NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE
School of Computer Science and Engineering
College of Engineering

Quantifying reputation and success of data scientists

Student: Isaac Soh Wei Yang Supervisor: Dr Sourav Saha Bhowmick

Background

In today's hyper-competitive world, people are obsessed with success. For decades, researchers and scientist have studied the issue of success and quantifying them. However, not all success can be easily quantified, such as the success of researchers or artists. Unlike sports, where there are many objective criteria to measure success.

Objectives

This project exploits network science techniques to understand and predict success of data scientist. We utilize Database Systems and Logic Programming (DBLP) datasets to track the career of data scientists in data management conferences, where they are known to be of high caliber and have low acceptance rate, over time and perform network analytics to understand and quantify their career success.

Key findings

1. Initial Success plays a huge role in defining a scientist's success
2. If scientists do not breakthrough within 8 years, they are unlikely to
3. Scientists who do not collaborate are stuck with limited success

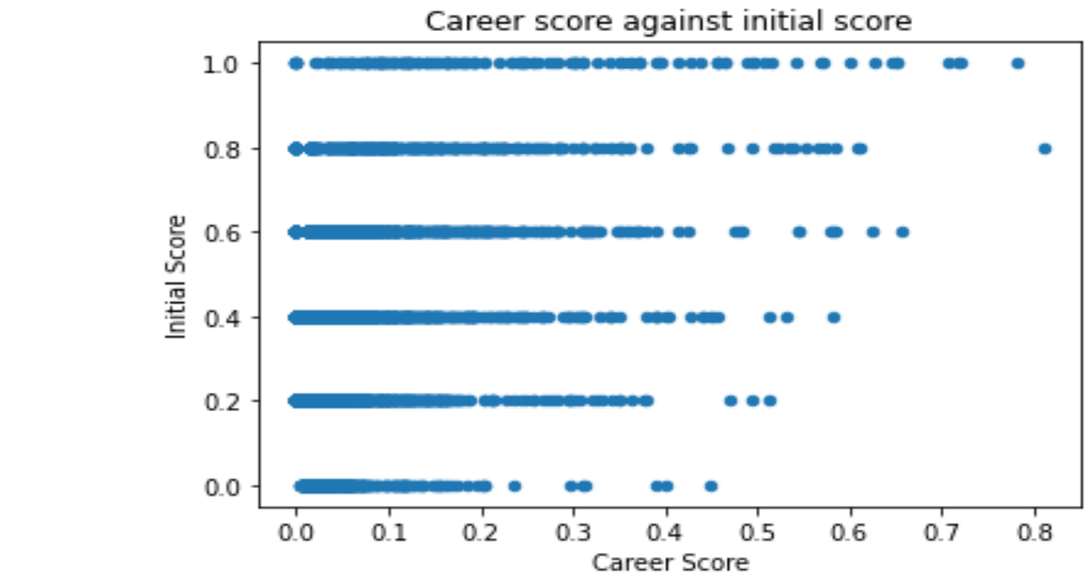

Figure above shows a strong positive correlation of 0.494 between initial score and career score

Figure above shows over 80% of scientist managed to breakthrough to a Tier 1 venue within 8 years

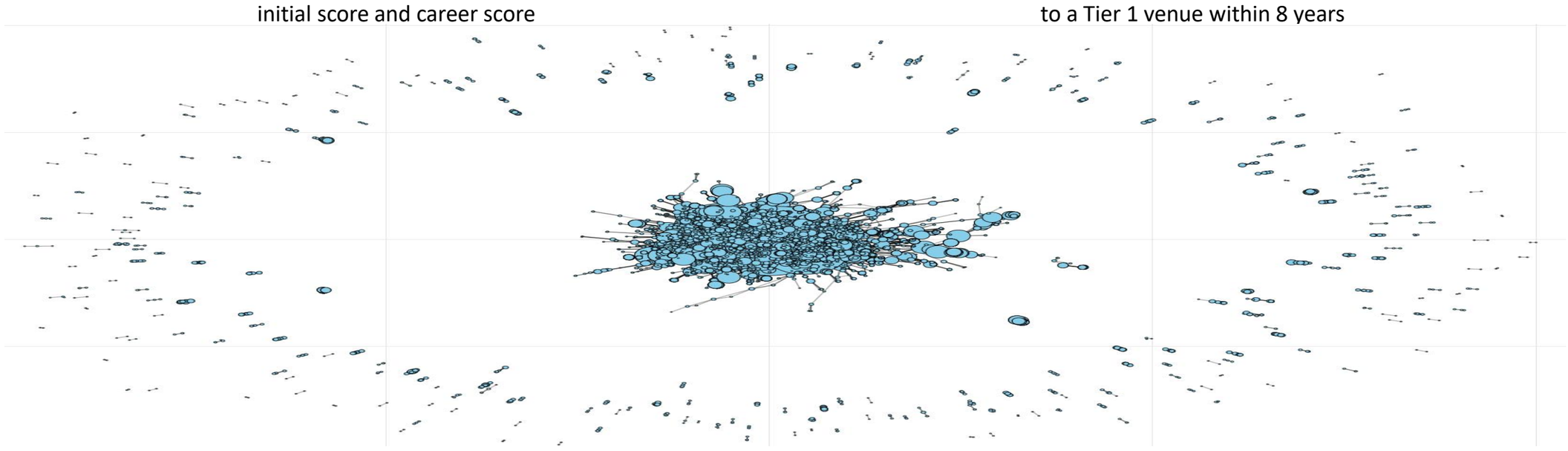

Figure above shows the co-authorship network among the scientist. As seen above, nodes that are not in the centre community are much more smaller, indicating limited success.