Swarm intelligence for counting the degrees of separation in Social Networks

Àlex Pardo Fernandez ALEXPARDO.5@GMAIL.COM

David Sánchez Pinsach SDIVIDIS@GMAIL.COM

Abstract

This is a great project and therefore it has a concise abstract.

Keywords: Swarm Intelligence, Ant Colony Optimization, Twitter, Degrees of Separation.

1. Problem statement and goals

One of the most important features of humans and in general, of lots of animals, is the sociability. The ability of communicating each other in order to share knowledge. Human social relationships form a network where everyone is connected with those that communicates often (considered as fiends).

The theory of the six degrees of separation was originally set out by Frigyes Karinthy (Karinthy, 1929) and explains that everyone is connected to any other person in the world by six degrees. That means, if you want to met someone in the world, you will need to pass by other five persons, as maximum, between you and your objective so the last one will be the one you are trying to reach. During these last decades, the six degree theory has been used in many fields like economy, social networks and markets and is a known property of small-world networks where most nodes are not neighbours of one another, but most nodes can be reached from every other by a small number of steps.

Twitter is one of the biggest social networks, that is over 200 million users and over 400 million tweets (the 140 character messages that are the main feature of this social network) every day [1]. The main idea is to create shorts messages of the 140 characters in order to express your ideas or opinions in a shorten way. Moreover, Twitter have introduced some concepts that are very popular now such as the *hash-tag* which is a way of tagging the messages i order to find all the related ones.

Another famous social network is Foursquare, a social network related to places in which you activate the application and notifies your friends that you have been in that place. The most active users in one place achieve some goals such as being the "major" of a certain place. Foursquare has over 45 million users and over 3 billion check-ins every day [2].

1. Information by March of 2013: https://blog.twitter.com/2013/celebrating-twitter7
2. Information by January of 2014: https://foursquare.com/about

The main objective of this work will be to estimate the degrees of separation in Foursquare and if it is computationally possible, in Twitter and see if it is possible to obtain six degrees of separation between two random people. In order to do this task, it will use some ideas of the Computational Intelligence like Swarm Intelligence. In particular we are going to use Ant Colony Optimization (ACO) (Colorni et al., 1991) for Shortest Path finding (SPACO) (Angus, 2005).

2. Previous work

The previous work is based on these ideas:

- ACO original (Colorni et al., 1991): Explain the algorithm of the Ant Colony Optimization.
- General theory (Watts and Strogatz, 1998): Explain the general theory about the six degree.
- On Twitter as (Cheng, 2010): Explain the six degree on the Twitter network.
- On other social nets such as

3. The CI methods

Swarm intelligence is a natural method based in the behaviour of the decentralized individuals who obtain solutions in some problem as result of their interactions. These agents normally are simple, with a few capabilities and follow simple rules. It includes ant colony optimization, flocking of the birds, bacterial grow or some on.

As it mention previously, it will use the Ant Colony Optimization algorithm in order to obtain the degree of separation of two random person. In this type of algorithm the individuals are the ants who have few capabilities. The ants cannot communicate with the others and only use the trace of pheromone as a probabilistic method in order to obtain the choice when they have different options. When a ant travels although the two points it leaves a few of pheromone. The other ants around these points feel the pheromone and try to follow this path. The paths with few pheromone will have less probability than the path with high pheromone. As shows the figure 1, the ants converges to the shortest path between two points and this is the main characteristic of this method. We will use this idea of the shortest path between two points, as a minimum degree between two person in the Twitter network environment. In the appendix A, it will show how the algorithm works in details.

4. Results and Discussion

In order to perform the tests we used two different datasets

5. Extensions, strengths and weaknesses

6. Conclusions

References

Daniel Angus. Solving a unique shortest path problem using ant colony optimisation. *Communicated by T. Baeck*, 2005.

Alex Cheng. Six degrees of separation, twitter style. 2010. URL http://www.sysomos.com/insidetwitter/sixdegrees/.

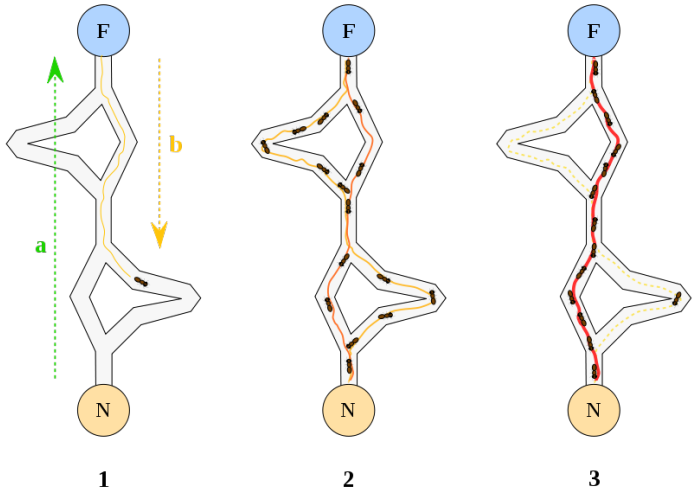

Figure 1: ACO algorithm

Alberto Colorni, Marco Dorigo, and Vittorio Maniezzo. Distributed Optimization by Ant Colonies. In *European Conference on Artificial Life*, pages 134–142, 1991.

Frigyes Karinthy. Chain-links. *Everything is the Other Way*, 1929.

Duncan J Watts and Steven H Strogatz. Collective dynamics of "small-world" networks. *nature*, 393(6684):440–442, 1998.

Appendix A. Implementation details

The application has been coded using Python 2.7. The system uses NetworkX library in order to represent the graph. Additionally, in order to create an small dataset based on twitter, we used Twython implementation of the Twitter API in order to acquire the data and again, NetworkX in order to save the resulting graph.

We based the code in the next basic pseudo-code algorithm of the ACO:
¡put here pseudocode!
Next to the pseudo-code we want to enter a few detail of some parts of the code and how this parts had been coded.

- **Pheromone**: The pheromone is the most important part of the our problem and our code. Remember that the ants try to follow the path which contain high quantity of pheromone. However the pheromone effect disappears with the time. For this reason, the algorithm need to define two parameters in order to determine the quantity of the pheromone that leaves the ant and the quantity of the pheromone disappears at each turn.

- **Start and final points of the ants**: We select always randomly the start and the final point of the ants using the standard Random methods of the python.