Comparing Unipartite and Bipartite Graphs

The networks that we have discussed thus far are all unipartite networks represented by unipartite graphs. This means that only one type of node is considered. Let's take corporate tax evasion fraud as an example. The nodes of the network are companies. Each company can be either fraudulent or legitimate. Companies can be linked by their resources, such as employees, equipment, physical address, and so on. Links or edges can be weighted using the Jaccard index, which is equal to the amount of shared resources divided by the total amount of resources. The advantage of this representation is that it is simple. A disadvantage is that we are missing resource-specific information.

In a bipartite graph, two types of nodes are included. In our fraud example, we now have both companies and resources in the network. Companies are represented as circles and resources as diamonds. Both companies and resources can be fraudulent or legitimate. Links or edges now imply an association between a resource and a company. They can be weighted using the recency of the association. A bipartite graph gives a more detailed representation because it also includes fraudulent resources. A disadvantage is that the complexity is increased.

Here you can see the difference between a unipartite and bipartite graph further clarified. In this unipartite graph, the white nodes represent legitimate companies and the black nodes represent fraudulent companies. We want to infer the fraud behavior of the companies that are labeled 1 and 2. Both have two fraud links and two non-fraud links, so according to the unipartite graph, they have the same risk. Here you can see the corresponding bipartite graph with the resources added. Company 1 has one connection to a resource that is non-fraudulent. It has two connections to resources that have each been engaged in two companies, of which one is fraudulent. Hence, company 1 has two resources that have a 50% fraud probability. It has no connections to resources that have a 100% fraud probability. Let's now look at company 2. This company has one resource that has a 0% fraud probability, one resource that has a 50% fraud probability, and one resource that has a 100% fraud probability. The latter is the resource indicated with the red circle. By adding resource-specific information and turning the unipartite graph into a bipartite graph, it becomes clear that company 2 is riskier than company 1. This clearly illustrates the added value of working with bipartite graphs.

This is another example of a unipartite graph. The red nodes represent the fraudulent companies and the green nodes are the legitimate companies. From the graph, it's difficult to see which companies are most risky. Here you can see the corresponding bipartite graph. By adding the resource information, it is now clear that companies 2 and 20 will most likely be fraudulent, as well.

Social Network Analytics

Close