

Identification of Influential Customers in Social Network based on BFO

Monica Sood, Preetpal Kaur

Abstract-In this paper we have proposed the implementation to identify the most influential customers in the social network. In Social network, different kind of people are communicate with each others and exchange their ideas, views about any products ,item or person. Any company or organization can increase the revenue of their product if the company identify such a customer in the social network that has the ability to influence to others in the social network . Influential customers whose connections, messages and opinion strongly influence to others in the specified social network .Such customers in the social network such as friendster, facebook can be identify by Swarm Intelligence algorithm-BFO. BFO has the strength to produce the optimal solution from the number of solution. We have followed the dataset from the social network site to find the most influential customers in the network. Bacterial Foraging Optimization(BFO) is the used to identify the optimal node in the social network .The evaluation based on the number of nodes with the highest simulation influence value to identify best nodes. Influence value based on number of friends, followers, number of messages reply, likes. The simulation influence point ratio is use to consider as the simulation influence value to identify the popular nodes in the social network with the help of optimized algorithm-BFO.

Keywords: BFO, Influential nodes, Optimized nodes, Swarm Intelligence

I. INTRODUCTION

Social networking sites like Friendster, face book are provides a platform not only to communicate different kind of people but also effective for the success in the business .In social network, from the initial adopters with its further communication links and influence value used to find the most influential customers.BFO algorithm is used for the optimal solution. Swarm Intelligence is the hot topic in different areas such as Data Mining, Image Processing, Wireless Networking, Management. Such as ACO, FSS, PSO. One of the widely used algorithm of SI is BFO. For Identification of users in the social network is perform by different researchers by using different methods like diffusion process ,Greedy methods and swarm intelligence algorithm ACO[3,4].The optimal solutions with fast speed is search by the latest algorithm is BFO.

A. Bacterial Foraging Optimization Algorithm

Foraging is the biggest challenge for the natural insects. The optimization food density on the best position or space is the goal of the natural insects.

Manuscript published on 30 April 2013.

*Correspondence Author(s)

Monica Sood, Assistant Professor, Dept. of CSE/IT, Lovely Professional University, Phagwara, Punjab, India.

Preetpal Kaur, M.Tech Student, Dept. of CSE/IT, Lovely Professional University, Phagwara, Punjab, India.

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an open access article under the CC-BY-NC-ND license <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

SI also suggest the behavior of the natural insects that can implement to solve the real world application. BFO is the optimization algorithm invented by Kevin Pasino in 2002 on the behavior of E.Coli bacteria for foraging a food[2].Like E.coli bacteria to search optimal food in the netruints.BFO works as same to find the optimal solution to find the number of solution to solve the real world problem. In BFO, the optimal solution will be find by initializing the chemo-taxis step with average value. Reproduction step in which only the healthier bacteria(solution) will be consider from the number of solutions. Elimination and dispersal step to eliminate the solutions that are not trapped in the nearest values of the fitness function

B. Influential Nodes in Social Network

In the social networking; the number of nodes are connected indicate individuals. Assume a social network as $G=(N,E)$; where

G=Graph of the social network,

N=Total number of nodes,

E=Total number of Edges

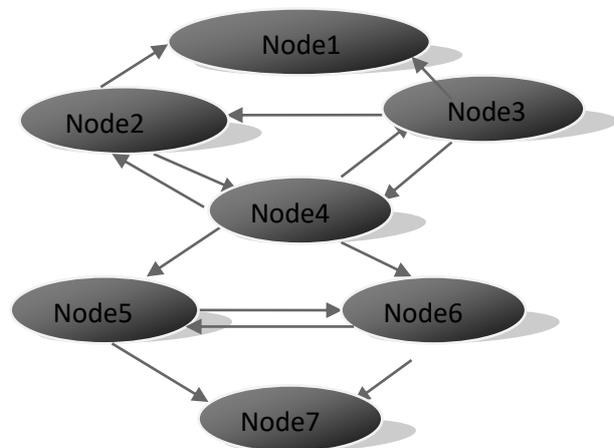


Figure1.1 :Social Network

Figure1 shows the nodes that are connected with each others on the basis of degree, distance and simulation influence value approaches[3]. Influence value can be identified by the number on links, strong ties-up or maximum number of links or friend list.

II. PROPOSED ALGORITHM

Bacterial Foraging is a population based algorithm .In this research, each node (customer or user) of the social network is considered as bacteria and simulation influence value is a food.



Need to identified the nodes that has a highest influence value with the simulation fitness value. According the dataset of the SN website ,we have 46 nodes and 879 connections with different simulation influence values. The purpose of this algorithm to find the optimal solution with maximum influence value. All the nodes are categorized in two parts on the basis of the health status (required value) –Popular nodes and Unpopular nodes. The nodes that has the maximum influence value will be consider as popular nodes others are unpopular. On the set of popular nodes ,BFO select the most popular or optimal nodes according to its property.

- BFO consists of the three principle of steps:
- Chemo-taxis
- Reproduction
- Elimination and dispersal

A. Chemo-taxis

Chemo-taxis initialize the number of the steps repeat to find the nodes whose influence value is near or equal to the fitness value. In this case, Swim operation perform otherwise Tumble if values varies in high range.

B. Reproduction

With the reproduction step, the highest value nodes are selected .The first half consider the top highest value nodes.

C. Elimination and Dispersal

The lowest second half of the nodes are eliminated by the number of elimination steps. The top highest nodes overlap on the lowest nodes of second half and performed by the dispersal steps. The process continue with iterations until the popular nodes achieved.

III. METHODOLOGY

To achieve the optimal nodes selection from the dataset of the social network;BFO algorithm is used to perform the selection of the nodes and find the best nodes from the social network.

A. Select random datasets from the social network

BFO select the specified population (datasets) from the social network that random datasets are the initial adopters or users to detect node- to-node connection .The position of the random datasets and simulation influence value of the nodes are also displayed in the output.

B. Find fitness value

All the nodes put in the fitness functions with the specified parameters to find the fitness value. That fitness value becomes our influence value that is used to compare with all the nodes.

C. Optimal nodes selected by different steps of BFO

After select the random datasets; different steps of BFO such as chemo-taxis, reproduction and elimination-dispersal performed to find the optimal solutions(nodes) from the number of solutions .The optimal nodes are selected with maximum simulation influence value and strong degree of connections.

D. Effect of simulation influence value in graph

The effect of simulation influence value perform well to find the influential nodes as compare to find in the

distance parameter. This is shown in the graph.

E. Simulation

The proposed algorithm is implemented in Matlab. The algorithm represent the best nodes from the number of nodes. The random nodes connection sets shown in table

Node1	10	29	22	19	38	28	14	18	25	38
Node2	4	36	15	36	34	44	22	2	2	13

3.1.

Table 3.1: Selected 10 random datasets of nodes

The following equation (1) is used to apply check on dataset fetch of the connected nodes.

$$\text{Chk}=\text{find}(\text{node1}=\text{main_data}(:,1) \ \& \ \text{node2}=\text{main_data}(:,2)) \tag{1}$$

The Equation (2) is used to return the result that are connected nodes.

$$\text{out}=\text{main_data}(\text{chk},3) \tag{2}$$

On the basis of the Equation1 and Equation2 ,the fitness function is implemented that match with influence values and return optimized nodes by follow the steps of BFO.Each node has different influence value.The highest value match with fitness value node selected.The best nodes are shown as in table 3.2.

Table 3.2: Final best nodes

19
36

The flow chart of proposed algorithm shown in Fig3.1 .

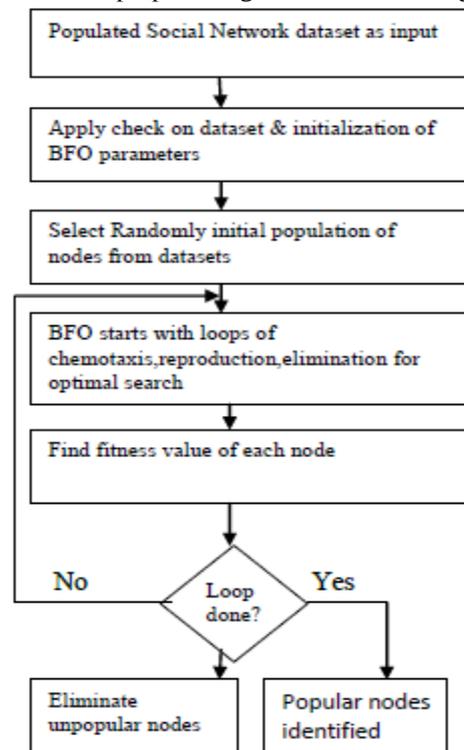


Figure 3.1 : Flowchart to identify optimal nodes in Social Network

The Pseudo code of the proposed algorithm is given in figure3 as follow:

- Step 1: Take the dataset of social network as input.
- Step 2: Apply check on the size of the dataset.
- Step 3: Initialize parameters of BFO S= Total number of nodes Nc=Total number of chemo-taxis steps(Repeat) Ns=Limit of swim(change) Nre=Number of reproduction steps Sr=Number of splits
- Step 4: Determine the position of the random dataset(ie node1,node2).
- Step5: Calculate the fitness value of the fitness functio.
- Step6: Calculate optimized fitness value and previous fitness value store.
- Step7: Repeat steps 5 & 6 until loop complete for reach at the optimized result.
- Step8: Final best nodes with highest influence value selected determined.
- Step 9: Stop.

- 2012, Pune-ISBN: 978-93-82208-18-1.
- 2. Passino Kevin, "Biomimicry of Bacterial Foraging:for Distributed optimization and control", Department of Electrical Engineering, The Ohio State University, 2002 Neil Avenue ,USA 0272-1708/02©2002IEEE.
- 3. Yang Wan and Weng Shi ,"Application Of ACO to influence Maximization Problem", International Journal of Swarm Intelligence and Evolutionary Computation Vol. 1 (2012), Article ID 235566.
- 4. D. Kempe, J. Kleinberg, and E. Tardos, "Maximizing the spread of influence through a social network", in Proc. of the 9th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '03), Washington, DC, 2003, 137–146.
- 5. Daniel M. Romero , Wojciech Galuba, Sitaram Asur, Bernardo A. Huberman ,"Influence and Passivity in Social Media", ACM 978-1-4503-0637- 2011, Hyderabad, India 113.

IV. RESULTS

Fig4 shows the graph of the performance that depict the best performance done by the simulation influence value. The identification of the nodes by distance and simulation influence value of the overall structure –simulation influence value works well. It represent the best results of the datasets. With the help of BFO optimal results found with simulation influence value.

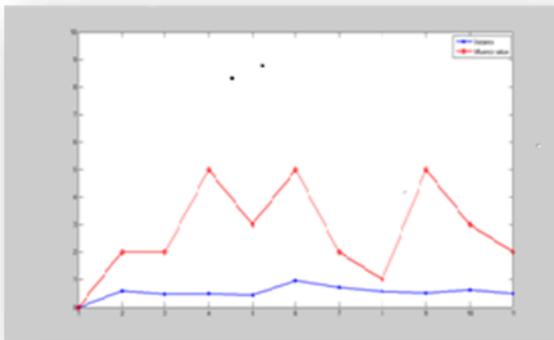


Figure 4.1. Graph of comparison of approaches of distance and Influence value.

V. CONCLUSION

After search results by following BFO algorithm on the application of social network, the following points is concluded as BFO is more feasible and efficient algorithm than ACO algorithm. Fast and optimized results are achieve and nodes selection is easy by BFO because of the fast conversion property of BFO algorithm. Social networking is not the way for communication of friends or exchange photos ,messages but it is also provides the best platform for the business organization to advertise of their product online. It is possible if the most popular person identify in the networking .For this purpose BFO algorithm perform well and Influence values are the best parameters to find the optimized nodes instead of follow the simulation of number of in degree. Company’s revenue will be increase by identifying the optimized customers in the initial stage.

REFERENCES

- 1. Chaudhari Chaitrali and Chaudhari D ,"Application of Swarm Intelligence Algorithm for solving optimization problems", ICEECS