Evolution of Patent Networks

Alun Meredith and Dr Markus Brede   MSc Data Science

Introduction

Bibliographic (citation) networks.

The patent system can be described as a network in which each patent is directionally linked to its "prior art". There are many similarities between this network and academic citation networks. Many relate innovation networks to evolutionary systems where the state of the art is mutation and combination of the prior art [2].

Preferential attachment

The preferential attachment model [3] which has been used to describe patent networks gives the likelihood of a patent receiving a new incoming citation as a function of the number of citations it already has. This 'rich get richer' mechanism describes how already successful patents continue to accumulate new citations. In this model the accumulation of citations in the early growth period is heavily stochastic as there will be many patents of similar age and low numbers of citations, before converging to a more predictable state.
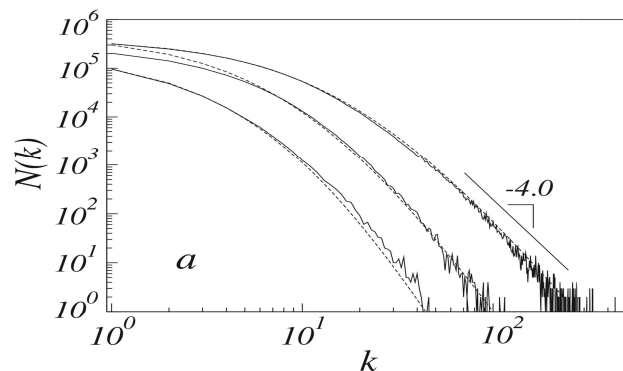

Fig. 1. In-degree distribution of patent network for 3 different years 1984,1992, 2002 (from left to right). Shows extended power law: $P_i(k) \sim (k + k_0)^{-\gamma}$ (P: probability of receiving new link, k: number of existing links). *Source: Valverde et. al [1]*

Aims

- Investigate the early growth of a patent
- Assess the predictability of success of a patent.
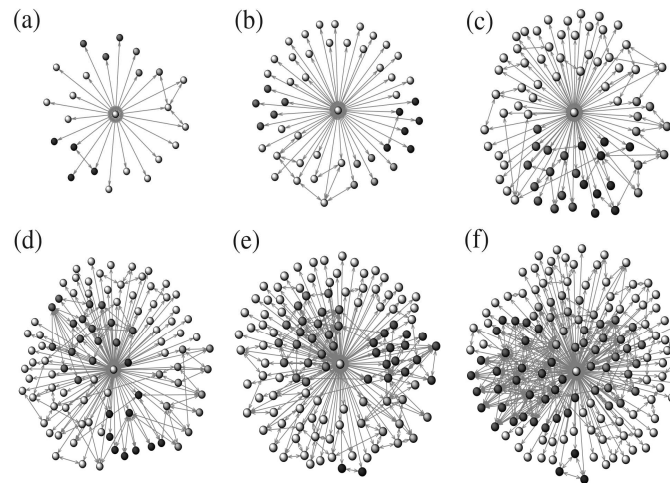- What is the relationship between predictability and the age of the patent?

Fig. 2. Evolution of computed tomography (CT) network, from a precursor hub node. Shades of grey represent final modules of the network. *Source: Valverde et. al [1]*

Why it Matters

- Better modelling of innovation networks can improve understanding of the systems which produce new technologies.
- Identifying key technologies earlier can lead to more agility on the part of innovators.
- Can lead to breakthroughs in other related areas such as modelling innovation networks as an evolutionary process.

Objectives

- Build a bibliographic network using the US Patent Office Dataset (USPTO)
- Reproduce the work of Valverde et. al. estimating the parameters of the preferential attachment model for this network (Fig. 1 & 2).
- Produce a predictive model of the future citation activity of a patent based on the time series of its early growth and implicit parameters such as industry code.
- Extract additional features of patents through text mining techniques such as computing measures of semantic distance, building a semantic network and integrating these features into the model.
- Evaluate this model, contrasting with preferential attachment. How do the initial features of the patent affect predictability?

Plan
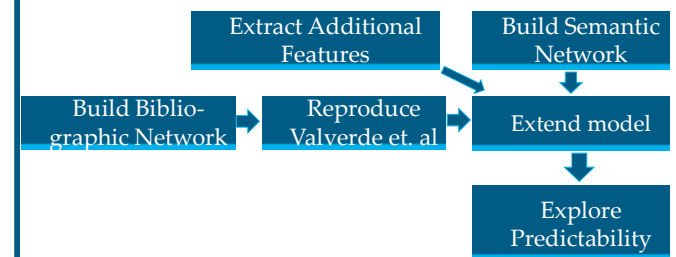
References

[1] S. Valverde, R. V. Solé, M. A. Bedau, and N. Packard, "Topology and evolution of technology innovation networks," Physical Review E, volume 76, no. 5, p. 056-118, 2007.

[2] W. B. Arthur and W. Polak, "The evolution of technology within a simple computer model," Complexity and the Economy, 2014

[3] D. d. S. Price, "A general theory of bibliometric and other cumulative advantage processes," Journal of the American society for Information science, vol. 27, no. 5, pp. 292-396, 1976