

Another article that makes bibliometric analysis a bit harder

J. Pfeffer

Carnegie Mellon University

J. Pfeffer

Stanford University

April 1, 2015

There are many authors publishing many articles in a wide variety of scientific fields. Some of these researchers analyze the collaboration among researchers by looking at co-publishing or citation behavior. "A key challenge when working with publication data is to disambiguate different authors carrying the same name, as accidentally merging multiple authors can distort results massively. This challenge is even harder when an initial is used in place of a first name. The goal of this article is to exacerbate this challenge and to make bibliometric analysis harder still.

1 Introduction

Researchers publish a lot, on many topics. They publish about people (J. Pfeffer, 1998; al. et J. Pfeffer, 1978; J. Pfeffer, 1994), organizations (J. Pfeffer et al., 2003; J. Pfeffer, 1992; J. Pfeffer et al., 1981), networks (J. Pfeffer et al., 2012; al. et J. Pfeffer, 2012; J. Pfeffer et al., 2011), time lords (al. et J. Pfeffer, 2013), as well as far more complicated topics (al. et J. Pfeffer, 1986; al. et J. Pfeffer et al., 1997). More recently, researchers are more and more interested in social media stuff (al. et J. Pfeffer, 2014; J. Pfeffer et al., 2014; al. et J. Pfeffer, 2014; al. et J. Pfeffer et al., 2013; J. Pfeffer et al., 2012).

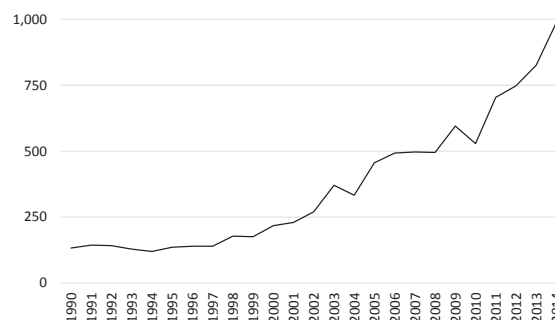


Figure 1: Research products by year for “J Pfeffer” on Google Scholar. There are an additional 1,660 results from year 0 to year 1989, resulting cumulatively in about 10,850 research products.

Intriguingly, publishing a lot is neither a new nor a completed phenomenon. As a matter of fact, *J. Pfeffer* (as well as others) has been creating more and more research products every year (see Figure 1). However, recursive citations can be seen as contemporary activity (J. Pfeffer & J. Pfeffer, 2015).

The contributions of our work are:

- We show that co-publishing analysis can be quite hard;
- We make co-publishing analysis even harder;
- By adding some random *J. Pfeffer* references to our paper, we create very interesting artifacts for future co-citation analysis.

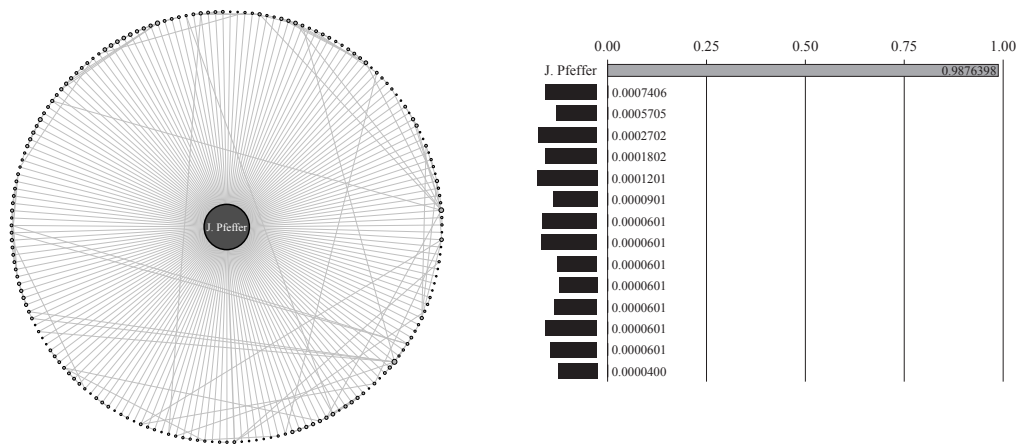
2 Some Analysis

In order to show a network visualization and a ranking, we extracted the first 1,000 results from a Google Scholar search with “J. Pfeffer” and constructed a network (al. et J. Pfeffer, 2012). Nodes in this network represent people, i.e. J. Pfeffer and others. Links connect these nodes in case of at least one shared publication. The result can be seen in Figure 2(a). Authors that share multiple publications are connected by a thicker line.

To quantify the importance of J. Pfeffer in this network, we calculated betweenness centrality and present the very obvious results in Figure 2(b).

3 Outlook

To assess the importance of this issue for bibliometric analyses of the years to come, we estimate future levels of activity based on the previous scientific productivity of *J. Pfeffer*. Using data from



(a) Impressive Collaboration Network

(b) Intimidating Ranking

Figure 2: The pretty impressive collaboration network of *J. Pfeffer*, based on top 1,000 results of a Google Scholar search as well as betweenness centrality ranking in this network. Out of respect to the co-authors of *J. Pfeffer*, we keep quiet about their identities.

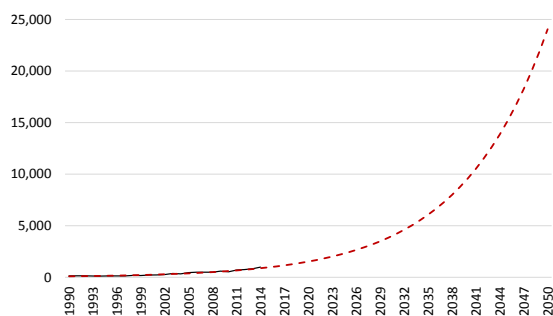


Figure 3: Estimated future productivity of “J Pfeffer” by year based on exponential growth fitted on historic data.

Figure 1, we first fit an exponential function to describe historic activity. It is important to notice that the following exponential function almost perfectly ($R^2 = 0.9471$) describes our historic data:

$$y = 88.43e^{0.0919*(year-1989)}$$

Figure 3 shows the full extend of the *J. Pfeffer* problem. Assuming a continued exponential growth, we expect about 270,000 cumulative research products by the year 2050.

4 Conclusions

It is okay to have a common family name and to give your child a boring first name. However, please consider one or more middle names to give her or him a head start in terms of an unmistakable identity. Exemplary behavior was shown by the parents of J.M. Pfeffer (J.M. Pfeffer et al., 2012; J.M. Pfeffer et al., 2013) and J.T. Pfeffer (J.T. Pfeffer, 1974; J.T. Pfeffer, 1992). In countries in which ~ 100 million people share the same family name, we recommend five middle names with at least two being picked completely at random.

5 Acknowledgments

The authors will always be grateful to Google Scholar for indexing this article. Without their greedy literature gathering approach, our work would be preposterous.

References

- Castillo, C., El-Haddad, M., Pfeffer, J., and Stempeck, M. (2014). Characterizing the life cycle of online news stories using social media reactions. *Proceedings of the 17th ACM conference on Computer supported cooperative work and social computing*, 211-223.
- Columbus, D. and Pfeffer, J. (2013). 50 years of team tardis. *Whotopia 25*, 34-37.
- Hennig, M., Brandes, U., Pfeffer, J., and Mergel, I. (2012). Studying social networks: A guide to empirical research. *Campus Verlag/University of Chicago Press*.

- Morstatter, F., Pfeffer, J., Liu, H., and Carley, K. M. (2013). Is the sample good enough? comparing data from twitter streaming api with twitter firehose. *Proceedings of ICWSM, 2013*.
- Pfeffer, J. (1974). Temperature effects on anaerobic fermentation of domestic refuse. *Biotechnology and Bioengineering*, 16(6):771–787.
- Pfeffer, J. (1992a). Managing with power: Politics and influence in organizations. *Harvard Business Press*.
- Pfeffer, J. (1994). Competitive advantage through people. *California management review* 36 (2), 9-28.
- Pfeffer, J. (1998). The human equation: Building profits by putting people first. *Harvard Business Press*.
- Pfeffer, J. and Carley, K. M. (2011). Modeling and calibrating real world interpersonal networks. *Network Science Workshop (NSW), 2011 IEEE*, 9-16.
- Pfeffer, J. and Carley, K. M. (2012a). k-centralities: Local approximations of global measures based on shortest paths. *Proceedings of the 21st international conference companion on World Wide Web ...*
- Pfeffer, J. and Carley, K. M. (2012b). Social networks, social media, social change. *Proceedings of the 2nd International Conference on Cross-Cultural Decision Making*, 273-282.
- Pfeffer, J. and Lammerding, C. (1981). Power in organizations. *Pitman*.
- Pfeffer, J. and Pfeffer, J. (2015). Another article that makes bibliometric analysis a bit harder. *Proceedings of the 9th SIGBOVIK conference*.
- Pfeffer, J. and Salancik, G. R. (2003). The external control of organizations: A resource dependence perspective. *Stanford University Press*.
- Pfeffer, J., Zorbach, T., and Carley, K. M. (2014). Understanding online firestorms: Negative word-of-mouth dynamics in social media networks. *Journal of Marketing Communications* 20 (1-2), 117-128.
- Pfeffer, J. M., Moynihan, P., Clarke, C. A., and Clarke, A. J. (2012). Control of lytic transglycosylase activity within bacterial cell walls. *Caister Academic Press, Norfolk, UK*, pp. 55-68.
- Pfeffer, J. M., Weadge, J., and Clarke, A. (2013). Mechanism of action of neisseria gonorrhoeae o-acetylpeptidoglycan esterase, an sgnh serine esterase. *The Journal of Biological Chemistry*, 288, 2605-2613.
- Pfeffer, J. T. (1992b). *Solid waste management engineering*. Prentice Hall.
- Ruths, D. and Pfeffer, J. (2014). Social media for large studies of behavior. *Science* 346 (6213), 1063-1064.
- Salancik, G. R. and Pfeffer, J. (1978). A social information processing approach to job attitudes and task design. *Administrative science quarterly*, 224-253.