Addressing issues may occur when suggesting items for new users and adopting items that are newly added or rarely viewed.

User oriented recommendation generating process

Step 1: Find the k -most liked users in the system and store all the user similarities.

This is a critical step in the algorithm where the basic idea is in similarity computation between two users, u_i and u_j where we first isolate the items viewed by both the users and then apply a similarity computation technique to determine the similarity. In this algorithm, a cosine-based similarity [13] is used. Two users are considered as two vectors in an m -dimensional item-space. The similarity between the two users is measured by computing the cosine of the angle between these two vectors. Formally, in an $m \times n$ matrix as in Fig. 3, the similarity between users i and j , $sim(i, j)$, is denoted by,

$$sim(i, j) = \cos(\vec{i}, \vec{j}) = \frac{\vec{i} \cdot \vec{j}}{\|\vec{i}\|_2 * \|\vec{j}\|_2} \quad (1)$$

where “.” denotes the dot-product of the two vectors[11].

Step 2: Identify similar users to the active user

U_2 and U_9 are similar to U_8 who is the active user as in the Figure 3.

Step 3 : Identify the items not yet viewed by the user

I_1 and I_m have not been viewed by the active user U_8 as shown in the Figure 3.

Step 4 : Recommendation generation

A prediction is made by calculating the weighted sum on I_1 and I_m as shown in Fig. 4. The most important step in this process is to generate the output interface in terms of the prediction. Once the set of most similar users are isolated based on the similarity measures, the next step is to look into views of target users and use a technique to obtain predictions.

The process is as follows,

- 1] Find k most similar users to u , save to U_{sim}
- 2] Get all items purchased by U_{sim} , save to $I_{candidate}$

- 3] Remove unavailable items in $I_{candidate}$
- 5] Get all items purchased by u , save to $I_{purchased}$
- 6] Take $I_{candidate} - I_{purchased} = I_{recmd}$
 Re-order items in I_{recmd} based on sum of user-user similarity

This process is denoted as [3]:

$$pred(u, i) = \frac{\sum_{v \in k-similarUser(u)} userSim(u, v) * r_{vi}}{\sum_{v \in k-similarUser(u)} userSim(u, v)} \quad (2)$$

		users												
		1	2	3	4	5	6	7	8	9	10	...	n	
items	1	1	1	1			1				1			
	2							1	1	1				
	3	1	1		1				1	1				1
	4		1			1			1	1				
	⋮			1				1						
	m				1						1			

FIGURE 3: Finding similarity between users.

		users											
		1	2	3	4	5	6	7	8	9	10	...	n
items	1	1	1	1			1		0.7		1		
	2							1	1	1			
	3	1	1		1				1	1			1
	4		1			1			1	1			
	⋮			1				1					
	m				1				0.9	1			

FIGURE 4: Predicting by taking a weighted sum.

Note that P number of items will be generated from the first step of this algorithm.

Let us try to apply this process to above user/item matrix where cell value 1 means the user purchased the item.

- 1]. If we take user 8 we could see that user 2 and user 9 are the most like-minded users with respect to user 8, $U_{sim} = \{U_2, U_9\}$
- 2]. Consider items purchased by U_{sim} , $I_{candidate} = \{I_1, I_2, I_3, I_4, I_m\}$
- 3]. Consider items purchased by U_8 , $U_8 = \{I_2, I_3, I_4\}$
- 4]. Calculate $I_{candidate}$, $I_{candidate} = \{I_1, I_m\}$
- 5]. Calculate user-user similarity, $pred(u_8, i_1) = 0.7$, $pred(u_8, i_m) = 0.9$

Assuming single item will be recommended, $I_{recmd} = \{I_9\}$

Item oriented recommendation generating process

Step 1: Compute and store all the item-item similarity

The adjusted cosine similarity will be used to compute the similarity in this mechanism because the basic cosine measure in item oriented case has a major drawback in that the difference in rating scale between different users are not taken into account. The adjusted cosine similarity offsets this drawback by subtracting the corresponding user average from each co-rated pair. Formally, the similarity between items i and j using this scheme is given by [14],

$$sim(i, j) = \frac{\sum_{u \in U} (R_{u,i} - \bar{R}_u)(R_{u,j} - \bar{R}_u)}{\sum_{u \in U} (R_{u,i} - \bar{R}_u)^2 \sqrt{\sum_{u \in U} (R_{u,j} - \bar{R}_u)^2}} \quad (3)$$

where R_u is the average of the u^{th} user's ratings.

Step 2: Find N items that will be most likely purchased by user u .

- 1] Get all items purchased by u , save to $I_{purchased}$
- 2] For each item in $I_{purchased}$, find k most similar items, save them to $I_{candidate}$
- 3] Remove unavailable items in $I_{candidate}$
- 4] Get all items purchased by u , save to $I_{purchased}$
- 5] Take $I_{candidate} - I_{purchased} = I_{recmd}$
- 6] Re-order items in I_{recmd} based on [3],

$$pred(u, i) = \frac{\sum_{j \in purchased\ items(u)} itemSim(i, j) * r_{uj}}{\sum_{j \in purchased\ items(u)} itemSim(i, j)} \quad (4)$$

Note that Q number of items will be generated from this step of the algorithm.

Let us now try to illustrate the process with the following diagram,

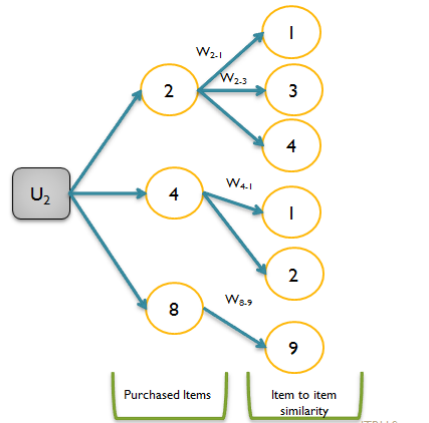


FIGURE 5: Illustrate Item based recommendation generation.

- 1]. All the items purchased by user 2, $I_{purchased} = \{I2, I4, I8\}$
- 2]. Then generate the $I_{candidate}$ with respect to $I_{purchased}$, $I_{candidate} = \{I1, I2, I9\}$
- 3]. Calculate I_{recmd} , $I_{recmd} = \{I1, I3, I9\}$

Feature vector oriented recommendation generating process

The problem with item oriented and user oriented recommendation process is that the items which are newly added or items that are rarely viewed and rated by users will not be

selected for the recommendation list even though they are largely relevant for a particular user and an item. By using a feature based recommendation process, this effect can be minimized.

This process is as follows:

Step 1: Get the P number of items generated at process 1 and Q number of items generated at process 2

Step 2: Eliminate the similar items and get the P+Q list of items

Step 3: Get the rarely viewed and newly added items and compare features with P+Q set of items and derive similar items.

R number of items will be generated from this process as shown in Fig. 5.

Features of the items

		Feature1	Feature2	Feature3	Feature m
Items from process 1	I _{p1}	1	1	0				0
	I _{p2}	0	1	0				1
+ Items from process 2							
+ Rarely viewed items	I _{q1}	0	0	1				1
	I _{q2}	0	0	1				0
.....								
	I ₁	0	0	1				0
	I ₂	1	1	1				0
.....								

FIGURE 6: Feature oriented similarity matrix.

Prediction of entire list (P + Q + R)

The final stage of the algorithm is to re-order the entire list of suggestions according to a weight as given below:

Step 1: Generation of Z, the number of suggestions intended to provide for the user. Z will be denoted by,

$$Z = \frac{(P*Z)+(Q*Z)+(R*Z)}{(P+Q+R)} \tag{5}$$

where $Z \leq (P+Q+R)$. If the number of suggestions selected from each process is p, q and r,

$$p = (P \times Z)/(P + Q + R) \tag{6}$$

$$q = (Q \times Z)/(P + Q + R) \tag{7}$$

$$r = (R \times Z)/(P + Q + R) \tag{8}$$

Step 2: Predict the items in combined list of P+Q using [3],

$$pred(u, i) = \frac{\sum_{j \in purc \text{ hasedItems } (u)} itemSim(i, j) * r_{uj}}{\sum_{j \in purc \text{ hasedItems } (u)} itemSim(i, j)} \tag{9}$$

Step 3: Insert items from list R to the ordered list considering the feature similarity value denoted in process 3 and generate the final list of suggestions for the active user.

6. RESULTS & DISCUSSION

The proposed algorithm provides a set of recommendations to a user who views a particular advertisement by considering both the user's history of views and the characteristics of the item. As such, these two dimensions need to be evaluated separately before evaluating the combined algorithm. A dataset of 68 users and 150 items from the data warehouse were selected randomly for the evaluation process. The evaluation was carried out considering the user feedbacks for the suggestions generated by the algorithm for these items.

Initially, the user-based recommendation generation process was evaluated depending on the feedbacks provided by the users. The accuracy of the algorithm depends on the view count of the users where this accuracy was higher for the users who have viewed more items. Fig. 6 illustrates the results of this evaluation.

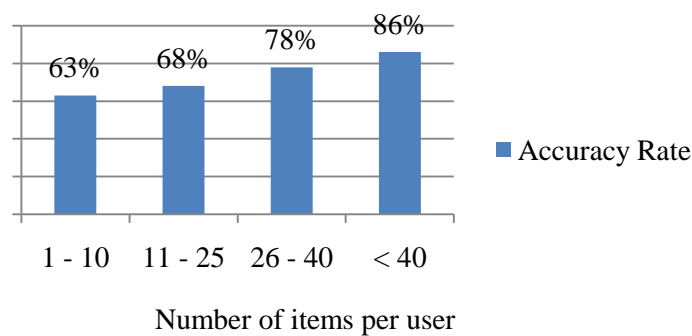


FIGURE 7: Accuracy of the user-oriented recommendation algorithm.

The next stage of the evaluation process is to evaluate the item-based recommendation generation where the accuracy is dependent on the number of users who view each item. As depicted in Fig. 7, the accuracy was higher when a particular item is viewed by many users.

Finally, the combined algorithm was evaluated using the same dataset with respect to the number of items viewed by users. We found that the accuracy rate of the final algorithm is dependent on the number of items the users view and the number of users who have viewed each item. As shown in Fig. 8, this accuracy is high for the user who views more items.

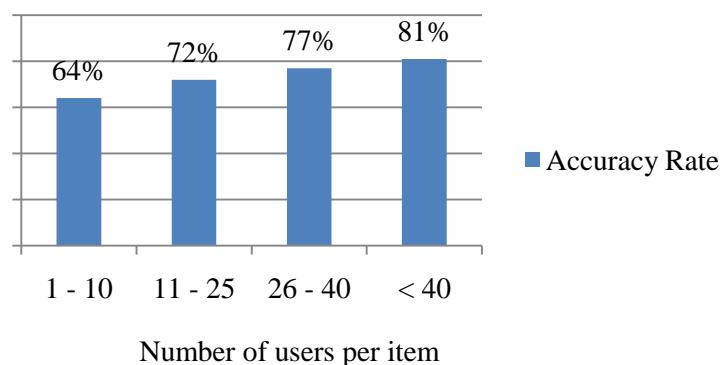


FIGURE 8: Accuracy of the item-oriented recommendation algorithm.

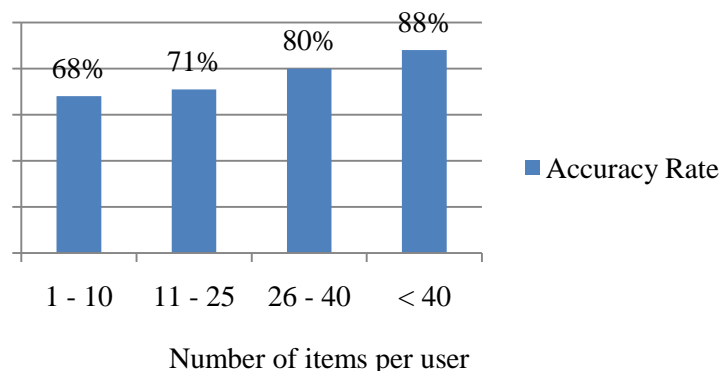


FIGURE 9: Accuracy of the combined recommendation algorithm with respect to item views per user.

The error rate of suggestions generated by the proposed algorithm somewhat depends on the number of views it has obtained. This could be minimized by increasing the percentage of items selected by feature vector oriented recommendation generating process but it could reduce the user relevance of the results.

7. CONSLUSION & FUTURE WORK

In this paper, we present a new mobile advertising framework with user recommendations as existing implementations of mobile advertising are limited to very brief services. In this research, we concentrated mainly on providing additional information about a particular product or service in multimedia format and giving suggestions about similar item a user views.

Advertisers are allowed to store multimedia contents such as video, audio, animation, graphic, etc. about advertisements published in printed publications. Consumers can use a mobile application to scan advertisements and retrieve additional information provided by the advertiser and recommendation provided by the system will reduce the time and effect that the user has to put on searching for similar items. The accuracy of the algorithm can be further increased if the system is tested with a large number of users. Further research is carried out to generate a more accurate recommender algorithm. Proximity suggestions provided by the system should be personalized according to the profiles of the user as a personalized proximity suggestion mechanism will definitely benefit both the user and the advertiser.

The main research work which is being carried out at present is to extend the inbuilt searching mechanism using semantics of the search queries provided by the user. User profiling and meta database will be used to improve the accuracy and relevancy of the search results.

8. REFERENCES

- [1] Bar code to 2D Code, <http://archive.is/20120915/http://www.qrcode.com/en/aboutqr.html>, 2010
- [2] Advantages and Disadvantages of QR code, <http://www.estateqr.com/advantages-disadvantages.html>, 2010
- [3] Lin, Recommendation engine demystified, neighborhood methods collaborative filtering, NYC predictive analytics, pp. 11, 2011.
- [4] Sarwar, G. Karypis, J. Konstan and J. Riedl, "Item-based Collaborative Filtering Algorithms", Technical Report ACM 1581133480/01/0005, Computer Science Dept., University of Minnesota., pp. 4, 2001.

- [5] Lee, S. Fung, "Online Social Networks and E-Commerce", student paper, TS085: Ethics and Law on the Electronic Frontier, MIT, pp. 8, 2007.
- [6] Blipper, "QR codes on steroids-augmented reality advertising", <http://www.digitaltimes.ie/2012/01/blipper>, 2012.
- [7] How to Build Trust and Improve the Shopping Experience, <http://knowwpcarey.com/article.cfm?cid=25&aid=1171>, 2012
- [8] L. Dubois, "How to Use Augmented Reality in Advertising", <http://www.inc.com/guides/201104/how-to-use-augmented-reality-in-advertising.html>, 2010
- [9] N. Ford, "Droid Does QR Coded in People Magazine", http://qranywhere.blogspot.com/2012_04_01_archive.html, 2010.
- [10] N. Ford, "Luxury Brand Wins with QR Codes", <http://qranywhere.blogspot.com/2010/04/luxury-brand-wins-with-qr-codes.html>, 2010.
- [11] P. Symeonidis, A. Nanopoulos, A.N. Papadopoulos and Y. Manolopoulos, "Collaborative Filtering: Fallacies and Insights in Measuring Similarity", Technical Report, Dept. of Informatics, Aristotle University of Thessaloniki, Greece, pp. 4-10, 2004.
- [12] P. Viswanathan, "Mobile Marketing – Pros and Cons of SMS Marketing", <http://mobiledevices.about.com/od/marketingapps/a/Mobile-Marketing-Pros-And-Cons-Of-Sms-Marketing.htm>, 2010.
- [13] N. Tan, M. Steinbach and V. Kumar, "Introduction to Data Mining", Addison-Wesley, ISBN 0-321-32136-7, 2005.
- [14] S. Zhu, J. Wu, H. Xiong and G. Xia, "Scaling up top-K cosine similarity search", Data and Knowledge Engineering, vol. 70, pp. 63-66, 2011.