Analysis of Social Network Using Clever Ant Colony Metaphor

Mohammad Al-Fayoumi, Soumya Banerjee, Jr., P. K. Mahanti

***Abstract*—**A social network is a set of people or organization or other social entities connected by some form of relationships. Analysis of social network broadly elaborates visual and mathematical representation of that relationship. Web can also be considered as a social network. This paper presents an innovative approach to analyze a social network using a variant of existing ant colony optimization algorithm called as Clever Ant Colony Metaphor. Experiments are performed and interesting findings and observations have been inferred based on the proposed model.

***Keywords*—**Social Network, Ant Colony, Maximum Clique, Sub graph, Clever Ant colony.

I. INTRODUCTION

A social network is a social structure between actors primarily individuals and organizations. Broadly, these networks implicate the possibilities in which they are connected through various social familiarities spreading from casual acquaintances to proximal business associations. During the last three years social network such as orkut[1], ryze[2], [3]linkedin have had phenomenal growth and impact on web traffic. The majority of these networks are for personal and socialization purpose, however social network analysis leads to mapping, measuring and modeling the relationships and flow between people, groups, organization or between any living or non living entities. From modeling point of view, the nodes in the network are the people or community, whereas the link exhibits relationships and flow between the nodes. Functionally, the number, size connection arcs among the subgroups may explore the number of information like conflict among multi groups, overlapping or entropy structure of information etc. Email or web traffic, transmission of pandemic diseases or terrorist plots all could be modeled as social network. In this context, the self-organized behavior of agent can be utilized for finding the similarity of different

Mohammad Al-Fayoumi,Department of CIS,Middle East University for Graduate Studies,Jordan– Amman(e-mail: mfayoumi99@yahoo.com)

Soumya Banerjee, Department of Computer Science,Birla Institute of Technology,InternationalCenter,Mauritius
e-mail:. sumo_bani@yahoo.com

P. K. Mahanti,Department of CSAS,University of New Brunswick, New Brunswick, Canada ,e-mail: pkmahanti@yahoo.co.in

[1] www.orkut.com

[2] www.ryze.com

[3] www.linkedin.com

community instances. Hence, clustering of subgroups and cliques seems to be more effective and practically mandate solution. Recently, ant colony optimization technique has demonstrated significant promises by solving several intelligent clustering based complex problems. Clustering is concerned with the division of data into homogenous subgroups. The objective of this division is bi-fold in implementation: data items within one cluster are required to be similar to each other, while those within different cluster should be dissimilar. Problem of this type arise in a varieties of disciplines encompassing from sociology to psychology to commerce and biology including computer science as well. Therefore, the flavors of clustering available are also versatile like data handling (numerical, proximity and categorical data), the shape of cluster, and the form of final partitioning hard versus fuzzy assignment). Behavior of real ant colony demonstrates the core clustering of corpses and larval through their dropping and picking up activities from one place to another place **[4].** This inspires the present research proposal and certain variant of core ant colony optimization algorithm has been developed to investigate different clusters and cliques of any standard social network. Similar to diversified ant colony types and scenario, there are also different forms of social networks available in practice e.g. web, weblogs (special subset), semantic web, blogs etc. These components have been combined in this proposal for validating the proposed model for finding the maximum cliques and sub graph. The remaining part of the paper is organized as follows: Section 2 defines the problem statement studied in this paper, followed by related works on social network and ant colony optimization to find out clique or sub graph paradigm in section 3. Section 4 describes the proposed *Clever Ant colony Metaphor (CACM)* and its implications. Section 5 discusses the experimental setup to test the result of proposed algorithm. Section 6 and section 6.1 elaborates benchmark, result and comparison with classic Ant Colony approach. Finally section 7 presents conclusion and mentions further scope of research in this context.

II. PROBLEM STATEMENT

A social network is considered as a descriptive framework to study and analyze social relation, behavior and measure social matrix.

Let A = $\{a_1, a_2, \ldots a_n\}$ be a finite set of individual of relation and ties. These interacting actors form a connected social

network. It has been observed that division of actor into clique or subgroups can be very important aspect of social structure. The clique becomes instrumental to point out the holistic behavior of the social network. Fundamentally, a clique of a graph is fully connected substructure. Cliques with the maximum number of vertices are the maximum cardinality cliques of the graph. Therefore if we assume the social relation to be symmetric, then there are M = n (n-1)/2 possible ties among the actors. Social networks are represented either by graphs called socio graph or adjancey matrix. The objective is using some form of textual data (which are associated with actor) and the new ties those could be explored to accomplish the completeness and enrichment of social network. This proposed model of *Clever Ant Colony Metaheuristics (CACM)* would like to explore maximum clique sub graph problem.

III. RELATED WORK

One reason for using mathematical and graphical techniques in social network analysis is to represent the description of networks compactly and systematically **[8]**. Data mining has been also used for mining social networks. For instance, the social networks in a message board are extracted by using the influence diffusion model. The social networks are categorized to study some social and psychological issues such as interactivity among members, their strong or weak choices etc. Another approach in extracting social networks is using usage and log data instead of textual contents **[5]**. In the context of selecting the variant of ant colony metaphor *(i.e. Clever Ant Colony Metaphor)* for treating these categories of problems, it will be worthy to provide a background on the same. Ant colony optimization, which is a class of algorithms, is a recently developed population-based approach. It has been successfully applied to several NP-hard combinatorial optimization problems **[3, 7, 1, 2].** As the name suggests, ACO was inspired by the observation of real ants' foraging behavior. Ants live in colonies. They use a cooperative method to search for food. While moving, ants initially explore the area surrounding their nest in a random manner. They initially leave a chemical pheromone trail on the ground. During the return trip, the quantity of pheromone that an ant leaves on the ground may depend on the quantity and quality of the food. This result in the amount of pheromone becomes larger on a shorter path. Then the probability that an ant selects this shorter path is higher. At last, the pheromone trails will guide other ants to the food source via the shortest path. Considering this ant colony background, the present problem has been interpreted into maximum clique sub graph problem. Subset selection problems are ideally to find an optimal feasible subset of an initial set of objects with respect to an objective function or some constraints. Stuetzle and Hoos introduced Max-Min Ant System (MMAS) in which two parameters, a maximum and minimum trail levels, were introduced to restrict possible trail values to the interval [τ_{max} τ_{min}]**.** Recently, Serge Fenet and Christine Solnon firstly proposed an ACO algorithm called Ant-Clique for the maximum clique problem **[6].** ACO is used as a heuristic for

choosing, at each step, the vertex to enter the clique. The proposed algorithm is the variant of the basic ant colony used for sub graph or clique customized in multicasting and secure environment.

IV. PROPOSED MODEL- CLEVER ANT COLONY METAPHOR (CACM)

The central theme of ACO is to model the problem as a search process for a minimum or maximum cost path in a graph. On each graph edge, the amount of pheromone is represented by τ (v_i , v_j) [if edge is (v_i , v_j)] and thus this becomes an identification of the learnt desirability for v_i, and v_j to belong the same path, hence the completeness of socio relation in social graph is achieved. The proposed version of *Clever Ant colony Metaphor (CACM)* has been introduced for secure multicast grouping. At each stage the vertex should enter the clique. This vertex is chosen with respect to a probability that depends on the pheromone trail laid between and the clique under construction, while pheromone trails are deposited in proportion to the quality of the previously computed cliques. *Clever Ant Colony* Metaphor proposed here has some basic differences over classical ant colony optimization:

- The path on which the ant agent travels in CACM is not an edge sequence, but it is a vertex and the pheromone is deposited on vertex.

The multicasting and secure environment in social network can also be defined as follows:

Secure multicast group is a triple (U, K, R) is a triple: Where U is a finite nonempty set of actors

K is a finite and non-empty set of keys between actors; R is a binary relation between U and K.

$$R \subseteq U \text{ X } K$$

Additionally, we assume CC: Current Clique

FNV: Feasible Neighbourhood vertex

It has to be mentioned that the pheromone factor $\tau(v_{ij})$ doesn't depend on the pheromone trail between last added vertexes in clique. The candidate vertex v_i, could be on all pheromone trail of the partial clique i.e.:

$$\omega(v_i) = \sum_{v_j \in \text{ Current Clique}} \omega(\tau(vi - vj))$$

ω is the vertex weight of partial clique

Based on this perspective, the *Clever Ant Colony Metaphor* (CACM) is designed on a social network according to the step mentioned below:

Step 1: At first step, we construct a map

f: (U,K,R) $\rightarrow$ (V,E; ω),

f established the completeness between (U,K,R) to an intermediate graph $G_s = (V_s,E_s)$ is a complete graph on a social network S

Step 2: There is no one to one correspondence between k in K and set of v in V represented as, v ↔ k.

Step 3: Vertex weight $\omega(v_i)$ doesn't only depend on the cost:

$$\tau(v_i) = \sum_{v_j \in \text{Current Clique}} \tau(vi - vj)$$

Step 4: Prior to devise the *CACM* algorithm some approaches defining substructures and groups in sociogram could be mentioned as prerequisite parameters. They are:

Clique: Actors who have all possible ties among themselves.

N clique: Actors are connected to every member of group at a minimum distance N

N class: All paths among members occur by the way of other members of N clique.

K plex: Cliques in which actors have ties to all but k of member of the group.

Algorithm.

Input: (U, K, R)

1. a. Initialize Pheromone trail τ (c) associated with each c $\in C$ to τ_{max}

   b. Initialize Pheromenal component C a pheromone strategy Φ on Social Network S

   2. Repeat

   3. {Generate a population of clever ants of size PopSize

   4. for each clever ant (C) in the population

   5. { Clique_Length= 0; N _Classs = 0;

      K_Plex=0;

   6. Randomly select the start CC (Current Clique)

   7. While candidate ≠ Φ do {

   8. Construct a feasible candidate neighbourhood pool of vertex FNV of CC.

   9. Evaluate the selection probability function v ∈ FNV

   10. Next Node = Select a feasible neighbourhood using random rule;

   11. Add Next Node to CC

       end while

   end for

12. CC_Length = CC_Length +1; }

13. Until clique is constructed and $G_s = (V_s, E_s)$ is obtained

14. Apply clever ant state transition rule for each c $\in C$ under the pheromone trail τ(c) or τ_{ij} as follows:

(While constructing the solution ants select the following vertex to be visited through a stochastic mechanism. When ant k is in vertex i and has so far constructed the partial solution s^p, the probability of going to next vertex j is given by):[4]

$$p_{ij}^k = \begin{cases} \dfrac{\tau_{ij}^{\alpha}\eta_{ij}^{\beta}}{\sum_{0}^{} c_{ij \in N(s^p)} \tau_{ij}^{\alpha}\eta_{ij}^{\beta}} & \text{if } c_{ij} \in N(s^p) \\ & \text{Otherwise} \end{cases}$$

15.} Apply local pheromone evaporation rule

16. Add to tabu ( )

17. Best_Tour = maximum clique_length per current population

18. $$\Delta\tau = \frac{1}{(1+|G\_best| - |best\_tour|}$$

19. if (best_tour> G_best)

20. G_best= best_tour //G_Best = Global Best

21. end if

22. Increment pheromone trail on the best tour by $\Delta\tau$

23. Update Current_best Clique, N class, Kplex

Update Tabu }

Until maximum cycles are reached for social network S or acceptable solution found.

Return the best solution since the beginning.

V. DESCRIPTION OF PROPOSED CACM ALGORITHM

Lines 5-13 of the proposed algorithm describe the procedure used by clever ants to construct cliques or subset of the social network under consideration. The pheromone factor τ evaluates the desirability of augmenting clique or subset in the given network based on the pheromenal strategy Φ. Similarly one heuristic factor η evaluates the promise of clique or subset based on the information local to the ant. Line 14 depicts the state transition rule followed by the ants at each cycle. Multiplying the quantity of pheromone on each pheromenal component by a pheromone persistence rate ρ simulates evaporation. The proposed algorithm follows the *MAX-MIN*

---

[4] Where, $N(s^p)$ is the set of feasible components; that is, edges (i, l) where l is a vertex not yet visited by the ant k. The parameters α and β control the relative importance of the pheromone versus the heuristic information η_{ij}, which is given by:

$$\eta_{ij} = \frac{1}{d_{ij}}$$

where d_{ij} is the distance between distance from the origin to the node under consideration.

Ant system. All components are initialized to the maximum allowed range (Line 1a) to keep exploration high and attractive in these cycles. The proposed algorithm has demonstrated certain additional features. It incorporates state transition, global update and local updating rule. In the global update the ant constructs the best tour and they re allowed depositing pheromone. Global updating is performed after all ants have completed their tours. The pheromone level is updated by: $\tau_j = (1-\alpha)\tau_j + \alpha\Delta\tau_j$ Where, $\Delta\tau = 0$ or

otherwise $(L_{Gb})^{-1}$ if $j \in$ Global Best Tour, Where $0 < \alpha < 1$ is the pheromone decay parameter and L_{Gb} is the size of visited vertex set. It is obvious that only pheromones on those nodes belong to the best tour, will enhance. Secure Multicast has been explained earlier. It is practically a scenario of intra-communication where, each user of the social network u has group key K and individual key I. Therefore the clique or subset in the network could perform secure group communication among them. As the final stage of local update, the algorithm follows: $\tau_j = (1-\rho)\tau_j + \rho\Delta\tau_j$,

where $0<\rho<1$ isa parameter, $\tau_j = \tau_0 = (mL_{mn}) -1$, here mL_{mn} is a rough approximation of the size of key covering set by clever ants over the social network.

VI. EXPERIMENTAL RESULTS AND TEST SUITE

TABLE I
SIMULATION RESULT OF CLEVER ANT COLONY METAPHOR

Name	N	Density	OCS	CACM Best	CACM Avg.	Std. Best With Ant Colony	Avg
MANN _a27	378	0.982	126	126	125.7	126	125.5
Brock400_2	400	0.733	29	29	25.5	29	24.7
Brock800_2	800	0.653	22	21	19.8	21	19.23
C500.9	500	0.906	57	55	54.3	56	55.5
Hammings8-4	256	0.647	16	16	15.8	16	16.0
P_hat300-1	300	0.243	8	7.9	7.7	8	8.0
P_hat300-2	300	0.493	25	25.0	25	25	25
P_hat300-3	300	0.749	36	36	36.0	36	36
P_hat700-1	700	0.244	11	11	11.0	11	11
P_hat700-2	700	0.488	44	44	44.0	44	44
P_hat700-3	700	0.753	62.2	62	62.1	62	61.12
Keller6	3361	0.826	59	57	52.23	55	53.1

For maximum clique problems, we demonstrate results obtained on the 5 DIMACS clique instances. The proposed *CACM* algorithm has been implemented in C++ on P-IV 2.4 G (512 MB RAM). In this proposed algorithm, 30 ants were used and other empirical parameters are as follows: $\alpha = 1$; ρ (0) = 1.0; $\Phi = 0.002$. Other parameters are taken from τ min = 0.01 and τ max =4

The performance of the proposed *Clever Ant colony Metaphor (CACM)* on the benchmark data set has been tested (Table 1). The column OCS represents the Optimum clique size and the standard result section contains the result of classic ant colony clique algorithm (Bypassing the pre-processing part with 30 ants) [6]. N represents number of vertices. From Table 1, certain important observations about proposed CACM could be drawn: Out of 12 graphs of DIMACS benchmark, at least in 8 cases the proposed *CACM* reaches close to the best solution. On 4 instances both the proposed and standard algorithm goes neck to neck. On Keller 6 the *CACM* proves to be better.

The improvement over classic Ant Colony Approach can also be highlighted. Based on the initial experiment and performance appraisal of the algorithm, the social random networks with maximum cliques have been envisaged. The data set was conceived from the popular social network orkut with approximate 2500 guest book entries over a day. Each member is represented as vertex and two vertices with an edge if at least one bi-lateral message exchange took place.

In this context, Fig.1 demonstrates the community structure (adjacency matrix shown) achieved by *Clever Ant Colony Metaphor (CACM)*. The weights of the intra community model have been presented by different gray degrees as shown. On the Keller 6 instance the *CACM* proved better result than classic Ant colony approach. From the standard result [6] it is evident that, if the number of ants is fixed to 7, the average quality of solution could be improved for 18 instances, but still the proposed *CACM* found to be better due to its traversal through vertex and the deposition of pheromone on vertex. It should be noted that a relatively low probability value within community link persists if the graph instance corresponds to Keller 6 (shown in Table 1)
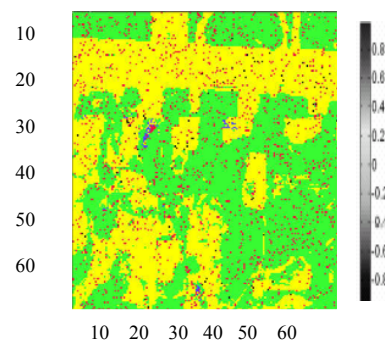

Fig. 1 Identifying community structure from maximum Clique random Social network

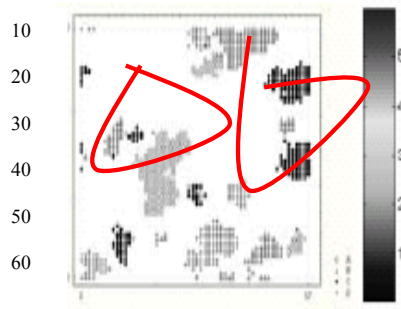http://. dimacs.rutgers.edu[5]

Fig. 2 4 predefined clusters in Social Network Detected by CACM

Fig. 2 presents certain predefined homogeneous clusters emerge due to *Clever Ant's* movements considering maximum cliques as criteria. The red mark indication emphasizes the cluster, which is of maximum cliques.

VII. CONCLUSION

In this work, a proposed algorithm called as *Clever Ant Colony Metaphor* is presented using social network structure. The algorithm has the potential to cluster social network structure through maximum clique and sub grouping criteria. The algorithm has demonstrated impressive performance on test run using the standard clique sequence data particularly for the clustering of the instance named as Keller 6. However, *CACM* can be further improved with respect to standard benchmark results by modifying current clique sequence and pheromone deposition strategy. In this context of combinatorial optimization problem, the proposed algorithm is an initial effort to cluster the data in the basic premise of classic Ant Colony System.

REFERENCES

[1] Bullnheimer B., Hartl R.F., and Strauss C., "An improved ant system algorithm for the vehicle routing problem", Annals of Operations Research, vol. 89, 1999, pp. 319-328.
[2] Colorni A., Dorigo M., and maniezzo V., "Distributed optimization by ant colonies", in proceedings of ECAL'91, European Conference on Artificial Life, Elsevier Publishing, 1991, pp.134-142.
[3] Costa D. and Hertz A., "Ants can colour graphs", Journal of the Operational Research Society, vol. 48, 1997, pp. 295-305.
[4] Dorigo M., Maniezzo V., and colorni A., "The Ant System: Optimization by a colony of cooperating agents",IEEE Transactions on Systems, Man, and Cybernetics-Part B, 26(1), 1996, pp.29-41.
[5] Emaili K. Neshati S., Jamali M., M., and Abolhassani H."Comparing performance of recommendation techniques in the blogsphere", in Proceedings of ECAI'06 Workshop on Recommender Systems, Riva del Garda, Italy, August 2006.
[6] Fenet S., Solnon C., "Searching for maximum cliques with ant colony optimization", Applications of Evolutionary Computing, LNCS 2611, Springer, 2003, pp. 236-245.
[7] GamBardella Tailard L., E., and Dorigo, M., "Ant colonies for the quadratic assignment problem", Journal of the Operational Research Society, vol. 50, 1999, pp. 167-176.
[8] Hanneman A. and Riddle M., "Introduction to social network methods," online at http://www.faculty.ucr.edu/hanneman/nettext/, 2005.

Mohammad Al-Fayoumi was born in 1951, in Jordan. He received his BSc. (Mathematics) MSc. (Mathematics) and PhD in Computer Information Systems, from Jordan University, Amman Jordan and University of Bucharest, Romania in 1974, 1977, and 1981 respectively. He is currently associated as full professor in Middle East University for Graduate Studies Faculty of Information Technology, Department of Computer Information Systems. He was involved in different applied technical training programs on different software applications and programming tools in Jordan and aboard. Prof. Fayoumi has more than 50 research paper published to his credit encompassing different areas like modeling and simulation, information security & privacy, Discrete structures and algorithm analysis and design.

Prabhat K. Mahanti is Professor, Department of Computer Science and Applied Statistics (CSAS), University of New Brunswick Canada. He obtained his M.Sc. (IIT-Kharagpur) India, Ph.D. (IIT-Bombay) India. His research interests include Software engineering, software metrics, reliability modeling, modeling and simulation, numerical algorithms, finite elements, mobile and soft computing, verification of embedded software, neural computing, data mining, and multi-agent systems. He has more than 200 research papers, technical reports to his credit. Dr. Mahanti is also involved with several consultancy projects in and around North America.

Soumya Banerjee obtained his bachelor in engineering (B.E. (Hones)) degree in Computer Science from Govt. Engineering College, Maharastra, India in 1995. He has submitted his Ph.d thesis on Bio-inspired Algorithm and hybrid intelligence in 2008. He is currently a reader with the Department of Computer Science, Birla Institute of Technology, International Center, Mauritius. His research interests include Bio–inspired algorithm, Soft Computing, and Hybrid intelligent system.