

Finding the perfect job: visualisation and exploration of the relationships between strengths and traits in personality survey data using networks

Supervisor: Colm Connaughton (Complexity and Mathematics)

Co-supervisor(s): James Clyne

External partner: Capp Ltd.

Scientific background

Founded in 2005, Capp have an established academic background in positive psychology and an extensive experience in the provision of strengths-based talent management solutions. The Capp vision is “to match the world to their perfect job”. Individuals’ strengths and traits are assessed via online questionnaires. The responses to these questionnaires are then used to score each individual on various skills and personality attributes which are then matched against known requirements for different professional roles and vacancies. Good matches are highly beneficial both to potential employers and to job seekers. The company is constantly improving its matching algorithms and seeks to incorporate best practice in modern data analytics into its operations.

Research challenge

Capp have a large database containing a large number of responses to questionnaires provided by job seekers. Since the questionnaires are done online, they can adapt the questions asked to try to explore an applicant’s interests. The responses have been used to score each applicant against a number of different strengths using in-house scoring algorithms. The data set therefore consists of a number of classes of discrete entities - applicants, questions, responses, strengths etc - and the relationships between them. This project has two strands:

- **Data representation and visualisation:** Capp are particularly interested in exploring whether it is useful to look at their data from the view point of complex networks. Clearly there are many different ways to represent the data as a network. For example, the nodes could be applicants and a link indicates that they applicants both scored the same value for a par-

ticular strength or the nodes could be questions and a link indicates that both questions are relevant to a particular strength. The first part of the project is to explore what types of network are representations are informative. Is a multi-layer network the appropriate model? Should the edges be weighted and how?

- **Dynamic assessment of information content of questions:** Capp have a strong sense that only a fairly small number of questions are required to establish the strengths and traits of an applicant. They are interested to know what is the best order in which to ask questions so that the strengths of an individual can be assessed accurately in as few questions as possible. The second part of the project is to use the data to construct a probabilistic model of the responses and associated strength scores. If correctly calibrated, such a model could be used to ask the questions in the most informative order using statistical techniques such as mutual information or decision trees.

Pre-requisites

This is a data-centric project so a prospective student should be proficient at handling and processing real data. An interest in networks and statistical analysis is also required. The project is quite open-ended so a willingness to experiment and think independently is important.

Additional considerations

If the project goes well there will be a follow-on PhD project available and the possibility of an internship with Capp as part of what is hoped to be an on-going collaboration between Capp and the Centre for Complexity Science.