Journal of Basic and Applied Engineering Research
Print ISSN: 2350-0077; Online ISSN: 2350-0255; Volume 1, Number 10; October, 2014 pp. 60-75

http://www.krishisanskriti.org/jbaer.html

Enhancement in Advertisement using Operational Research and Community Clustering

Vivek Prasad[1], Ujjwal Jain[2], Meghna Patiyawala[3]

[1,2,3]*Department of Computer Science and Engineering, Nirma University, Ahmedabad*

Abstract: **E – Commerce and M – Commerce have remained the fastest growing form of commerce compared to physical retail stores. More than half the Internet user population join an online social network, share photos and shop online. The next generation M – commerce and E – commerce is dominated by smart advertisement and publicity. In this paper, we propose an algorithm to cluster community, that can be implemented by classification based on their past behavior of their likes and dislikes on social network, and send advertisements to those clusters. The key idea of the algorithm is to note the consumer actions related to the videos on YouTube and the followings on Facebook and a community can be formed on the base on common likings of the customers. New members can be added to the community based on the link with the existing member of the community. The new member is added to the community if the similarities and the relation with the member is more than the threshold value. The algorithm gets the data of the user from the server side scripting and cookies maintained by the browser of the user. This information is used to relate the user with a specific product and also to other user that are linked with the user which is added already in the community. Concepts of OPT such as Minimum Spanning Tree (MST) and clustering algorithm, that forms the community while crawling the users and adding those which are relevant to the clusters that are being formed, are used. The relative advertisement could be shown to the targeted group and this can facilitate more precise connection between the advertisers and the consumers.**

Keywords: **E-commerce, M-commerce, Community Clustering, OPT, Minimum Spanning Tree.**

1. INTRODUCTION

E-commerce [18] and M-commerce [1] have been the fastest growing form when compared to physical retail stores, services and entertainment. More than half of the internet user joining an online social network contribute to social bookmarking sites, share photos and are involved in E-commerce and M-commerce.

As the online commercial business is increasing day by day, more business firms are getting digital and online advertisement has become an important aspect for the firms to increase sales. High revenues are spent on online advertisement so enhancement in advertisement has become an important issue for digital firms.

In this paper, we have proposed a way to enhance the advertisement [2] methods using the help of community clustering [9]. The users on the internet are increasing day by day. So there have been many approaches made by the companies to find optimal users which may be more appropriate than others for their products. The users can be found using community clustering [13].

A community can be defined as a group of "densely connected" people that are loosely connected to others [14]. The formation of this community using some algorithm is called community clustering.

In this paper, we propose an algorithm that can be useful to find the relevant customers for showing the online advertisement. Section 2 shows the related work which is undertaken by big commercial websites. The current measures taken by large firms are shown in this section. Section 3 shows the concepts used to implement the proposed algorithm. Concepts of co-citation [11] and bibliographic coupling [10] and modularity [14] are introduced. In section 4, our main algorithm is introduced, which can be used to form the clusters of the users. The algorithm uses the concepts of operational research such as Minimum Spanning Tree (MST) [7, 8] and the algorithm is similar to the incremental clustering algorithm [3, 4, 5, 6] to form a group of the customers. These customers are than shown the advertisement. The algorithm can be implemented on weighted graphs. Section 5 shows the algorithm which can be implemented to improve the cluster quality. In this algorithm, two identical clusters are combined and a more optimal cluster can be obtained. In section 6, we conclude the paper and outline the future work.

2. RELATED WORK

Today, the business has become much powerful as new strategies are being implemented by various firms using the newly developed improved technology. The competition level has been increased between the business firms. So new

methods are developed by the firms to increase the sales and generate more revenue.

M – Commerce [1, 2, 21, 22] and E – Commerce [18, 19, 20] involves the use of mobile phones and other devices using which the users perform transactions for buying something or some service. This field of business has been increasing with a great speed as it is easy to use and relatively faster for the users to buy anything.

With the growth of online business, the advertisement and marketing aspect has also become important for the companies and firms. Today, no company could survive without the use of technology and online business.

The advertisement field has remained an important aspect for the companies. The sales of the firms depend largely on the way of advertisement by the firms. Social networking sites have become a useful place for these advertisers. The firms usually post relevant advertisement on social networking sites to garner attention for their products through likes and shares. The growth of products via these means is directly proportional to the amount of social publicity. To explain the above statement the following example can be stated.

YouTube for instance previews advertisement as prequel to most of its video. These are Pay-Per-View videos which are means of promotions of the products. When any user likes those videos, our main protocol is to accumulate the data from those likes through cookies and spywares which are stored on the server and cluster those data. The cluster can be formed by the below mentioned algorithm.

From the cluster we can distinguish the individual's likes and dislikes and then showing them only those products he/she intends to use or wants to view. Here we are further optimizing the problem and create a community of likeminded individuals. YouTube can then show them only those advertisements which they are intended to view neglecting the rest.

Another instance of community cluster is Facebook where in likeminded people form a cluster and shares a page or advertisement or any other link that they think that it is useful for a particular group of people. This is another way of propagating a chain to increase the market value of the products and gaining higher social media points.

The communities that are formed in these ways are not optimal and have low quality. On Facebook, the clusters or the groups formed are on random base and no similarity is there among the members of the group. So in this paper, we have used a clustering algorithm, so that a more optimal and high quality cluster of the users can be obtained.

3. CLUSTERING CONCEPTS

The community of users can be formed by finding users that are related to each other. They have a similar quality among them.

The extent of relation between the users can be found by two metrics – co-citation [11] and bibliographic coupling [10].

The co-citation between two users $u1$ and $u2$ is defined as the number of users that are linked with both the users $u1$ and $u2$. The number of users which relate to both the users $u1$ and $u2$ is called the co-citation value of the users.

The bibliographic coupling concept is similar to the co-citation but it can be defined as the number of users to which both $u1$ and $u2$ relate. The number of users that both the users $u1$ and $u2$ link with is called the bibliographic coupling value of the users.

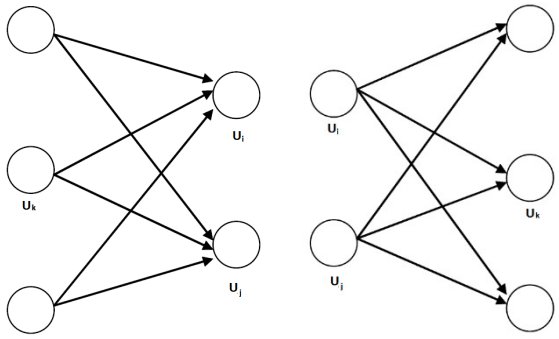

Fig. 1: Co-citation Figure 2 Bibliographic coupling

With the help of these two metrics, the relation between users for the addition in the tree can be obtained.

The quality of the cluster formed can be obtained by the criteria called modularity Q [14, 15, 23]. Modularity is a well-known criteria proposed specifically for the context of community mining. The modularity of any cluster can be found by following formula:

$$Q = \frac{1}{2E}\sum_{l=1}^{k}\sum_{i,j\in Cl}\left[Aij - \frac{\sum_j Aij \sum_i Aij}{2E}\right]$$

Modularity considers the difference between the fraction of edges that are within the community and the expected such fraction if the edges were randomly distributed. Here E denotes the number of edges in the network i.e.

$$E = \frac{1}{2}\sum_{ij} Aij$$

Aij denotes the weight between node ni and nj and C_l denotes the cluster l.

4. CLUSTERING OF USERS

Community clustering is forming a group of people on the basis on some common feature. A community can be defined as the group of "densely connected" individuals which are "loosely connected" with the rest [14]. A great number of clustering algorithms have been defined in the last few years. The basic approach of these algorithms is to divide the groups into sub-groups based on the network of interest. The "goodness" of the algorithm can be found by an objective known as modularity Q [15, 16, 17]. Several algorithms have been proposed for optimizing the modularity Q. There is no well – defined criterion for evaluating the result of the algorithms but the standard procedure is to compare the results of the algorithm with the actual optimal results.

The proposed algorithm is made using concepts of Operational Theory such as Minimum Spanning Tree (MST) and clustering algorithm. The clustering algorithm is similar to incremental clustering crawler algorithm used for clustering the relevant web pages. In this algorithm, the information of user is obtained from Medias like server side scripting and cookies that are maintained on the user browser. The users that are related with the previous users, which are already included in the tree, are also added to the tree. Then the algorithm assigns the weight on the relevance between the users. A spanning tree is formed by joining all the users as the vertices of the tree and the edges are assigned weight. The weight of the edges is inverse of the relation between the two users. The less the weight, the more is the similarity between the users. Then based on the weight of the edges, the vertices are removed from the graph having the highest weight first till the maximum number of required users are obtained or the minimum similarity criteria is reached.

Clustering algorithm :-

Input :-
Graph $G\ (V, E)$, where V – vertices of the graph representing the user and E – edges of the graph representing the relation between a single user with the other users.
K – Maximum number of required users
Wt. – weight threshold for the relation between the users.
V_o– The output set of users which is initially Ø.

Output:-

The graph with K users or with maximum weight of edge not more than Wt.

Begin

1. Find the user meeting the criteria with the product.

$Vo = Vo \cup \{ s \}$

2. Find some other user related to the user already in the tree.

$$Vi = \{ u \mid u \neq s \text{ and } (s, u) \in R \}$$

$$Vo = Vo \cup Vi$$

3. Repeat step 2 while all $u \in Vo$, for which $(s, u) \in R$.
4. Assign weights to all the edges of the graph depending on the relation between the users.
5. Remove edge $\{ w \}$ with the highest sum of weight in the graph.
6. Repeat step 5 until $K > |Vo|$ or there isa $w < wt$.
7. $Vo$is the output of the required group of users.

End

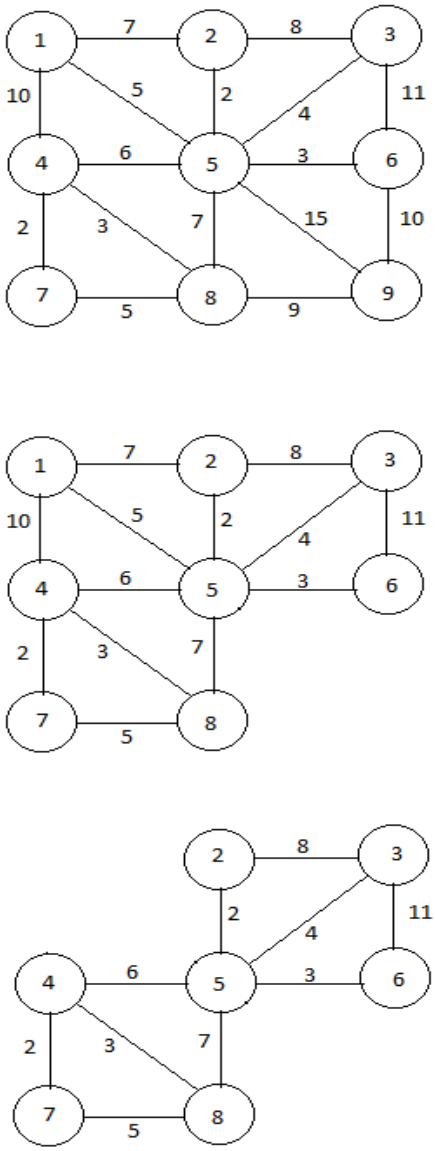

Fig. 3. Example of Minimum Spanning Tree

In this algorithm, initially the graph G is taken as the input. A user s, meeting the relevant criteria is added in the set Vo. An identical user u, related with the user s, is also added into the set. If a user s has some relation with user $u, (s, u) \in R$. R is the set of the related users. Then weight is assigned between all the users showing the level of relation, shown as the edges of the graph. Maximum weight is given if there is no relation between the users. Then the user w, with the maximum total weight is removed. This process is done until K number of users remain or the maximum weight threshold is met. The output is the graph with the optimal combination of users to whom the advertisement can be sent.

5. ALGORITHM TO ACHIEVE MORE OPTIMALITY

A large number of clusters can be formed by the above algorithm. The users on the internet are increasing day by day. The clusters formed from above algorithm are in large number and so it would be helpful to reduce the number of clusters and also increase the optimality of the clusters by merging the clusters formed.

There are users which can be found in many clusters formed by the previous algorithm. The clusters formed may have very less difference in the users included in them. These clusters can be merged to form a single cluster as the users in them may have similarity. Thus a larger cluster with more optimality can be obtained [12].

Algorithm:-

1. Start with two clusters $Vo1$ and $Vo2$.
2. Find a relational coefficient between clusters $Vo1$ and $Vo2$.
3. If the coefficient is larger than the threshold value than merge $Vo1$ and $Vo2$.
4. Repeat from 1 until all clusters have been considered.
5. Output the final cluster Vo of users which is more optimal.

In the preliminary cluster set, there exists situations that the two clusters share the same users which may be huge in numbers proportional to the size of these clusters. These users results in the clusters overlapping each other. This relational coefficient between two clusters in the above algorithm can be obtained by the following formula:-

$$Relational\ Coefficient = \frac{|\ Vo1\ \cap Vo2|}{\min\ \{\ Vo1, Vo2\ \}}$$

Where Vo1 represents the first cluster and Vo2 the second cluster. | Vo1 □ Vo2 | represents the number of users that are present in both the clusters and min { Vo1, Vo2 } means the minimum number of users from both the clusters.

Using this algorithm, the number of clusters formed at the end decreases and a more optimal cluster of users can be obtained.

6. CONCLUSION

Advertisers must identify an efficient means to communicate with consumers to increase revenue. Today, online methods are more commonly employed than offline means. In this paper, we proposed an effective algorithm which can be used to form the clusters of the users to which an online advertisement regarding a specific product can be sent and a second algorithm to improve the quality of the cluster. In this the users are first searched, put in the graph, are assigned the weights and then removed on the basis of the criteria and an optimal cluster of the users are formed. Generally high quality clusters can be formed.

The E-commerce and M-commerce have a great scope in future and this algorithm can be extended to increase the revenue and the sales of the firm. Thus, such methods involving the concepts of operational theory can be employed to build a better relationship between the customer and the firms.

REFERENCES

[1] Kalle Lyytinen, "M-commerce – Mobile commerce: a new Frontier for E – business", 2001.

[2] Jung woo Lee, Choongsik Lee and Yong suk Park, "Research on the Advertisement Effect of Push Type Mobile Advertisement", Yonsei University, South Korea.

[3] Gye-Jeong, Kyu-Young Whang, Min-Soo Kim, Hyo-Sang Lim and Ki-Hoon Lee, "An incremental Clustering Crawler for Community-Limited Search", KAIST, Yuseong-Gu, Daejeon, Korea.

[4] Kim, G., Kim, M., Kim, Y., and Whang, K., "Web Crawling and PageRank Calculation for Community-Limited Search" *Proc. Korea Computer Congress of KIISE* (in Korean), pp.1-3, July 2005.

[5] Kim, G., Whang, K., Kim, M., Lee, B., Lim, H., "Incremental Clustering Crawler for Community-Limited Search, " Tech. Report CS-TR-2008-283, Dept. Comp. Sci., KAIST, Yuseong-Gu, Dejeon 305-701, Korea, Jan. 2008.

[6] Kim, M., Lee, J., Kim, M., and Whang, K., "Site-, Domain-, Community-Limited Search, and Home Search in a Web Search System Using the ODYSSEUS Object-Relational DBMS, " *Proc. 2005 Fall Conf. KIISE* (in Korean), pp. 175-177, Nov. 2005.

[7] Mecca, G., Raunich, S., and Pappalardo, A., "A New Algorithmfor Clustering Search Results, " *Data & KnowledgeEngineering*, 62(3):504-522, 2007.

[8] Zahn, C., "Graph Theoretical Methods for Detecting andDescribing Gestalt Clusters, " *IEEE Trans. Computers*, C-20(1):68-86, Jan. 1971.

[9] Google Help, http://www.google.com/help/refinesearch.html

[10] Kessler, M., "Bibliographic Coupling between Scientific Papers, "*American Documentation*, 14(1):10-25, 1963.

[11] Small, H., "Co-citation in the Scientific Literature: A NewMeasure of the Relationship between Two Documents, " *Journal of the American Soc. of Inf. Sci.*, 24(4):265-269, 1973.

[12] Bingjing CAI, Haiying Wang, Huiru Zheng and Hui Wang, " An improved random Walk based Clustering Algorithm for Community Detection in Complex Networks", University of Ulster, N. Ireland, UK.

[13] Schaeffer, S.E., *Survey: Graph clustering. Computer Science Review I (2007)*, 2007: p. pp. 27-64.

[14] Reihaneh Rabbany, Mansoreh Takaffoli, Justin Fagnan, Osmar R. Zaiane and Ricardo J. G. B. Campello, " Relative Validity Criteria for Community Mining Algorithms", University of Sao Paola, Sao Carlos, SP, Brazil.

[15] M. E. J. Newman and M. Girvan, "Finding and evaluating community structure in networks, " *Physical Review E*, vol. 69, no. 2, p. 026113, 2004.

[16] Newman, M.G.a.M.E.J., *Community structure in social and biological networks.* Proc. Natl. Acad. Sci. USA, 2002. **99**: p.7821–7826.

[17] S. Fortunato and M. Barth´elemy, "Resolution limit in communitydetection, " *Proceedings of the National Academy of Sciences*, vol. 104, no. 1, pp. 36–41, 2007.

[18] Wu Xiaoli and Tian Hui, "Discussion on E-commerce key technology", Henan Institute of Urban Construction.

[19] Tairan Liu, "E-Commerce Application model based on Cloud Computing", School of Business Administration, Changchun University of Technology, China.

[20] Shailesh Kumar Shivakumar, "Maximizing Knowledge Management Returns in E-Commerce", School of Computer and Information Sciences, New Delhi.

[21] Xiaobo Wu, Qi Chen and JingjingGuo, " Impact of M-Commerce on Business models", School of management, Zhejiang University, Hangzhou, China.

[22] ShafiqurRehman and Naveed Ejaz, "Towards an efficient Shopping Mechanism for M – Commerce", FAST National University, Islamabad, Pakistan.

[23] Newman, M.E.J., *Modularity and community structure in networks.* Proc. Natl. Acad. Sci., 2006. **103**: p. 8577–8582.