ESIGELEC – Project S8

The Overview of the thesis

An approximation algorithm for shortest path based on the hierarchy networks

Project efficient itinerary – Team 1 LIU Jixiong

28/03/2018

The tremendous scale of the social networks mined from Internet is the main obstacle of a social network analysis application. The bottleneck of many network analysis algorithms is the extortionate computational complexity of calculating the shortest path. Real-World networks usually exhibit the same topological features as complex networks such as the "scale-free" and etc, which indicate the intrinsic laws of the shortest paths in complex networks.

After studied the thesis "An approximation algorithm for shortest path based on the hierarchy networks" which allow us to build a new hierarchy network to make the calculate more rapid in the real

Hypothesis and limits

- o   Hypothesis:
    - -   There are important point who is already defined
    - -   The distance between two important points are correct
- o   Limits:
    - -   The points and the destination
    - -   The distance between points
    - -   The important points

Used approach

Based on the topological features of real-world networks, a novel shortest path approximate algorithm which uses an existent short path passing through some local center nodes to estimate the shortest path in complex networks, is proposed. This paper illustrates the advantage and feasibility of incorporating the proposed algorithm within the network properties, which suggests a new idea for complex social network analysis.

The basic idea of the algorithm is to iteratively construct higher level hierarchy networks by condensing the central nodes and their neighbors into super nodes until the scale of the top level network is very small. Then the algorithm approximates the distances of the shortest paths in the original network with the help of super nodes in the higher level hierarchy networks.

Results

The proposed algorithm is focused on undirected and unweighted networks. Its central node based on the hierarchy networks. Moreover, they will consider an adaptive algorithm for different sorts of networks.

The algorithm approximates the distances of the shortest paths in the original network by means of super nodes in the higher level hierarchy networks. They test the performance of the proposed algorithm on four real networks. The results show that our algorithm can maintain high accuracy while achieve the runtime per query within a few milliseconds on large scale networks.

The proposed algorithm has been evaluated both on synthetic network stage and real world network stage. Experimental results show that the proposed algorithm can largely reduce the computational complexity and remain highly effective in complex networks.

Advantages, disadvantages and errors

- o   Advantages:

- Developer-friendly, easy to understand:

The algorithm used in the article does not require a complex implementation. T tried to simplify the complex network to a new network with just a few important points. So it makes the developers easier understand what is going on in this algorithm.

- High speed of calculate :

With less points and less paths to calculate, of cause the speed of the procedure will be raised up. This is a great advantage if the data scale is way too big.

- Implementable:

We can join same others arguments like price and time in this model, and to have a new network which have other dimensions or a mixed dimension. And with this method, we can have an system dynamic.

o Disadvantages:
- Non-dynamic:

The method proposed in the article is optimal only if the input values of arguments are fixed and unique (distance).

But we can change a little bit and join some arguments to make it dynamic

- Low precision:

In this algorithm, they just neglect the base network called i, and they pay more attention on they new network called i + 1, which means some the shortest path in the network i may be ignored in the network i + 1. Some times the error is acceptable(because it is an approximation algorithm), however, if two points is too close, the error may became large. So they have already considerate the situation if two points share the same important point, but didn't considerate what would be if their important points aren't the same, but very close, one next to another.

- Difficult to choose important point:

In their thesis, important points are already defined, but in real word, it wouldn't be simple like this, it's still hard to find an data base which has already gave you all the important points, so when we want to us their algorithm we should write another function to choose the important points .

Conclusion

This algorithm try to build a new approximation network, scarify the precision and achieve the efficiency. But it still have some problems, therefor, we can't directly just copy what they have done. However we are inspired by this thesis, we can use their model, with some personal modification to make it suitable to our project.
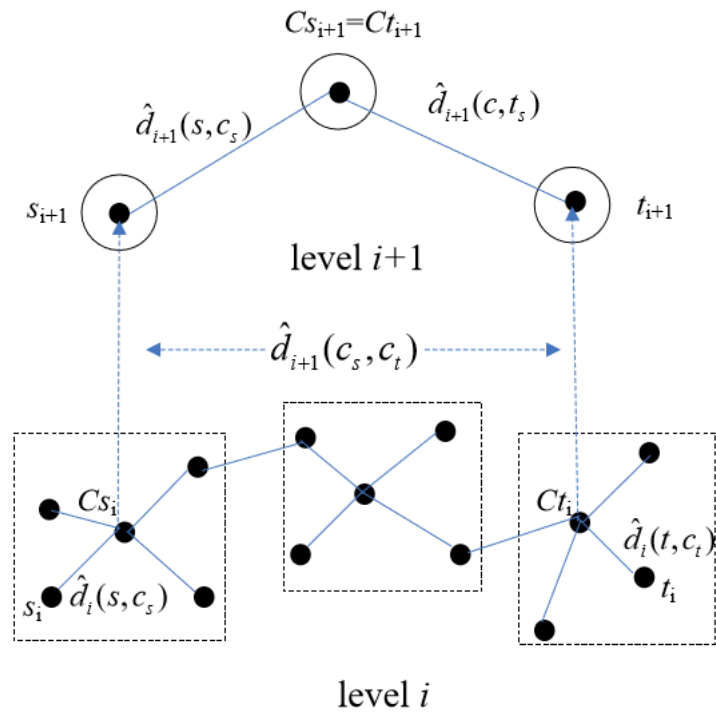
Annex

Figure1 : Illustration of the iterative approximation