

Modularity Based ABC Algorithm for Detecting Communities in Complex Networks

Thet Thet Aung and Thi Thi Soe Nyunt

Abstract—Network structure is a representation of interactions among actors. Transportation network, biological network and social networks have complex networks structure. Detecting community in complex networks is an important research to gain insight information. Effective optimization algorithms are needed in network community detection. Modularity based artificial bee colony (MABC) algorithm is proposed to uncover community in complex network. The proposed MABC algorithm is evaluated by Modularity and Normalize Mutual Information (NMI) metrics. Three real world datasets are used in the experiment. The proposed approach effectively detects community structure and produces noticeable good result than other previous algorithms in sample complex networks.

Index Terms—Artificial bee colony, community detection, modularity and normalize mutual information.

I. INTRODUCTION

Nowadays, complex networks analysis is important research field in big data's era. Community structure in complex network's topology is defined as for divide the complex network into a set of group which have common features such as the most relation together vertices. Community can be defined as a cluster of nodes which have denser intra connections and sparse inter connections.

Community detection can be described as module detection, graph partitioning and network clustering. Researchers use many types of community detection algorithms in different kind of networks. Most of them are based on clustering principles adapted to graphs, using divisive and agglomerative clustering, graph partitioning and optimization methods. Divisive clustering divides communities by considering nodes distance and agglomerative clustering merges groups by considering node similarity between them. Community detection in complex networks is NP-hard because it is optimization problem. If the numbers of nodes increase, whose computational complexity will arise exponentially. In optimization based method, quality function is used to measure the result quality of network community. The algorithm's efficiency depends on the search ability of the algorithm and quality function.

Optimization problems are solved by using nature inspired based swarm intelligence algorithms. The algorithms can

find optimal solution in many problems. This algorithm contains a numbers of possible solution. These solutions are numerically encoded and fitness function is also used to optimize on the solution space. Nidhi Arora presented a framework to implement natural intelligence based optimization strategies to detect communities in complex networks [1]. Recently applied nature inspired based algorithms for community detection include swarm intelligence based algorithm such as GA (genetic algorithm) [2] and PSO (particle swarm optimization) [3], and EA (evolutionary algorithm) [4].

The remaining parts are described as follows. Firstly, related research works of community detection algorithm is discussed and then introduce the background theory of paper such as artificial bee colony algorithm, community detection problem, evaluation metrics and solution representation schema in nature inspired based algorithms. Then, it presents the detail stage of the proposed works and experimental result on some complex networks. Final part includes conclusion and future work.

II. RELATED WORKS

A fast algorithm called fast greedy is proposed to detect communities in the complex network by Newman (2004). The algorithm is bottom-up approach and modularity is being optimized in a greedy fashion. In this algorithm each node has an own community and then communities are merged iteratively that yield the optimal modularity result. When the modularity result doesn't increase, it will be stop [5]. This method is fast and get optimal modularity result. Khaled Ahmed *et al.* [6] have proposed adaptive approach that is based on chicken swarm optimization algorithm(CSO). In CSO algorithm, as an objective function of the algorithm Modularity and NMI are used and then locus-based representation are used for solution representation. Bi-Objective Genetic Algorithm based on maximizing modularity and community score is presented by Rohan Agrawal (2011) to detect effective communities in networks. For solution representation, it uses one kind of array based representation as locus-based. Author used Modularity and NMI to compare other community detection algorithms.

Artificial bee colony swarm optimization and locus-based encoding schema are used by Hafez *et al.* [7] to detect community in social networks. They used many kinds of objective functions. The results of first type objective functions were minimizing and maximizing the second types functions.

Manuscript received October 20, 2019; revised January 13, 2020. This work is important for author research.

The authors are Cloud Lab Research Lab of University of Computer Studies, Yangon, Myanmar. (e-mail: thetthetaung@ucsy.edu.mm, thithi@ucsy.edu.mm).

All of the above community detection algorithm have achieved efficient community results on various complex networks. All of these algorithms use locus based solution representation. The locus based representation is used for solution encoding, decoding and representation results as a set of community. This representation needs to decode the result community. To reduce the decoding step, label based solution representation is used in the proposed algorithm.

In 2015, Dervis Karaboga [8] introduced Artificial Bee Colony algorithm to solve numerical optimization problem. It gets optimal result in many numerical optimization problem and clustering problem [9]. In this paper, ABC will be used to solve combinatorial problem such as community detection. Combinatorial optimization problem finds values for discrete variables such that to find the optimal solution for a given objection function. MABC algorithm use modularity as the fitness function of ABC. NMI is used to test the community detection algorithm’s accuracy but this metric can only be used in some datasets which have ground truth communities and modularity will also be used to judge communities’ quality. Experiments on three networks reflect the effective of MABC.

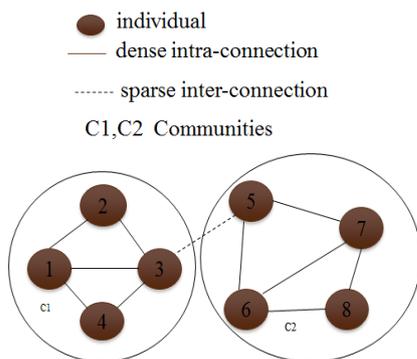


Fig. 1. Example community structure of a complex network.

III. COMMUNITY DETECTION IN COMPLEX NETWORK

A community in complex network means nodes within group are tightly connected and loosely connected to the outside groups. Community detection is the process by which nodes in a network are clustered based on connections between nodes. Complex network community detection can give valuable information to understand the structure of ties influences individual and its connections. Practical application areas are recommendation, social network analysis, link prediction and criminal detection. For this problem, input network is represented as graph structure and the output is the set of nodes groups representing communities. There are two types of community detection, disjoint community (node will be in at most one community or none) and overlapping community detection (nodes can belong to multiple community). In this paper, MABC algorithm is proposed to detect the disjoint community in complex network. A graph is denoted as $G = (V, E)$, in which V is the set of vertices and E is the set of edges. Fig. 1 shows the example of community structure present in a complex network. Identify set of communities such that $C = \{C_1, C_2, \dots$

$C_{nc}\}$, where nc is the total number of communities and $V = \cup_{a=1}^{nc} C_a$ and $\cap_{a=1}^{nc} C_a = \emptyset$;

IV. ARTIFICIAL BEE COLONY (ABC) ALGORITHM

In the last year, various nature inspired algorithms have been proposed to solve real world optimization problems in various fields. This nature inspired based algorithms can be inspired by evolutionary phenomena, the collective behavior of powerful intelligence technology, physical rules, and concepts related to humanity [10]. They got efficient results in solving various optimization problems. These algorithms are applied in data mining, information retrieval and engineering. Artificial bee colony (ABC) algorithm is a population based metaheuristic algorithm that mimics foraging behavior of artificial bees. ABC algorithm works with the set food sources in which each food source corresponds a possible solution to the optimization problem. Food source quality represents the nectar amount of these food source. The artificial bee colony algorithm contains three type of bees called employed bees, onlooker bees and scout bees. An Employee bee find a new food source that contains more nectar near a memorable food source. Then they share the information about these food source to onlooker bees. The onlooker bees select good food sources based on the probability which is associated with the nectar amount of food sources given by employed bees. If the quality of the food source is high, it is more likely to be selected by onlooker bees. Employed bee whose food source don't improve a predefined trial number, it becomes a scout and the food source is abandoned. And then, scout bee finds new random food source position [11].

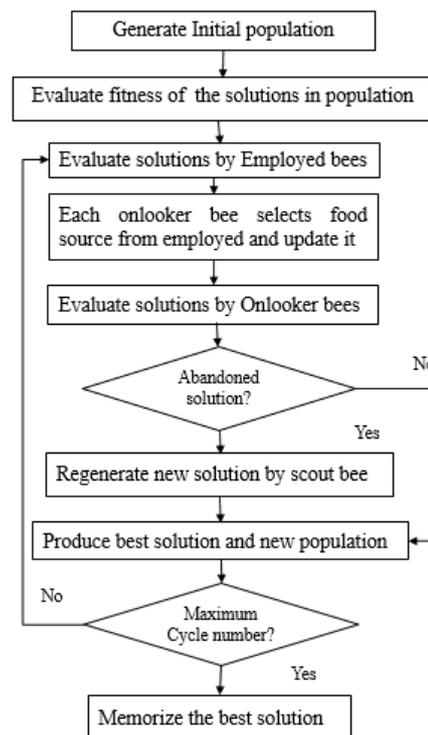


Fig. 2. Flow diagram of ABC algorithm.

These three bees’ processes are iterated until maximum cycle number is reached. In ABC algorithm, employee bees

and onlooker bees make exploitation and exploration is controlled by scout. Fig. 2 shows the flow diagram of the ABC algorithm.

V. EVALUATION METRICS

Complex network's community detection has become NP-hard optimization problem. It cannot be solved in polynomial time. It need evaluation metrics to get optimal communities result. Many metrics have been proposed to detect the quality of communities within a given network. Some metrics based on internal connections such as internal density, fraction over median degree, edge inside, average degree, and triangle participation ratio, external connection based metrics such as expansion and cut ratio, internal and external based metrics such as Maximum-ODF, Average-ODF, conductance, normalized cut and Flake-ODF and some metrics considering a network model as modularity, separability, density, cohesiveness and clustering coefficient. After finding the communities in the network using one community detection algorithm, the result community structure will be evaluated. The evaluation is easier when the network has actual or ground-truth community structure. At that time, ground-truth based validation metrics can also be used. Ground-truth based validation metrics are F-Measure, Normalize Mutual Information and so on. If the algorithm detects a community structure whose result is very similar to the actual community structure, the algorithm is accepted and used in other networks where the underlying true community is not available [12]. Modularity metric is used as the fitness function of MABC and it and normalize mutual information is also used to evaluate the result communities' quality.

A. Modularity

Modularity accesses the quality of result cluster, it finds the different between actual links among all vertex pairs in the same community, and the links present those pairs in a randomly created model network of the same network. The equation 1 is used to calculate the modularity Q .

$$Q = \frac{1}{2m} \sum_{ij} [A_{ij} - \frac{d_i d_j}{2m}] \delta(c_i, c_j) \quad (1)$$

where,

- $\delta(c_i, c_j)$ is 1 if i and j present in the same community label, else $\delta(c_i, c_j)$ equals to 0 .
- A is the adjacency matrix, $A_{ij}=1$ if there is a connection between i and j .
- d_i is the degree of node i and d_j is the degree of node j .
- m means the number of connections.
- c_i is the community number of node i , c_j is the community number of node j [13].

High quality of modularity indicates the good community structure. Q value is between 0 and 1. Generally, real world network modularity values between 0.3 and 0.7. This values indicate the presence of strong network modularity.

B. Normalize Mutual Information

Normalize mutual information (NMI) is used to measure

community detection algorithm's performance. It can be used to measure the similarity between the detected community and true community of the same network. For two network structures A and B of same network, NMI (A, B) can be calculated using equation 2 [14].

$$NMI(A, B) = \frac{-2 \sum_{i=1}^{C_A} \sum_{j=1}^{C_B} C_{ij} \log(\frac{C_{ij} N}{C_i C_j})}{\sum_{i=1}^{C_A} C_i \log(\frac{C_i}{N}) + \sum_{j=1}^{C_B} C_j \log(\frac{C_j}{N})} \quad (2)$$

where,

- N is the total number of nodes of the network,
- C is a confusion matrix.
- C_A is the number of communities in partition A and C_B is the number of communities in partition B.
- C_{ij} is the number of nodes present in community i of A (community j of B).
- C_i is the number of members in row i (or column j).

If the similarity of true and detected communities has greater value of NMI, the result communities will get better quality result. For truth and detected communities being same, the $NMI(A, B) = 1$ and if they are different communities, $NMI(A, B) = 0$.

VI. SOLUTION REPRESENTATION SCHEMAS

Solution representation is important for nature inspired based community detection algorithm. Two popular encoding schemes for the community detection problem are label-based representations and locus-based representations. In the two schemas each solution or individual is represented by a vector of D dimension in which D is the number of nodes in the given network. In label based solution representation, each solution is based on the indices values in the vector. The value in each index indicates to which community the node belongs. This value is an integer value in the range $\{1, 2, \dots, k\}$ in which k is the number of communities in the network. The label based encoding schema is defined as follow

1) S_a is a solution or individual, S_a^k is k^{th} position of S_a , and its' index value represent the community label of node k in the given network that is denoted by $L(k)$.

2) $S_a^2 = S_a^3 = 1$ means node 2 and node 3 contain in the community label 1. So, $L(2) = L(3) = 1$.

This representation automatically determines the number of communities and equals the number of unique labels, it can make to faster algorithm convergence and make code and decoding easier.

In locus-based representation schema, each locus in the vector corresponds to a neighbor node in the given network.

1) In each solution S_a , each index vale takes in the range $\{1, 2, \dots, n\}$ where n is the number of node in the network.

2) $S_a^3 = 2$ and $S_a^2 = 1$ mean node 3 has a link to node 2 and node 2 has a link to node 1 in the corresponding graph. This means that node 3 and 2 (node 1 and node 2) might be in the same community in the network.

The locus based solution representation have a major

advantage that automatically detect the community's number. Some drawbacks are it needs to decode the result community and it may have difficulty in evolution phase of nature inspired algorithms [10]. For MABC's food source representation, label based solution representation will be used.

VII. PROPOSED WORK

All the stage of proposed modularity based artificial bee colony (MABC) algorithm for detection communities in complex network are describe in this section. If the algorithm is designed based on modularity function, it is called a modularity based community detection algorithm, In ABC, the number of food sources are the number of solutions in the population. The food source with highest nectar amount represents a better solution of the optimization problem. The population initialization, modularity based fitness function of proposed algorithm and population search strategy are discussed as follow.

A. The Population Initialization

For the community detection problem, the input is undirected and unweighted network structure and the output is strongly connected community membership number for each node. Network is represented as graph structure as $G(V, E)$, V contains the set of n nodes and E is the set of m edges. In the initialization of solutions in the population, label based solution representation is used. In which, all elements in the solution is an integer value in the range $\{1, 2, 3, \dots, k\}$ in which k is the community number in the network. Each j^{th} value of food source specifies the cluster label they belonged. If i^{th} value of food source and j^{th} value of food source have the same community number, the value of $x_{ii}=x_{ij}$ will be same. The population initialization procedure algorithm is described as follow.

```

Algorithm 1 : The Population initialization
Input: Network structure as graph:  $G = (V, E)$ , Population Size:  $p$ ,  $\alpha < 0.3$ ,
number of community:  $NC$ 
Output: the set of solution in the population
Begin
  Generate a population of solutions  $SN = \{X_1, X_2, X_3, \dots, X_p\}$ , each food
  source  $X_i = (x_{i1}, x_{i2}, x_{i3}, \dots, x_{in})$  in which ( $n = \text{no. of vertices}$ ),  $i = \{1, 2, 3, \dots, p\}$ 
  For each food source in  $SN$  do
     $tc \leftarrow 0$ ;
    For each element in  $X_i$  do
       $x_i \leftarrow$  Random number of  $NC$ ;
    End for
    Repeat
      Randomly select a vertex  $x_{ik}$ ;
      label  $\leftarrow$  Random number of  $NC$ ;
       $x_{ik}$  and its neighbor vertices  $\leftarrow$  label;
       $tc \leftarrow tc + 1$ ;
    Until  $tc = (\text{int}) (\alpha * n)$ 
  End for
End
    
```

Algorithm 1 discusses about the initialization of the solutions in the population. Some large-scale network can't know the exact number of community. At that time, the proposed algorithm can also detect the acceptable community because label-based encoding schema can also

detect community number automatically. Prior knowledge of networks (the predefined number of community) is used to improve the performance of algorithm and to get more accurate result. And then, the proposed algorithm will try to get better modularity value. Fig. 3(a) shows a sample network diagram that contain 9 nodes and Fig. 3(b) is the example of label-based solution representation for each food source. In which, NodeID is the node name and CID means the community label/ community identification number. Community label can be get from the set of solutions in population initialization using algorithm 1.

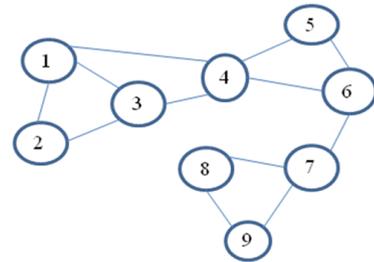


Fig. 3. (a) Sample network containing 9 nodes.

NodeID	1	2	3	4	5	6	7	8	9
CID	1	1	1	1	2	2	3	3	3

Fig. 3. (b) Label based representation of each food source.

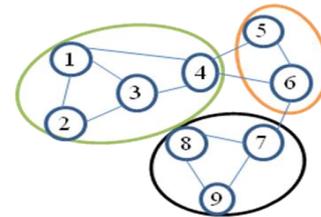


Fig. 3. (c) Community structure of the proposed food source.

B. Fitness Function

For swarm intelligence based algorithm, fitness function is important factor of various optimization problem. Modularity Q is used as the fitness function of MABC algorithm. This function is described in equation (1).

C. Search Strategy

Employed bees find new food sources which have the more nectar amount than old food source. Firstly, a neighbor food source is randomly selected. Then randomly select one node in the old food source and select other nodes in these food source which have same community number as selected node. Exchange these genes value between original food source and neighbor food source. This strategy is similar to the crossover operation on genetic algorithm. Choose one food source from the two child food sources; it has more nectar amount than other. Then keep the best one between current and candidate solutions using greedy selection. Example of creating new food source is described as follow. One of the neighbor food sources (Fig. 4(b)) in the set of solutions is selected. Then randomly selected a node is in original food source (Fig. 4(a)). Chosen node is node 3. Node 3 have same community label with nodes 4 and 5. Swap these genes values between two food sources, the two child food sources are $\{111223333\}$ and $\{112223333\}$. Evaluate fitness of two

child food source, results are 0.3194 and 0.22569. Better food source {111223333} is chosen as candidate solution. Then more nectar amount solution is kept between original and candidate solution for new food source. The original food source fitness is 0.22569 and candidate is 0.3194. So, {111223333} with fitness 0.319 is chosen as new food source.

NodeID	1	2	3	4	5	6	7	8	9
cid	1	1	2	2	2	3	3	3	3

Fig. 4. (a) Original solution/ food source.

NodeID	1	2	3	4	5	6	7	8	9
cid	1	1	1	2	2	3	3	3	3

Neighbor solution/ food source.

Fig. 4. (b)

Then the probability of each food source is calculated using equation 3.

$$P_i = (0.9 \times (\text{fit}_i / \text{maxFit})) + 0.1 \quad (3)$$

where,

- P_i is the probability of food source i ,
- fit_i is food source i 's fitness result
- maxFit is the maximum fitness value in the population

Then, onlooker bee chooses a food source depend on the quality of food source given by employee bees. After choosing a food source, it selects neighbor food source to create new food source as employee. If the corresponding food source cannot improve, it will be abandoned. Scout bee generates a new random food source as solution initialization. In each iteration, the proposed algorithm keeps the best solution found in the set of solutions. Those three bees process is repeated until the maximum number of iteration is reached.

Finally, the proposed algorithm produces the community structures of the given network as an output. In example network gets three communities. Community label 1 has node 1, 2 and 3, community label 2 has node 4, 5 and 6 and community label 3 has node 7, 8 and 9 and that are produce as an output. Its fitness value is 0.41319. Fig. 5 displays the final output community result for the sample network with optimize modularity result.

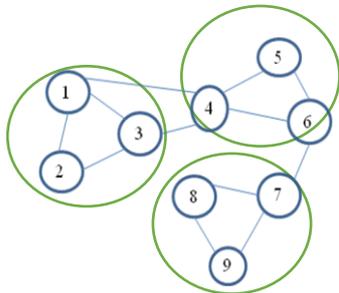


Fig. 5. Output community structure.

VIII. RESULT ANALYSIS

The experiments are implemented on a laptop with Core i7, 8GB of RAM and 64-bit Window operating System. Population size SN= 50, maximum cycle number = 100, abandon limit = 20 and swap operator rate=0.3 are set as the parameter of the algorithm. Three complex networks such as

Zachary karate club, bottlenose dolphins, and college football networks are used to show the effectiveness of algorithms. These complex networks are represented as graph structures, each contains different number of vertices and edges. The evaluation of the efficiency of MABC algorithm is compared with the result of other algorithms. NMI is used to measure performance of algorithm.

Zachary's karate club is a complex network that show the inside relationships of a karate club. dataset contain 34 members of club and 78 links between members. The club is divided into two groups because of argument between two teachers [15]. The graph structure of Zachary karate club dataset is described in Fig. 6.

The ground-truth communities and MABC algorithm's detected communities result of karate club dataset is shown in Table I and Table II as a sample. The proposed algorithm finds above 97% correct communities structure. In result communities, node 10 places in wrong community. There is ambiguous for node 10 because node 10 has connection to node 3 in community number C1 and connection to node 34 in community number C2.

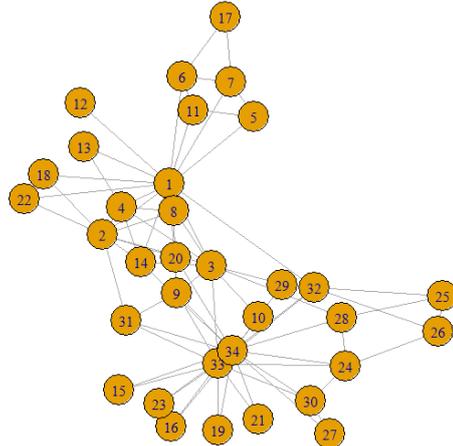


Fig. 6. Network structure of Zarachy karate club.

TABLE I: GROUND TRUTH COMMUNITIES AND ITS MEMBERS IN ZACHARY KARATE CLUB DATASET

Community Number	Vertices	Total
C1	1,2,3,4,5,6,7,8,11, 12,13,14,17,18,20,22	16
C2	9,10,15,16,19,21, 23,24,25,26,27,28, 29,30,31,32,33,34	18

TABLE II: DETECTED COMMUNITIES RESULT USING MABC IN ZACHARY KARATE CLUB DATASET

Community Number	Vertices	Total
C1	1,2,3,4,5,6,7,8,10, 11,12,13,14,17,18,20,22	17
C2	9,15,16,19,21,23,24,25,26,27,28,29,30,31,32, 33,34	17

Dolphin network is the network of 62 dolphin lives in Doubtful sound, New Zealand. It contains 159 link and split into two large groups [16]. Football network is game network of American Colleges. Nodes represent teams and edges represent matches. It contains 115 teams and 613 games that grouped in 12 communities [17]. Table III describes about three datasets information.

TABLE III: THREE REAL-WORLD NETWORKS INFORMATION INCLUDE VERTEX NUMBER, EDGE NUMBER AND COMMUNITY NUMBER

Network	No. of vertices	No of edges	Ground truth Community Number
Zachary Karate	34	78	2
Dolphin	62	159	2
Football	115	613	12

TABLE IV: COMPARISON OF VARIOUS ALGORITHMS BASED ON MODULARITY FOR THREE TYPES OF COMPLEX NETWORKS

Algorithm	Modularity value of dataset		
	Karate	Dolphin	Football
Fast greedy	0.38	0.49	0.54
GA	0.38	0.49	0.577
BOCD	0.419	0.507	0.577
EA	0.38	0.46	0.56
MABC	0.371	0.38	0.59

TABLE V: COMPARISON OF VARIOUS ALGORITHMS BASED ON NMI FOR THREE TYPES OF COMPLEX NETWORKS

Algorithm	NMI value of dataset		
	Karate	Dolphin	Football
Fast greedy	0.69	0.55	0.70
GA	0.692	0.573	0.762
BOCD	0.695	0.615	0.878
EA	0.83	0.78	0.76
MABC	0.837	0.753	0.916

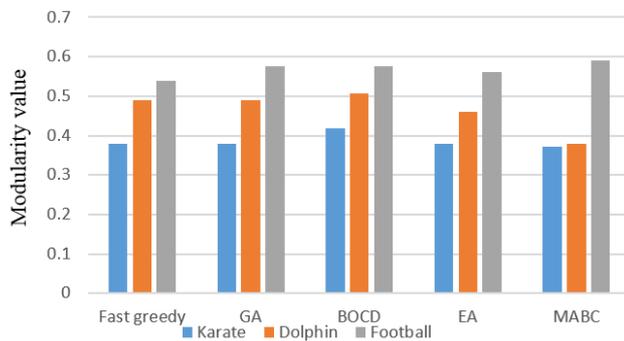


Fig. 7. Modularity values comparison using five community detection algorithms.

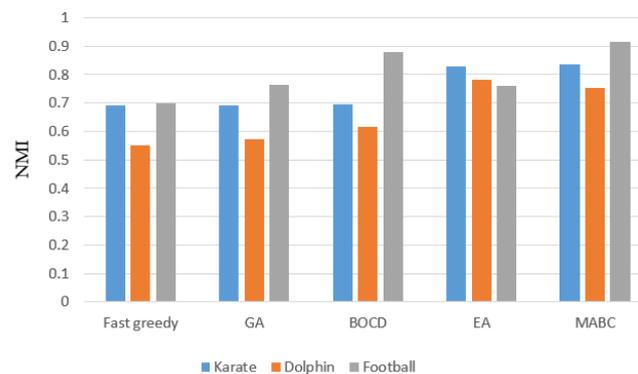


Fig. 8. NMI results comparison using five community detection algorithms.

The result obtained by MABC is compare with the results of previous community detection algorithms such as fast greedy, Evolutionary algorithm (EA), Genetic algorithm (GA), Bi-Objective community detection (BOCD) based on modularity and NMI quality metrics. Modularity results for real world datasets are compared in Table IV. Table V

describes the comparison of algorithms on Normalize Mutual Information results for real world networks. In Table IV, BOCD gets better modularity results in karate and dolphin dataset than other algorithm but its' result community number is little different with ground truth community number. BOCD algorithm get effective results in small complex networks.

MABC use ground truth community number of three networks that can makes the algorithm more targeted and can improve the stability and accuracy of detecting community. MABC get better modularity result in largest dataset football. In performance comparison table, MABC algorithm was better than the results of other algorithms in karate and football datasets. The proposed algorithm gets the highest NMI and Q values in football dataset among the algorithms. Fig. 7 and Fig. 8 show Modularity and NMI compare results chats using five community detection algorithms.

IX. CONCLUSION

One of the swarm intelligence based algorithms called modularity based ABC (MABC) algorithm is proposed. ABC algorithm is widely used in various type of numerical optimization problem. It can also be used in combinatorial optimization problem as community detection problem. The proposed algorithm takes modularity as a fitness function. Modularity and NMI are used to measure the quality and performance result of algorithm. Some conventional community detection algorithms and nature-inspired algorithms are used to compare results. The experimental results show that MABC can detect communities that is closer to ground truth in large network. Future development is focused on distributed community detection on spark framework to reduce time complexity and to get more effective and accurate results. Multi-objective community detection algorithm will also be considered to improve more quality of results.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

Author proposed modularity based artificial bee colony algorithm to find community structure for unweighted and undirected networks. ABC is proposed to solve numerical optimization problem. It got effective results in these problems. So it is used to solve combinatorial optimization problem as community detection. For solution representation, label propagation based initialization is used. It makes to faster convergence of algorithm. Use the modularity based ABC to find optimal community structure on various complex networks.

ACKNOWLEDGMENT

I would like to thank a lot to my supervisor and all my teachers for their encouragement, helpfulness and

recommending the research. I specially thank to my parents, my hubby and friends who have supported me with their kindness, love, support and encouragement.

REFERENCES

- [1] N. Arora and H. Banati, "A framework to mine communities using nature inspire algorithms," in *Proc. ICRICN*, 2016.
- [2] R. Agrawal, "Bi-objective community detection (BOCD) in networks using genetic algorithm," *Contemporary Computing*, pp. 5-15, Springer Berlin Heidelberg, 2011.
- [3] M. Gong, Q. Cai, X. Chen, and L. Ma, "Complex network clustering by multiobjective discrete particle swarm optimization based on decomposition," *IEEE Transactions on Evolutionary Computation*, vol. 8, no. 1, pp. 82-97, 2014.
- [4] S. Bilal and M. Abdelouhab, "Evolutionary algorithm and modularity for detecting communities in networks," *Physica A*, 2017.
- [5] M. E. Newman, "Fast algorithm for detecting community structure in networks," *Physical Review E*, vol. 69, no. 6, p. 066133, 2004.
- [6] K. Ahmed, A. E. Hassanien, E. Ezzat, and P. Tsai, "An adaptive approach for community detection based on chicken swarm optimization algorithm," *J. Pan Etal.*, 2017.
- [7] A. I. Hafez, H. M. Zawbaa, A. E. Hassanien, and A. A. Fahmy, "Networks community detection using artificial bee colony swarm optimization," *Advances in Intelligent Systems and Computing*, vol. 303, pp. 229-239, 2014.
- [8] D. Karaboga and B. Basturk, "Artificial bee colony (ABC) optimization algorithm for solving constrained optimization problems," *IFSA*, pp. 789-798, 2007.
- [9] D. Karaboga and C. Ozturk, *A Novel Clustering Approach: Artificial Bee Colony (ABC) Algorithm*, 2009.
- [10] M. Guendouz, "A discrete black hole optimization algorithm for efficient community detection in social networks," *IGI Global*, chapter 9, 2018.
- [11] D. Karaboga, "An idea based on honey bee swarm for numerical optimization," Technical Report-TR06, Department of Computer Engineering, Engineering Faculty, Erciyes University, 2005.
- [12] T. Chakraborty, A. Dalmia, A. Mukherjee, and N. Ganguly, "Metrics for community analysis: A survey," *ACM Comput. Surv.*, vol. 50, no. 4, August 2017.
- [13] A. Arenas, A. Fernandez, and S. Gomez, "Analysis of the structure of complex networks at different resolution levels," *New Journal of Physics*, vol. 10, pp. 1-23, 2008.
- [14] H. U. Baofang, "A cultural algorithm based on artificial bee colony optimization for community detection in signed social networks," in *Proc.*

10th International Conference on Broadband and Wireless Computing, Communication and Applications, 2015.

- [15] W. Zachary, "An information flow model for conflict and fission in small groups," *J. of Anthropological Research*, vol. 33, pp. 452-473, 1977.
- [16] D. Lusseau, K. Schneider, O. J. Boisseau, P. Haase, E. Slooten, and S. M. Dawson, "The bottlenose dolphin community of Doubtful Sound features a large proportion of long-lasting associations," *Behavioral Ecology and Sociobiology*, vol. 54, pp. 396-405, 2003.
- [17] M. Girvan and M. E. J. Newman, "Community structure in social and biological networks," in *Proc. Nat. Acad. Sci.*, vol. 99, no. 12, pp. 7821-7826, 2002.

Copyright © 2020 by the authors. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited ([CC BY 4.0](https://creativecommons.org/licenses/by/4.0/)).



T. T. S. Nyunt has got the B.Sc Physics(Hons:) degree from Yangon University in 1994, she got the master of information science (M.I.Sc.) degree and Ph. D (IT) from UCSY in 1998 and 2004 respectively. She is currently working as a professor and the head of department in Faculty of Computer Science, UCSY.

Her research interests include knowledge & software engineering, database, computer graphics, big data analytics, IoT and artificial intelligence and neural network.



computing.

T. T. Aung is a Ph.D candidate and doing her research at Cloud Computing Lab, University of Computer Studies Yangon, Myanmar (UCSY). She got B.C.Sc (Hnos:) and M.C.Sc from Computer University in 2009 and 2011, respectively. She joined the University of Technology, Pyin Oo Lwin, Myanmar since 2010. Now, she is an assistant lecturer at Computer University. Her research interests currently include social network analysis, big data mining, machine learning and evolutionary