Complex Systems and Networks
Project Report
To Find Maximum Clique Using Preferential Attachment Based On Ant Colony Optimization

Omar Rawashdeh
Harshal Raut
Utsav Shah

April 20, 2017

Abstract

As the size and number of online social networks are increasing day by day, social network analysis has become a popular issue in many branches of science. In this paper, we propose a link prediction method using preferential attachment based on the ant colony optimization. Interaction between users in online social networks plays a key role in social network analysis and various domains such as sociology, anthropology, information science, and computer sciences. One on important types of social group is full connected relation between some users, which known as clique structure. Therefore finding a maximum clique in the dynamically changing complex network like social network has become an essential aspect of social network analysis, such as privacy protection, citation analysis, cohesive subgroup analysis et al. With the development of big data, the mass of nodes in the graph and complexity of analysis set a higher requirement for solving the maximum clique problem. In the algorithm, artificial ants are employed to travel on a logical graph. Each ant chooses its path according to the value of the pheromone and heuristic information on the edges.

Contents

List of Figures

1 Introduction

1.1 Goals of the Project

To find the maximum clique in a complex network consisting of multiple interconnecting of nodes by edges amongst the cliques formed in a complex network viz Social Network. One of the emerging topics in social network analysis is link prediction as the social network keep increasing by the day. Prediction of a new connection or link between two nodes based on attributes of existing nodes and links in the graph is called link prediction. This is done using preferential attachment.

1.2 Motivation

Complex networks surround us everywhere and play a major role in our daily life. From social networks and biological networks such as metabolism networks and ecosystem networks to brain networks such as connectome.

Studying the properties of such networks will help us understand them better. Furthermore, studying these properties in real life situation where the structure of the complex network is continuously changing by time, increases the difficulty of the task.

Finding cliques (or maximum clique) in graphs has many applications, including locating communities in social networks and solving the graph coloring problem[1]. Using the knowledge learned from the course, we want to simplify the intricate tasks and prolifically implementing it for applications such as community tracking in a network, recommender systems in domains like music genres, movies, books, and shopping and more. Moreover, it has applications in divergent fields including coding theory, geometry, the identification of faulty processors in multi-processor systems, computer vision and printed circuit board design[2].

1.3 Approach

The basis of the project is Ant Colony Optimization. A network has been created where it keeps growing exactly like the real-world social network. With the amount of pheromones on the edges between the nodes, the clique is formed. The maximum clique is formed as the path followed is one with more pheromone deposition.

2 Background

Only few studies were made on the evolution of social networks on microscopic level. That is why the authors are studying the evolution of social networks by mining for frequent local structural changes patterns by time. they show few examples of local rules extracted from real networks (Digital Bibliography and Library project - network of authors and their collaboration work)[3]. They give the definition "link prediction" as an estimation of the probability that two nodes will form a link between them in the future. following that definition, they explained how link prediction can be calculated based on features extracted from the link structure in a snapshot at a given moment of time[2]. however the authors proposed a different approach that made a better prediction, by introducing scores for each embedding, which based upon a better prediction can be made.

One paper explains why studying the online social networks should be approached in a different way than the regular social networks (it should be approached as a CAS - Complex Adaptive System). It discusses the two main characteristics of complex systems, Self-organization and Emergence, and how these characteristics apply to online social networks[4]. The authors of this paper talk about the two lines of research on complex networks, the study of the topology dynamics of such networks, and the study of the dynamics on a network with static topology (non changing topology). They introduce the concept of adaptive networks (networks in which the dynamics on the network affect the topology and the topology affects the dynamics back), the models used to study the development and evolution of adaptive networks, and the methods used to model social networks dynamics[5].

The authors of another paper start their paper by talking about how the different popularities of individuals results in a new concept "Network Centrality" to emerge, due to what is called preferential attachment (PA), which means new nodes prefer to connect to old nodes with higher degree[6]. They report that even though the idea of PA was made clear as a result of previous studies, yet, there is no clear explanation on how the PL mechanisms affect the network features. The authors here, performed a case study using a dataset from a social network (Wealink), they found that Wealink network is scattered into unconnected subgraphs, and that the data shows a power low feature in the degree distribution for Wealink. they state that such behaviour can be a result of a linear model. they introduce a generalized model in which the model considers the possibility that two new nodes can forms a link between them.

Another paper presents an alternate method that uses Ant Colony Optimization (ACO) algorithm and Particle Swarm Optimization (PSO) algorithm to attain better results. The proposed method is a relative enhancement of the standard ACO algorithm and the simulated results are compared to standard social network datasets. Now, the ACO algorithm finds a solution to complexities involved in finding the fastest route between point A and point B, simulating the quickest path between them[7]. The PSO algorithm on the other hand, uses iterations by creating initial particles and assigning initial velocities to them, and evaluates the objective function at each particle location to determine the best function value and location, only to choose new velocities based on the current velocity to keep updating particle locations. Both, the ACO and PSO algorithms are explained in the paper to understand their hybrid algorithm to find a maximum clique in social networks. The newly introduced algorithm overcomes the traditional high complexities involved that significantly adds to the volume of calculations. The proposed algorithm places some ants initially on the graph to follow paths to find maximum clique. After the evaporation of the pheromones, the proper path is updated based on the amount of pheromone on the edges using PSO algorithm and this procedure is repeated until the optimum clique is obtained on the graph.

One other paper explains and depicts the intricate analysis of Improved Ant Colony Algorithm to find the maximum clique in the social networks. Various algorithms like Annealing Algorithm (Ant-Clique Algorithm) and PSO-ACO (Particle Swarm Optimization and Ant Colony Algorithm) were developed and modified to get optimal solution[8]. Although, Annealing Algorithm and PSO-ACO speeded up pace of constructing the clique, but the convergence rate needed to be improved. Dynamical adjustments of these parameters resulted into improvement in convergence rate and ability of searching the optimal solution of algorithm. However, during experimenting with the complex datasets like social networks dataset these algorithms had a drawback of not falling within the range of local optimum. Inferring, from these analyses the authors concluded that the performance of algorithm can be further increased and can be optimized, developing a modified and efficient algorithm called Improved Ant Algorithm based on mainly two aspects. Firstly, selecting a node as per different conditions by means of probabilistic approach including roulette wheel and inverse roulette wheel strategy. Secondly, Local Improvement of the Clique by updating cliques. The results of this algorithm produced accurate solutions and improved results experimentally to find maximum clique, but did not simplified methods to find global optimum.

3 Methods

Created an algorithm to generate a complex network that we can use to study its function and structure (complex network properties). Optimal usage of algorithms that are associated with complex systems (such as swarm algorithm or ant colony optimization) to study the complex network generated by the previous step. There are many platforms to implement algorithms and currently we are working on ideal method. Java code using the graphic libraries, and MATLAB to run simulations and plots.

An ant will follow the pre-existing path of pheromone to form a clique. This clique won't be entered by another ant. Clique rolling is not used because we are only focused on finding the maximum clique in the network. We have trouble with synchronizing different networks. Design of our approach, for each agent, at each iteration will check if there is a nearby position. If there is, a link or edge to be checked with small amounts of pheromone. Once the agents move to the node where the highest priority agent will form the maximum clique. Every edge will be treated like a newly added edge and will find clique at both ends. And it'll check for any cliques which need to me merged. If it is, then it merges, and if not, then it'll move to the next stage. We try to find a newly formed clique by three points for instance A, B, and C. We check if there exists edges between all of these points, and if it exists, it'll contact the hive.

The probability that node i with degree k_i is chosen can be expressed as:

$$\prod(k_i) = \frac{k_i^{\beta}}{\Sigma_j(k_j^{\beta})}$$

Compute the probability $\prod(k)$ that an old node of degree k is chosen, and it is normalized by the number of nodes of degree k that exist just before as:

$$\prod(k) = \frac{\Sigma_t[e_t = \wedge_v^k(t-1) = k]}{\Sigma_t|u : k_u(t-1) = k|}\alpha k^{\beta}$$

With probability p, we add a new node with on edge that will be connected to a node already present in the network. The probability that the new node will be connected to old node i with degree k_i is

$$\prod(k_i) = \frac{k_i}{\Sigma_j k_j}$$

With probability q = 1-p, we add one new edge connecting two old nodes. The two endpoints of the edge are also chosen according to linear preference.

After t time steps the model leads to a network with mean number of nodes N(t) =mθ +pt. For large t, N $\approx$ pt.

And the total degree of the network k all(t) =2eθ +2t θ 2t.Applying mean-field approach for user i, we obtain

$$\frac{\delta k_i}{\delta t} = p\frac{k_i}{\Sigma_j k_j} + 2q\frac{k_i}{\Sigma_j k_j} = \frac{p+2q}{2t}k_i$$

Thus, the probability density or P(k) is:

$$P(k) = \frac{\delta(k_i < k)}{\delta k} \alpha k^{\frac{p-4}{p=2}}$$

The ant colony algorithm for the maximum clique can be improved mainly from two aspects[8].

- Selecting Node According to Different Strategies: The basic ant colony algorithm employs the l method to select node according to the probability of each node, this makes the probability of the node with high pheromone being selected increasing, leading to the algorithm easily falling into local optimum in the later period.

- Local Improvement of the Clique: The proposed method for selecting node increases the diversity of solution, and yet slows down the convergence speed simultaneously. In order to conquer this shortcoming, we plan to improve the clique in a local manner. Local improvement of the clique is to further search around the clique constructed by ant to get a locally optimal one . This local improvement of the clique is carried out after an ant constructs the clique, but before the pheromone is updated.

4 Results

The following plot is a growing network of a typical social network. It can be observed from the first plot of the image that, the network consists of seven nodes. With every iteration, the network grows a certain amount. The second plot of the image shows that the network has grown who now has twenty-one nodes. The third plot of the image resembles how an actual social network grows with multiple iterations. This plot has about forty-four nodes in the network. The increase of the network from seven to forty-four nodes shows us the growth rate.

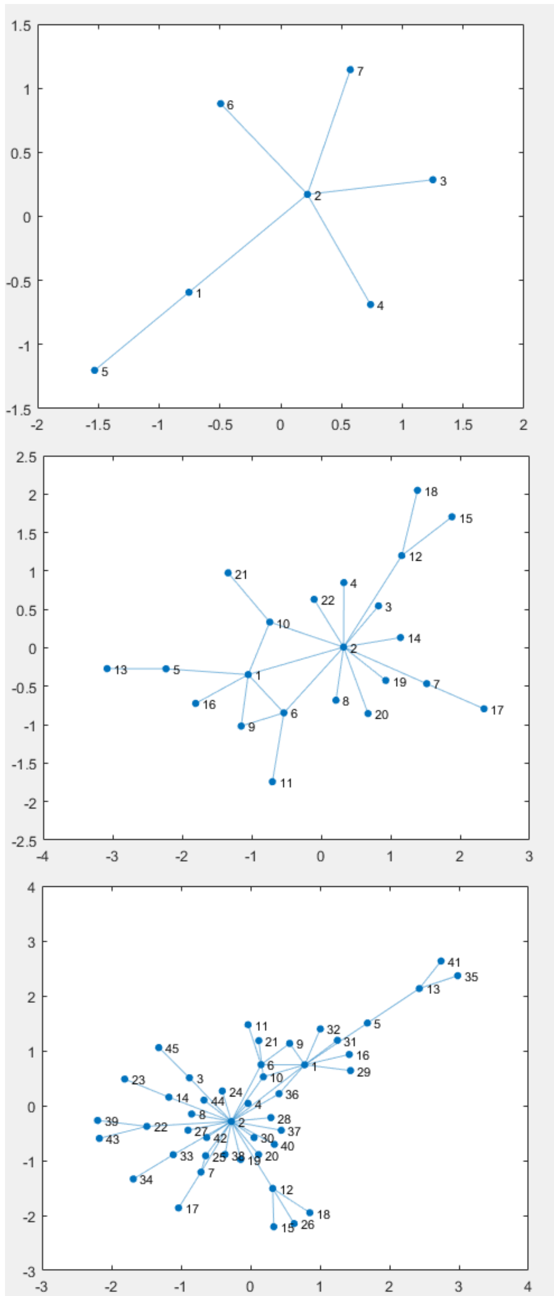

Figure 1: Growing Social Network Plots

5 Conclusion

A new algorithm has been implemented and tested on a user generated network using a blend of Ant Colony Optimization and Particle Swarm Optimization for finding a maximum clique in social networks using preferential attachment with nodes. In the algorithm,artificial ants (agents) are used to randomly walk on the logical graph based on pheromone deposition. Each ant chooses its path as per the value of the pheromone and heuristic information on the edges. The initial value of pheromone on each edge on the logical graph is parametrized and set according to the link connection on the network. The initial value of heuristic information on each edge is set according to the common neighbors of the two nodes or more nodes it connects to form a complex network.

Bibliography

[1] Mhand Hifi Didier El Baz and Xiaochuan Shi Lei Wu. *A Parallel Ant Colony Optimization for the Maximum-Weight Clique Problem*. Tech. rep. 2016 IEEE International Parallel and Distributed Processing Symposium Workshops.

[2] Berlingerio M. Bonchi F. Bringmann B. and Gionis. A. *Learning and Predicting the Evolution of Social Networks*. Tech. rep. IEEE Intelligent Systems, 2010.

[3] Ling Chen BolunChen. *A Link Prediction Algorithm Based on Ant Colony Optimization*. Tech. rep. 3. Oct. 2014, pp. 694–708.

[4] I. Lymperopoulos and G. Ioannou. *Understanding and modeling the complex dynamics of the online social networks: a scalable conceptual approach*. Tech. rep. Evolving Systems, 7:207-232., 2016.

[5] Masoud Asadpour Ehsan Sherkat Maseud Rahgozar. *Structural link prediction based on ant colony approach insocial networks*. Tech. rep. 2014.

[6] Guo J. Liu X. Hu H. and X. Wang. *A generalized theory of preferential linking*. Tech. rep. 2014.

[7] Mohammad Reza Meybodi Mohammad Soleimani-Pouri Alireza Rezvanian. *Finding a Maximum Clique using Ant Colony Optimization and Particle Swarm Optimization in Social Networks*. Tech. rep. 2012.

[8] Jun Yin Jingjin Guo Suqi Zhang Yongfeng Dong. *Improved Ant Colony Algorithm for Finding the Maximum Clique in Social Network*. Tech. rep. 2015.