BDA 6-5

1

Hello everyone, I am Haiying Che, from Institute of Data Science and knowledge Engineering

School of Computer Science, in Beijing Institute of Technology

2

In this chapter, we introduced some useful platform，Spark Mllib, TensorFlow and a typical big data analysis application, **Recommendation System.**

in this session we will talk about Social Network analysis

3

In this session we will learn social network, social network analysis, the **Metrics of social networking analysis, and Force-directed graph drawing**

At last we will do Social network analysis hands

4

A **social network**is a social structure made up of

* a set of social actors (such as individuals or organizations),
* sets of dyadic ties,
* and other social interactions between actors.

The study of these structures uses social network analysis to identify local and global **patterns**, locate **influential entities**, and examine **network dynamics**.

It characterizes networked structures in terms of nodes (individual actors, people, or things within the network) and the ties, edges, or links (relationships or interactions) that connect them.

Let us watch a video about social network overview

5

**Social network can be presented using a graph.**

A node refers to a social actor, and an edge refers to a relationship.

The weight of an edge measures how close the relationship is.

**Social Network has Low average distance.**

If we draw a social graph of everybody in the world, there will be more than 7 Billion nodes.

But according to Six Degrees of Separation, the minimum distance of two nodes is very probably less than or equal to 6.

Thanks to the development of Internet, the average distance of two nodes is going lower.

**Social network analysis (SNA)** is the process of investigating social structures through the use of **networks and graph theory**.

Social networks and the analysis of them is an inherently interdisciplinary academic field which emerged from **social psychology, sociology, statistics, and graph theory**.

Let us watch a video about What is Social Network Analysis

6

Some Metrics of social networking analysis,

The metrics about Nodes and edges include Average Clustering Coefficient; Eigenvector Centrality; Average Path Length etc.

The metrics about Network include Average Degree, Average Weighted Degree, Network Diameter, Graph Density, Shortest Path, Hits, Modularity, PageRank, Connected Components and etc.

7

Visualization is an important way to analyze the social Network.

Force-directed graph drawing is a good way to visualize the social Network.

Force-directed graph drawing is an important technology on social network visualization.

Force-directed graph drawing algorithms are a class of algorithms for drawing graphs in an aesthetically-pleasing way.

Currently force-directed graph drawing algorithms are Force atlas algorithm, Fruchterman-Reingold algorithm, Yifan Hu algorithm, etc.

The basic theory is assuming that there are forces between nodes and fits in physical models.

And the goal is to make the whole system stable or has the lowest energy.

8

Fruchterman-Reingold algorithm is one of the most commonly used algorithms.

This algorithm uses electromagnetism theories, and generates a circular graph.

Repulsive force divided by k ; Gravitation force

Final state is to make force balanced

The picture on the left is the Original graph and the picture on the right is Fruchterman-Reingold style, which is more aesthetically-pleasing.

9

We will hands on based on the data set of social network of “Game of throne” , using NetworkX and Gephi,

NetworkX is a Python package for the creation, manipulation, and study of the structure, dynamics, and functions of complex networks.

Gephi is an open-source and free visualization and exploration software for all kinds of graphs and networks.

Experiment 1 includes Loading data and generate graphs, Network analysis, Visualization using Networkx and Edge bundling using datashader.

Experiment 2 includes Load data and analysis, Graph simplifying using Fruchterman-Reingold algorithm and Visualization and beautification

Experiment 3 includes social network recommendation background setting, biased random walk, node representation learning and recommendation.

all the experiments material including the manual and codes are provided on the platform, which can help you to do the hands-on.

10

In this session, we learned social network analysis,

thank you for your attention, if you have any question, feel free to contact me.