Collective Credit Allocation in Science

Hua-Wei Shen and Albert-László Barabási
Presenter: Chunheng Jiang

Problem: Credit Allocation

Interdisciplinary collaboration has been a trend, e.g. the Human Genome Project. Papers are the primary form of intellectual products. However, there is no general guideline to place authors. Different fields may have extremely different traditions (corresponding author, contribution or alphabetical ordering), and the length of the list also varies widely from a solo to a giant orchestra with over 5,000 authors [?].

To make accurate estimation of their contribution to a joint work, a.k.a the Credit Allocation Problem, a simple but effective algorithm was proposed [?].

Algorithm

- Target Paper: p_0
- Authors: $\mathcal{A} = \{a_1, a_2, \ldots, a_m\}$
- Papers Citing: $\mathcal{D} = \{d_1, d_2, \ldots, d_l\}$
- Papers Co-cited: $\mathcal{P} = \{p_0, p_1, p_2, \ldots, p_n\}$

The credit shares of a coauthor is determined by whether she continues working with other coauthors on the topic of p_0. Here, the set $\mathcal{D}$ plays as a committee from the same community, and the citations are their votes. A bipartite network can be constructed.
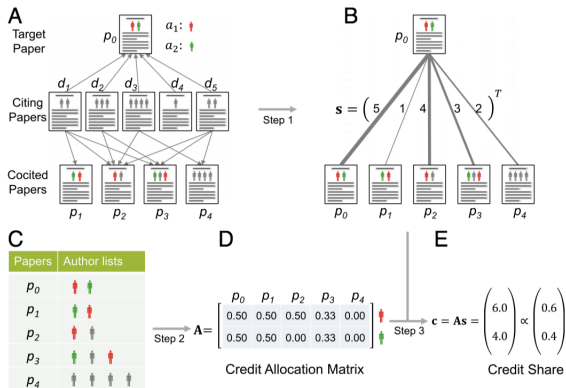
Algorithm

The credit share of a_i in p_0 is computed with the formula:

$$c_i = \sum_{j=1}^{n} A_{ij} s_j,$$

where A_{ij} is the credit a_i earns from the cocited paper p_j, and each coauthor gets the same amount of shares from p_j. The relevance level (or cocitation strength) s_j of p_j to p_0 is defined as the number of documents in $\mathcal{D}$ referring both p_0 and p_j.

Algorithm

Credit Allocation Algorithm

Evaluation

To evaluate the performance, the Nobel-prize winning papers in Physics (1995 - 2013), Chemistry (1998 - 2013), Medicine (2006 - 2013) and Economics (1998 - 2013) are collected to build the citation network. The credit allocation algorithm produces a credit share for each coauthor appeared in the awarding Nobel Prize papers.

The experimental results indicated that the algorithm performs very well, and successfully assigns the Nobel-prize winners the highest credit share in the awarding papers.

Evaluation

Disciplines	Priors				
	Count	Fractional	Harmonic	Axiomatic	Zhang
Physics	76%	76%	52%	44%	68%
Chemistry	71%	83%	67%	67%	88%
Medicine	86%	86%	43%	36%	100%
Overall	76%	81%	56%	51%	82%

Accuracy of the proposed algorithm with different prior rules for A

Evaluation
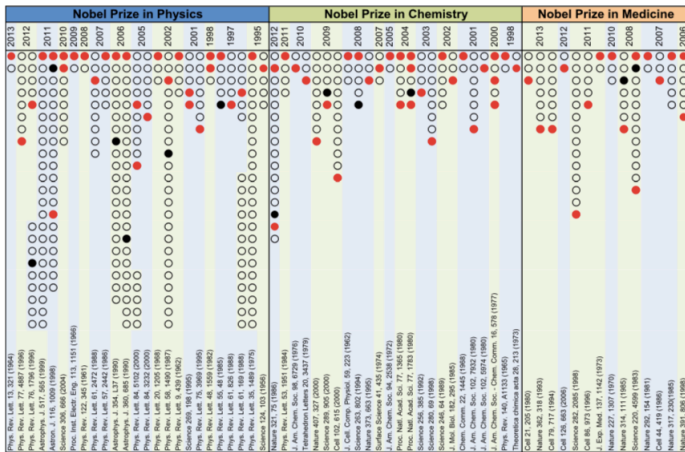
Highest credit share earners (●) and the Nobel Prize winners (●)

Further Studies

- Integrating both the order of coauthors and the contextual info to design better credit allocation matrix $A = (A_{ij})_{i=m}^{n+1}$
- Differentiate the values of the citing papers $\mathcal{D}$ because of the reputation differences, produce more reliable relevance estimation of s (link-analysis based analyze over all papers)
- Authors' institutions and contextual text of the citation(definition, ideas)