RESEARCH ARTICLE OPEN ACCESS

Modified Fuzzy C Means Clustering To Study the Willingness Maximization Using Discrete Multi-Valued Particle Swarm Optimization (DPSO) For Social Activity Planning

Alagulaskhmisubha. A [1], Mohanapriya. C [2]
Research Scholar [1], Assistant Professor [2]
Department of Computer Science
Tirupur Kumaran College for Women, Tirupur
Tamil Nadu - India

ABSTRACT

In the recent work demonstrate that a person is ready to join a social group action if the particular activity is attractive, and if a number of close friend's moreover also join the activity as companions. From the survey it shows with the intention of the interests of a person and the social rigidity between friends is able to be successfully derived and mined from social networking websites. On the other hand, even by means of the above mentioned two categories of information extensively presented, social group activities still necessitate to be corresponding physically ,and the development is tedious and time-consuming on behalf of users, particularly for a large social group activity, because of the several difficulty of social connectivity and the variety of feasible interests between friends .To solve above mentioned problems in this research work presents a new Modified Fuzzy C Means (MFCM) for the grouping of similar users particularly for large s ocial group activity. MFCM clustering method suggests prospective attendees of a social group activity, which might be extremely helpful, designed for social networking websites as a value-added service. For this grouping of similar user activities first need to specify a new problem, named Willingness mAximization for Social grOup (WASO). To solve the problem of WASO points out that the solution obtained by Discrete multi-valued Particle Swarm Optimization (DPSO) is expected to be attentive in a local optimal solution. Thus, new DPSO randomized algorithm to successfully and proficiently solve the problem. Specified the presented computational budgets, the proposed WASO-DPSO algorithmic capable to optimally assign the resources and discover a solution by means of an approximation ratio.

Keywords:- Social Network, Query Processing, Optimization, Social Media; Data Mining; Social Data; Social Media Mining;

I. INTRODUCTION

Studies demonstrate with the purpose of two major factors are frequently concerned in the result of a person joining a group activity on her obtainable time. Primary, the person is concerned in the inherent property of the activity, which might exist in line by means of her preferred hobby or do exercises. Subsequent, other people who are significant to the person, such as her close friends, determination join the activity as accompanying person. For instance, if a person who appreciates conceptual art has kind tickets designed for an up to date art demonstration on MoMA, she would most likely desire to attract her friends and friends of friends by means of this common interest. At the present time many people are familiar to sharing information through their friends on social networking websites such as Facebook, twitter, MyYearbook, and LikeALittle, and some of the work done in the literature [1-2] introduces a several methods to measure the interests of a person related to the importance attributes in her personal profile and the suitable information in her interaction by means of friends.

Furthermore, social connectivity schemas have been developed in [3] used for determining the tightness among two friends along with their social websites. Nevertheless, even by means of the exceeding knowledge accessible, to year there has be neither published work nora real structure discover how to control the above two important factors with automatic grouping and commending of a group activity, which is extremely helpful designed for social networking websites. At current, several social networking websites simply acts as a platform designed for information sharing and substitute in activity planning. The attendees of a group activity at rest need toward be chosen physically, and such physical management is usually difficult and time-consuming, particularly designed for a huge social activity, and specified the complex social link structure and the diverse interests. To solve this problem this research work integrates the interests of people and their activities to discover a group of attendees designed for automatic planning and management. It is popular to choose more attendees who like and have the benefit of the activity by means of the shared interest in the activity as companions. In reality, Psychology [4] and some work in the literature of social networks [5-6] have formed the willingness to be present at an activity as the sum of the interest of every attendee on the activity and the social tightness among friends with the intention of possible to join it. It is visualize with the intention of the chosen attendees are additional leaning in the direction of join the activity if the willingness of the group increases.

To solve the above mentioned complication problems data mining provides an efficient solution for handling huge amount of information available in the websites of social Medias. Traditional data mining methods is applied to several applications such as bioinformatics, data warehousing, business intelligence, investigative analytics, and decision support systems. Most important aim of the DM procedure is to successfully handle large-scale information, mine important patterns. Since the social media is extensively designed for various purposes, huge amounts of user-generated information exist and be able to be finished obtainable designed for data mining. For example, DM techniques be able to assist recognize the important people in the huge blogosphere, distinguish concealed groups in a social networking site, analysis sentiment users are designed for proactive planning, develop suggestion systems designed for tasks ranging beginning export specific products in the direction of making new friends, identify with network development and varying entity relationships, defend user privacy and security among users and entities. Mining social media is a rapidly increasing multi corrective area where researchers of various backgrounds are able to make significant contributions with the purpose of matter designed for social media research and improvement.

Clustering is a common example of unsupervised learning which provides a correct solution for social network activity which is determined based on the similarity or dissimilarity among data objects. Similarity or dissimilarity among the user or their activities is determined based on the distance measures such as Euclidean distance, Minkowski , Mahalanobis distance. Some other metrics are Jaccard coefficient, cosine similarity, and Pearson’s correlation has been used to determine the similarity or dissimilarity among the users. K-means, hierarchical clustering, and density-based clustering are used to determine the similarity or dissimilarity among data objects. Semi supervised learning algorithms follow the procedure of classification methods. Two most importantly used classification methods are semi supervised classification and semi supervised clustering. Active learning algorithms permit users to participate an active position in the learning procedure via labeling. Characteristically, users are domain experts and their skills are working to label a number of information instances designed for which a machine learning algorithm are positive concerning its classification. Some other techniques in the data mining methods are association rule mining, feature selection, instance selection, and visual analytics which is studied in the literature Han et al. [7], Tan et al. [8], Witten et al. [9], Zhao and Liu [10], and Liu and Motoda [11].

From this motivation of data mining methods, formulate a new problem as Willingness mAximization for Social grOup (WASO) for social networks. The WASO difficulty is represented as social graph G, where each node in a graph denotes as the candidate person and is connected by means of an interest score of the person designed for the activity, and every edge has a social rigidity score to specify the mutual knowledge among the two persons. Let k represent the number of predictable attendees. Specified the user defined value as k, the objective of regular activity planning is to make best use of the willingness of the preferred group F, at the same time as the induced graph on F is a associated sub graph designed for each attendee to develop into familiar by means of an additional attendee related to a social path. In favor of the social activities not including an a priori predetermined size, it is practical designed for users in the direction of denote a appropriate range used for the group size, and proposed algorithm be able to discover the solution for each k inside the range and revisit the solutions designed for the user to make a decision the most appropriate group size and the equivalent attendees. Naturally, too incrementally construct the group, a deterministic discrete multi-valued Particle Swarm

Optimization (DPSO) successively decides an attendee with the purpose of leads to the leading increment in the willingness on each iteration. Modified Fuzzy C Means (MFCM) is proposed for the grouping of similar users particularly for large social group activity. MFCM clustering method proposes potential attendees of a social group activity, which capacity is particularly cooperative, designed for social networking websites as a value-added service. Note the DPSO, although easy, tends to be attentive in a local optimal solution, because it facilitates the choice of nodes simply appropriate at the equivalent iterations.

II. RELATED WORK

A modern line of study has been proposed toward locate cohesive subgroups during social networks through dissimilar criteria, such as cliques, n-clubs, k-core, as well as k-plex. Sar´ıy¨uce et al [12] proposed an efficient equivalent algorithm toward find a k-core sub graph, anywhere every vertex is connected toward at least k vertices within the sub graph. Propose the primary incremental k-core decomposition algorithms designed for streaming graph data. These algorithms establish a little sub graph so as to is guaranteed toward contain the list of vertices whose maximum k-core values encompass to be updated, furthermore efficiently process this sub graph toward update the k-core decomposition. Experimentation results illustrate a significant reduction during run-time compared to non-incremental alternatives. Show the competence of our algorithms taking place different types of real as well as synthetic graphs, at different scales. For a graph of 16 million vertices, examine speedups accomplishment a million times, comparative to the non-incremental algorithms.

Xiang et al [13] proposed a branch-and-bound algorithm toward attain each and every one maximal cliques so as to cannot be pruned through the search tree optimization. Enables us toward successfully use Map Reduce is a recursive partition method with the purpose of partitions the graph into several sub graphs of similar size. After partitioning, the highest cliques of the dissimilar partitions be capable of computed independently, furthermore the computation is sped up with a branch as well as bound method. Furthermore, finding the maximum k-plexes was expansively discussed in [14]. Derives a latest upper bound scheduled the cardinality of k-plexes as well as adapts combinatorial clique algorithms toward find maximum k-plexes. On the other hand, community detection as well as graph clustering encompass been exploited toward identify sub graphs through desired structures [15]. They study problem of discovery clustering's through maximum modularity, consequently providing theoretical foundations used for past as well as present work base on this measure. More precisely, we establish the conjectured hardness of maximizing modularity together in the common case as well as through the restriction cuts as well as give an Integer Linear Programming formulation. This is complemented by first insights addicted to the behavior and efficiency of the frequently applied greedy agglomerative schema.

The quality of a community is calculated according toward the organization within the community as well as the structure among the community along with the rest of nodes inside the graph, such as the density of neighboring edges, deviance as of a random null model, along with conductance [16]. Discover is so as to the neighborhood communities be able to exhibit conductance scores to facilitate are as good as the Fiedler cut. Also, the conductance of neighborhood communities shows comparable behavior since the network community profile computed with a personalized PageRank community detection method. Neighborhood communities present us a simple as well as powerful heuristic used for speeding up local partitioning methods. While finding good seeds used for the PageRank clustering scheme is complex, the majority approach involve an exclusive sweep over a great many starting vertices.

In addition toward dense sub graphs, social groups through dissimilar characteristics have been explored used for varied useful applications. Expert team arrangement during social networks has concerned extensive research interest. The problem of construct a specialist team is to discover a set of people possessing the required skills, whereas the announcement cost with the choose friends is minimized toward optimize rapport between the team members toward ensure efficient operation. Communication costs be capable of be represent by the graph diameter, the range of the minimum spanning tree, as well as the whole length of the shortest paths [17]. By contrast, minimizing the whole spatial distance with R-Tree from the grouping through a given amount of nodes to the rally point is also studied [17].

Given the growing importance of different social network application, present has been a latest push on the learning of user concentration scores as well as social tightness scores since real social networking data. Wilson et al. [18] consequent a innovative representation toward quantify the social tightness among some two friends in Facebook. The number of wall postings is furthermore established to be an efficient display for social tightness. Consequently, the exceeding study present a sound foundation toward quantify the user interest as well as social tightness scores in social networks. Moreover, Yang [5] considers two major factors at the same time as willingness designed for marketing and recommendation. On the other hand, the above mentioned factors are difficult in social networks enclose not been leveraged designed for automatic activity planning. Since the Expert team development in social networks has concerned widespread research interests. The difficulty of constructing an expert team is to discover a set of people owning the particular skills, while the communications cost amongst the chosen friends is reduced to guarantee the relationship in the midst of the team members designed for a well-organized operation.

Sozio et al [19] introduces a new model for social community measurement by means of reducing the total degree of a community through specified nodes. On the other hand, the objective function of WASO is varied from community detection. Every node and every edge in WASO are related by means of interest score and social tightness score which is used in this research work to reduce the willingness of the attendees by means of a specified group size, which is able to be very helpful for social networking websites as a value-added service. On the other hand, the above mentioned methods doest measure the interest score and the social tightness scores among users, which is focused in this research work. Furthermore, the activity cost designed for the group is not integrated throughout the evaluation

III. PROPOSED METHODOLOGY

From this motivation of data mining methods, formulate a new problem as Willingness mAximization for Social grOup (WASO) for social networks. The WASO difficulty is represented as social graph G, where each node in a graph denotes as the candidate person and is connected by means of an interest score of the person designed for the activity, and every edge has a social rigidity score to specify the mutual knowledge among the two persons. Let k represent the number of predictable attendees. Specified the user defined value as k, the objective of regular activity planning is to make best use of the willingness of the preferred group F, at the same time as the induced graph on F is a associated sub graph designed for each attendee to develop into familiar by means of an additional attendee related to a social path. In favor of the social activities not including an a priori predetermined size, it is practical designed for users in the direction of denote a appropriate range used for the group size, and proposed algorithm be able to discover the solution for each k inside the range and revisit the solutions designed for the user to make a decision the most appropriate group size and the equivalent attendees. Naturally, too incrementally construct the group, a deterministic discrete multi-valued Particle Swarm Optimization (DPSO) successively decides an attendee with the purpose of leads to the leading increment in the willingness on each iteration. Modified Fuzzy C Means (MFCM) is proposed for the grouping of similar users particularly for large social group activity. MFCM clustering method proposes potential attendees of a social group activity, which capacity is particularly cooperative, designed for social networking websites as a value-added service. Note the DPSO, although easy, tends to be attentive in a local optimal solution, because it facilitates the choice of nodes simply appropriate at the equivalent iterations.

Problem Definition

Let consider the social network and it is represented as graph model $G = (V; E)$, where each node $v_i \in V$ and each edge $e_{i,j} \in E$ are related to the interest score h_i and a social tightness score $t_{i,j}$ found in the literature [20] and [1] respectively, from this define a new problem WASO designed for finding a set F of vertices by means of size k to make best use of the willingness W(F), i.e.,

$$\max_F W(F) = \max_F \sum_{v_i \in F} \eta_i + \sum_{v_j \in F : e_{i,j} \in E}^{1} \tau_{i,j} \quad (1)$$

where F is a associated sub graph in G to support every attendee to be familiar with an additional attendee related to a social path in F. Notice with the intention of the social tightness among v_i and v_j is not essentially symmetric, i.e., $t_{i;j}$ be able to be dissimilar by means of $t_{j;i}$. Consequently, the willingness in Eq. (1) regard as both $t_{i;j}$ and $t_{j;i}$. In this work introduce clustering methods to group similar user nodes in the graph model. Nodes into various clusters and includes the learning of the social network data or nodes information as well as measuring the similarity between two nodes in the social network model. The centroid values in the MFCM are chosen randomly. The separation of nodes $N = (n_1, .. n_k)$ into c cluster is done by the standard FCM objective function which is given as

$$J = \sum_{i=1}^{c}\sum_{k=1}^{N} \mu_{ik}^{p} ||n_k - v_i||^2 \quad (2)$$

Where $\{v_i\}_{i=1}^{c}$ are the centriods values of the clusters and the array $[\mu_{ik}^{p}] = U$ is denoted as fuzzy membership matrix $U \in u$. The parameter p is each fuzzy membership weighting exponent it calculates the fuzziness value. The minimization of FCM objective function is done based on the nearest node value founded in the social network model with high membership values, and the nodes which is not nearest to centroid value have low membership values, alteration is proposed to (2) by means of using a term α with the intention of allows the control of labeling of a cluster nodes to the labels in its direct neighborhood nodes. The new objective function of FCM is given by

$$J_m = \sum_{i=1}^{c}\sum_{k=1}^{N} \mu_{ik}^{p} ||n_k - v_i||^2 + \frac{\alpha}{N_R}\sum_{i=1}^{c}\sum_{k=1}^{N} \mu_{ik}^{p}\left(\sum_{n_r \in \mathcal{N}_k} ||n_k - v_i||^2\right) \quad (3)$$

Where $\mathcal{N}_k$ is the set of neighbors nodes of similarity function from equation (1-10) and N_R is cardinality of the $\mathcal{N}_k$ .The parameter α is controlled by the effect of the neighborhood node term . This MFCM algorithm involves the similarity between nodes as well as the study of the information of users in social network model.

Discrete multi-valued Particle Swarm Optimization (DPSO)

DGreedy offensively selects v_1 as the initial node and discover simply a solitary sequence of nodes in the solution space. To enhance the search space, randomized methods is introduced in the literature as a straightforward however efficient approach to solve the problems by means of large instances. To avoid local optimal selection problem for WASO is to indiscriminately select multiple number of nodes as start nodes. Each start node is measured as limited solution, and a node adjacent the limited solution is indiscriminately selected and further to the limited resolution on each iteration subsequently, until k nodes are integrated as a final results. This method is well-organized than DGreedy, since the computation of willingness is not concerned throughout the collection of a node. Designed for the difficulty by means of a large k, frequent candidate nodes neighboring the incomplete solution are essential to be inspecting in DGreedy to sum up the willingness, in order to discover the one with the intention of produces the largest willingness. In contrast, the randomized algorithm basically decides one neighboring node on random.

On the other hand, this assignment is comparable to the DGreedy with the purpose of it limits the range to the confined information related to each node and is not predictable to produce a solution with high willingness. To solve this problem introduce a new randomized algorithms, called Computational Budget Allocation designed for Start nodes (CBAS) and Computation Budget Allocation designed for Start nodes by means of Neighbor Differentiation (CBAS-ND), by satisfaction of initial node selection problem and increasing the partial solutions, respectively. This research work follows the procedure of Optimal Computing Budget Allocation (OCBA) [22] in randomization to reduce the computational cost in the start nodes by means of the prospective to create the solutions by means of elevated willingness. CBAS initially chooses m nodes as start nodes and subsequently randomly adds neighboring nodes in the direction of develop the partial solution stage-by-stage, awaiting k nodes are integrated as an absolute solution. Every start node in CBAS is extended to numerous final solutions. To appropriately provide the computational cost budgets, CBAS at every step recognize the start nodes significance more computational budgets related to sampled results of the preceding stages. Operational by means of the allocation strategy of computational resources, CBAS is improved to CBAS-ND in the direction of adaptively allocate the probability toward each neighboring node throughout the expansion of the limited solution related to cross entropy method. Demonstrate with the intention of the allocation of computational budgets designed for start nodes in CBAS and CBAS-ND, correspondingly.

Explain the IP formulation designed for WASO. Binary variable x_i represent if node v_i is preferred in the solution F, and binary variable $y_{i;j}$ represent if two neighboring nodes v_i and v_j are equally selected in F. The objective function is

$$\max_F W(F) = \max_F \sum_{v_i \in F} \eta_i \, x_i + \sum_{v_j \in F : e_{i,j} \in E}^{1} \tau_{i,j} \, y_{i,j} \quad (4)$$

where the primary term is the total interest score, and the subsequent term is the sum social tightness score of the chosen nodes. The above mentioned functions are optimized using DPSO. Particle swarm optimization (PSO) is an optimization technique developed by James Kennedy and Russell Eberhart [25]. PSO derives a set of potential problem solutions as a swarm of particles which is moving about in a virtual search space. In this work each node is assigned by a particle in the social network group and above mentioned function (4) is considered as the fitness function . PSO is an algorithm is performed based on the movement of flocking birds and their interactions by means of their neighbors in the social network group. Initially define the random position (np_i) and (possibly) randomized velocity (pv_{ij}) is allocated to n-dimensional search space to every node in the social network clusters in the swarm, where $np_{i,j}$ represents the nearest distance point value in the cluster i in the j-th dimension of the optimal social network search space. The optimization of candidate node value is determined by flying the particles with objective function of the nodes in the search space. The best nodes value is determined and represented as $(np^*_{i,j})$ is remembered by each node in the cluster (particle). Every node in the social network cluster (particle) determination be a neighborhood of other nodes particles consequently it is considered as a member. This neighborhood of node is considered as a subset of the particles, or each and every one the particles. The results of local nodes in the social network cluster are compared to equation (4) in the search space. Here the particle is represented as binary value is given by:

$$pv_{i,j} = pv_{i,j} = pw.rand().(np^*_{i,j} - np_{i,j}) + nw.rand().(np^*_{i,j} - np_{i,j}) \quad (5)$$

$$np_{ij} = 1 \; if \left(rand(\;\;) < r(v_{ij})\right), 0 \; otherwise \quad (6)$$

A new node is selected from the cluster , here automatically assign new velocity value $|v_{i,j}| < V_{max}$, where V_{max} is integer value typically close to 6.0 and probability is determined between various nodes in the social network group. In DPSO each particle's node is represented in the form of 2-dimensional to 3-dimensional: $np_{i,j,k}$ is the probability of social network node position i, element j between 0 or 1. It is determined using the following equation:

$$np'_{i,j} = \sum_{k=0}^{n} rv(np_{i,j,k}) \quad (7)$$

$$P(rv_j = k) = \frac{rv(np_{i,j,k})}{np'_{i,j}} \quad (8)$$

here $np'_{i,j}$ refers normalizing node rank value coefficient designed for node particle i and element j. By using this MCDPSO, the particle generates any minimum cost localization problem solution with varying probabilities depending on its terms. An adjustment is applied to node particle terms after each modification of the particle values. This adjustment is given by ,

$$np'_{i,j,k} = np_{i,j,k} - cn(i,j) \quad (9)$$

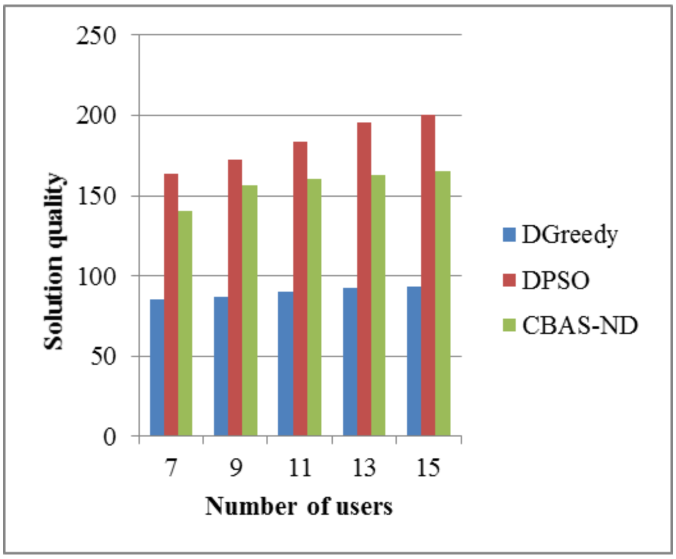
for all k, where $np'_{i,j,k}$ is the probability indicator of particle i, constituent j presumptuous value k, with $cn(i,j)$ selected with the purpose of

$$\sum_{k=0}^{n} rv(np'_{i,j,k}) = 1 \quad (10)$$

MCDPSO technique differs from standard PSO, inertia coefficient w linearly decreases from 1.2 to 0.6 at the end. Ring topology is second-hand for the swarm and allocated to the nearest particle on each side to be present a neighbor, pw and nw are together set to 2.0.

IV. EXPERIMENTATION RESUTLS

In this section measure the performance accuracy of various optimization methods to Facebook social network it invite 137 people beginning various communities, e.g., schools, government, technology companies, and businesses. Furthermore, each user in the social network is asked to preparation 10 social activities by means of the social graphs extracted from their corresponding social networks in Facebook. These 10 activities explore differed network sizes and different numbers of attendees in three different scenarios.



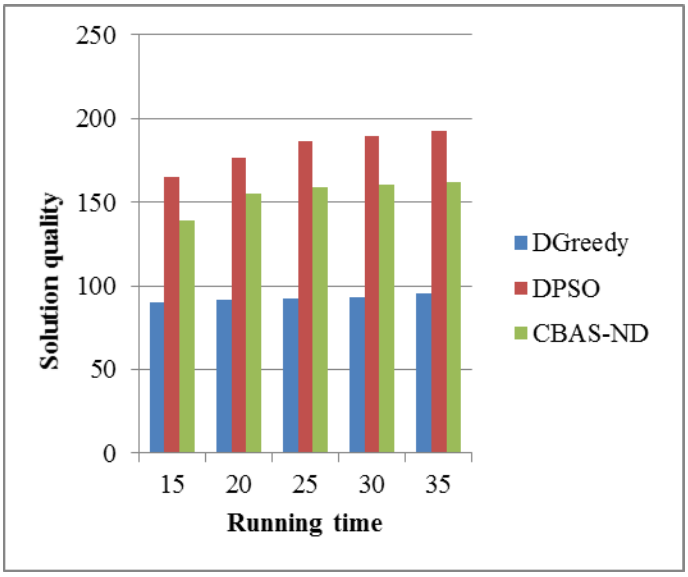


Fig. 1. Results of user study

Figs. 1a and 1b shows the performance comparison results of various methods with their solution quality and running time with varied network sizes, where the expected number of attendees k is 7. The experimentation results demonstrate that the proposed DPSO provides best optimal results for solving the WASO problem when compared to existing DGreedy and CBAS-ND method .Since the existing methods are hard for manual coordination, even when the network size is high.

V. CONCLUSION AND FUTURE WORK

In this paper work define a new problem definition based on Willingness mAximization for Social grOup (WASO) to obtain a set of attendees and make best use of the willingness. Proved with the intention of WASO is NP-hard and devised new Discrete multi-valued Particle Swarm Optimization (DPSO) with an approximation ratio. Naturally, too incrementally construct the group, a deterministic DPSO successively decides an attendee with the purpose of leads to the leading increment in the willingness at every iteration. Modified Fuzzy C Means (MFCM) is proposed for the grouping of similar users particularly for large social group activity. MFCM clustering method proposes potential attendees of a social group activity, which capacity is particularly cooperative, designed for social networking websites as a value-added service. Note the DPSO, although easy, tends to be attentive in a local optimal solution, because it facilitates the choice of nodes simply appropriate at the equivalent iterations. A number of users suggested with the intention of incorporate the proposed DPSO system with automatic obtainable time extraction to filter busy users, such as by means of integrate the proposed DPSO system by means of Google Calendar. Because candidate attendees are related by means of multiple attributes in Facebook, e.g., location and gender, these attributes be able to be particular as input parameters in the direction of advance filter out inappropriate candidate attendees. In addition it also is extensive in the direction of permit users to identify some attendees with the intention of should be incorporated in a certain group activity.

REFERENCES

[1] Clauset, C. R. Shalizi, and M. E. J. Newman. Power-law distributions in empirical data. In SIAM, 51(4):661–703, 2009.

[2] Mislove, B. Viswanath, K. P. Gummadi, and P. Druschel. You are who you know: Inferring user profiles in online social networks. In WSDM, pages 251–260, 2010.

[3] V. Chaoji, S. Ranu, R. Rastogi, and R. Bhatt. Recommendations to boost content spread in social networks. In WWW, pages 529–538, 2012

[4] M. F. Kaplan and C. E. Miller. Group decision making and normative versus informational influence: Effects of type of issue and assigned decision rule. Journal of Personality and Social Psychology, 53(2):306–313, 1987.

[5] D. N. Yang, W. C. Lee, N. H. Chia, M. Ye, and H. J. Hung. Bundle configuration for spread maximization in viral marketing via social networks. In CIKM, pages 2234–2238, 2012.

[6] M. Ye, X. Liu, and W. C. Lee. Exploring social influence for recommendation - a probabilistic generative model approach. In SIGIR, pages 671–680, 2012. Han, Moncepts and Techniques. Morgan Kaufmann, San Francisco, 2011

[7] P.-N. Tan, M. Steinbach, and V. Kumar. Introduction to Data Mining. Pearson AddisonWesley, Boston, 2006.

[8] H. Witten, E. Frank, and M. A. Hall. Data Mining: Practical Machine Learning Tools and Techniques. Morgan Kaufmann, San Francisco, 2011.

[9] Z. A. Zhao and H. Liu. Spectral Feature Selection for Data Mining. Chapman & Hall/CRC Press, Virginia Beach, VA, 2012.

[10] H. Liu and H. Motoda. Computational Methods of Feature Selection. Chapman & Hall, Boca Raton, FL, 2008.

[11] E. Sar´ıy¨uce, B. Gedik, G. Jacques-Silva, K.-L. Wu, and U. V. C¸ ataly¨ureks. Streaming algorithms for k-core decomposition. In VLDB, 6(5):433–444, 2013

[12] J. Xiang, C. Guo, and A. Aboulnaga. Scalable maximum clique computation usingmapreduce. In ICDE, pages 74–85, 2013.

[13] McClosky and I. V. Hicks. Combinatorial algorithms for max k-plex. In Journal of Combinatorial Optimization, 2012.

[14] F. Gleich and C. Seshadhri. Vertex neighborhoods, low conductance cuts, and good seeds for local community methods. In KDD, pages 597–605, 2012.

[15] M. Kargar and A. An. Discovering top-k teams of experts with/without a leader in social networks. In CIKM, pages 985–994, 2011.

[16] N. Yang, C. Y. Shen, W. C. Lee, and M. S. Chen. On socio-spatial group query for location-based social networks. In KDD, pages 949–957, 2012.

[17] C. Wilson, B. Boe, A. Sala, K. P. N. Puttaswamy, and B. Y. Zhao. User interactions in social networks and their implications. In Proc. Eurosys, 2009.

[18] M. Sozio and A. Gionis. The community-search problem and how to plan a successful cocktail party. In Proc. KDD, pages 939–948, 2010.

[19] V. Chaoji, S. Ranu, R. Rastogi, and R. Bhatt, “Recommendations to boost content spread in social networks,” in Proc. Int. Conf. World Wide Web, 2012, pp. 529–538.

[20] M. Mitzenmacher and E. Upfal, Probability and Computing: Randomized Algorithms and Probabilistic Analysis. Cambridge, U.K.: Cambridge Univ. Press, 2005.

[21] C. H. Chen, E. Yucesan, L. Dai, and H. C. Chen, “Efficient computation of optimal budget allocation for discrete event simulation experiment,” IIE Trans., vol. 42, no. 1, pp. 60–70, 2010.