4th International Conference on Computational Social Science IC^{2}S^2, July 12–15, 2018
Northwestern University's Kellogg School of Management, Evanston, IL. USA

Evolution of employment in the United States: a half-century of polarization

Judy Hanwen Shen, Morgan R. Frank, Esteban Moro, Iyad Rahwan

Keywords: Networks Science, Employment, Data Science, Information Retrieval, Automation

Extended Abstract

As new technology advances, particularly in artificial intelligence and robotics, the question of how these technologies will reshape the workplace becomes more urgent as more jobs become susceptible to automation. The Skill-Biased Technical Change (SBTC) hypothesis [2], attributes a rise in income inequality to new technology entering the job market which raises the demands for high skilled workers over low skilled workers. This hypothesis has been widely adopted to explain job polarization: the increase in employment in low and high skilled occupations and the decrease of middle skilled occupations.

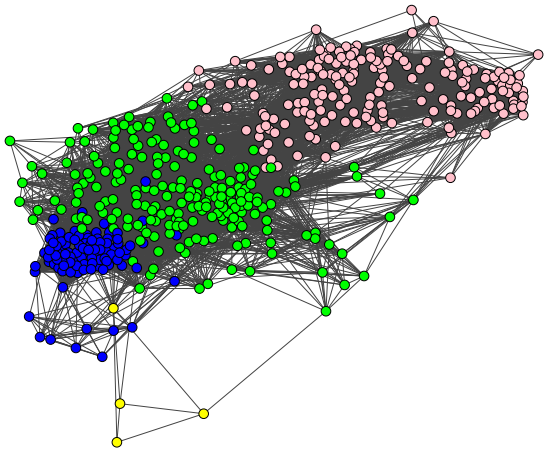
To understand job markets, analysis beyond the coarse-grained division of jobs into income-based high, middle, and, low skilled categories. For example, MacCrory et al. (2014) identifies technological advances as a driver for changed skill composition in different occupations. Using skill decomposition provides a better description of labor markets that helps decompose the forces that previously impacted job polarization and forces that could shape occupational changes in this coming era of automation. Unfortunately, most of the available data comes from recent years, which prevents us from understanding long term trends.

To address those questions, we extend the study of job polarization across the 20th century from 1939 to 1991 in the United States using the Dictionary of Occupational Titles. This occupational manual was developed by the United States Employment Service in efforts to standardized occupational information and job placement. Each title in the dictionary is accompanied by a description written by occupational analysts observing workers on the job. This dictionary was released in 4 editions: 1939, 1949, 1965, and, 1991. There are also a number of supplements intermittently produced between these 4 editions. The number of jobs in each edition varies from $11,480$ in 1991 to $22,028$ in 1949. Each edition was published in order to capture the changes in occupational content due to technological advancement [4] and helps explain job polarization through the underlying changes in skill requirements for different occupations across the years.

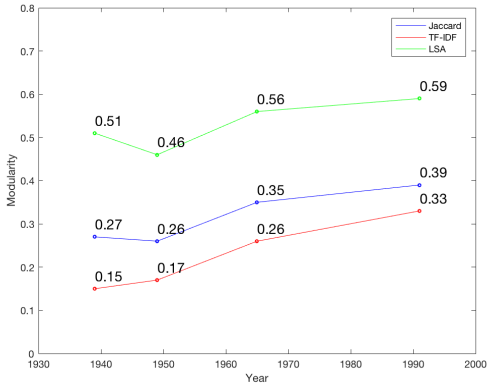
Since all three editions before 1991 are only available in scanned format, the files we used were digitized by the Hathi Trust Digital Library [1]. To compare job similarities, we use three interpretable document similarity measures: Jaccard Distance, TF-IDF and Latent Semantic Analysis. We calculate pairwise similarity measures between all jobs in each edition of Dictionary and use a z-value of 2 to only keep statistically significant document similarities. We then construct a graph on the remaining edges ($10\% - 30\%$) to represent the network structure of the jobs. To build community structure, use fast greedy optimization of modularity [1].

We find that this job similarity network shows a highly modular structure across all years. For example, figure 1a shows a job similarity graph of the 500 titles in the 1991 version of the Dictionary of Titles. The pink cluster represents managerial and communication jobs with most frequent words such as *record*, *direct*, and *determine* in the occupational descriptions. The blue cluster represents heavy industry jobs with frequent words such as *tool*, *pipe*, and *machine*. The green cluster represents various manual labour jobs with frequent words such as *press*, *cut*, and *place*. Furthermore, there is a significant rise in modularity of the job similarity network

[1] www.hathitrust.org



(a) Job similarity network: managerial and communication jobs (pink), heavy industry jobs (blue), general manual labor jobs (green), and specialized public service jobs (yellow)



(b) Change in modularity from 1939 - 1991

Figure 1: Modularity of jobs in the Dictionary of Titles

across all three different document similarity measures. Figure 1b illustrates the increase in job description modularity from 1939 to 1991. This growth in modularity suggests a polarization of jobs overtime by skill requirements; more and more occupations share skills between the cluster than across different clusters. As job clusters become more polarized, the barrier of entry in entering an occupation in a different cluster also rises.

Our work illustrates the modularity in United States occupations between 1939 and 1991 through document similarity analysis of job descriptions. Across three different document similarity measures, we find that modularity in the job similarity network increases. Future work will focus on decomposing the substructure of each cluster of jobs and on understanding how the same jobs change or disappear with time.

References

[1] Aaron Clauset, Mark EJ Newman, and Cristopher Moore. "Finding community structure in very large networks". In: *Physical review E* 70.6 (2004), p. 066111.

[2] Lawrence F Katz et al. "Changes in the wage structure and earnings inequality". In: *Handbook of labor economics*. Vol. 3. Elsevier, 1999, pp. 1463–1555.

[3] Frank MacCrory et al. "Racing with and against the machine: changes in occupational skill composition in an era of rapid technological advance". In: (2014).

[4] United States. Employment Service. *Dictionary of Occupational Titles 4th ed. rev. 1991*. Washington D.C.: United States Department of Labor, 1991.