Analysis of Career Pathways Based on the Social Network Analysis

Xulu Wang

Master of Science in Data Science and Analytics (DSAN) Program, Georgetown University, Washington, D.C., United States

Contents

1. Abstract

Changing careers is quite an unusual event. There are always ways to move from one position to another, whether within the same industry or across industries. This study tries to figure out the features and patterns of trajectories from one occupation to another one. Due to the advantages of social network analysis (SNA) for tracking the evolution of a behavior or state, social network analysis is used in this study. The network is constructed with different occupations, which are defined by the 2002 Census Occupation Code, as nodes. Each node has attributes representing the importance of some skills. An edge between A and B represents the existence of transition from occupation A to B, and the weight of an edge is proportional to the number of times the transition.

Some community detection algorithms are used to attempt to explore potential clustering relationships between different occupations. Based on this community detection, by comparing to the clustering results derived from K-means, it is conducted that career transitions do not exhibit significant clustering based on required skills.

This study finds that occupations that have low entry requirements are linked to a wider variety of other occupations, while the flexibility to enter and leave positions leads to high outflow rates in these occupations; occupations with high entry requirements are linked to a more limited number of occupations, but they are less mobile and have the ability to absorb recruits.

Keywords: Social network analysis, Career trajectory

2. Introduction

Switching to a different work is quite common for people, no matter whether these works are in the same industry or not. It is claimed that people usually change jobs seven times in their lifetime, but there's no real data to support it. However, according to a poll conducted by Harris Poll in 2021 exclusively for Fast Company, it shows that more than half of American employees were considering changing their careers that year, and 44% had already planned to make the switch (Dishman 2021). LinkedIn published a blog with the topic of the most common career transitions for recruiters. It found that approximately half (51%) of former recruiters still stayed in the field of human resources only making a transition to another HR role. The top two fields that the former recruiters turned to outside human resources are sales and business development (LinkedIn 2023). From there, it can be said that making a career change to a lesser or greater extent is a choice for many people, and discovering the characteristics and patterns is a topic worth exploring.

Some research aims at the reasons for career change for some specific groups, either in a particular field, age group, gender, etc. For people who have fieldwork experiences, the existence of violations of rules defining appropriate conduct would result in a change of career path. One of the reasons that encourage people to pursue their academic work is productive and enjoyable field experiences. Contrarily, negative field experiences would directly lead to career stalling, moves, or leaving (Nelson et al. 2017). For workers in STEM, even the birth of a child can affect the career path more than some other fields. 43% of new mothers and 23% of new fathers leave full-time STEM employment switching to part-time work or exiting the labor force after their first child (Cech and Blair-Loy 2019). Moreover, the choice of career path is affected by the perception gained during childhood. Children's occupational activities are affected by what they perceived from academic and social activities, which brings them the sense of self efficacy (Bandura et al. 2001).

Some tools including different models and methodologies are utilized for job mobility-related topics. The article (Xu et al. 2015) contributed to make predictions for job change based on the career mobility and daily activity pattern of individuals. They tried to predict whether an employee will change jobs or not sometime in the following months, based on personal resumes and historical check-in records in location-based social networks. With these prior data, some classification algorithms are used, for example, regression trees (CART), support vector machine (SVM), AdaBoost, and random forest. The comprehensive importance of features indicated people might keep job change preferences stable based on their historical experience and follow the pattern of career change. (Xu et al. 2014) proposed a technique to calculate the professional similarity between two individuals by modeling individuals' profiles on LinkedIn as a time-series sequence of positions.

Social network analysis (SNA) is a method based on sociology and graph theory, which is a powerful tool for tracking the evolution of behavior or state. It has been applied to many career-related fields, such as e-recruitment. (Milovanović et al. 2022) used SNA for the preselection of candidates. A network was created by the most frequently used terms. (Toteva and Gourova 2011) built a network based on social network sites. A node could represent a person or a web page. A node with a high degree could be a person who has many connections to high-level specialists or who has wide interests and hobbies or published content on many pages.

In this article, social network analysis is applied to explore the features and patterns of trajectories from one occupation to another one no matter what industry they are in. Data used for this study is from the official website of the U.S. Department of Labor. The datasets include data on occupational transitions from online career profiles and occupation clusters ("Career Pathways Descriptive and Analytical Study Data," n.d.), which provide individual information about career trajectories over time. The network is constructed with different occupations as nodes. The occupations are defined according to the 2002 Census Occupation Code (US Census Bureau 2023), which is a four-character code that identifies the generic occupation. Each node has attributes that come from the dataset representing the importance of some skills. An edge between A and B represents the existence of transition from occupation A to B, and the weight of an edge is proportional to the number of times the transition. The following research questions were formulated:

(1) Which occupations have the highest/lowest outflow/inflow rates?
(2) Are there occupations that are suitable as intermediate bridges from one occupation to another?
(3) Are there any occupations that can be seen as core occupations?

(4) Are there clear categories for career transitions?
(5) Does the occupation transition show a tendency to cluster according to the skills required?
(6) What characterizes a career change in the field of computer technology?

3. Methods

3.1 Data Preparation

The datasets are public use datasets, originally released by the U.S. Department of Labor. The main dataset used for this study records workers' career trajectories and transitions within three years. Their work status, including occupation, wage, sector, etc. was updated every month during the time when the data was collected. The granularity of the occupational division is based on the 2002 Census Occupation Code Lists. The OCC is a 4-digit number representing primary occupations, coded into a contemporary census classification scheme.

Additionally, this dataset contains variables denoting the importance of skills, such as problem-solving, communication, teaching, etc. These variables are extracted on a per-occupation basis.

3.2 Social Network Analysis

Social network analysis is a powerful method for visualizing and analyzing complex connections among lots of entities. It has applications in various fields, including sociology, anthropology, business, information science, and more. It helps to identify patterns of relationships and information flow among them to gain a deeper understanding and insights into the whole structures and the features of the research object.

For this study, a directed network is constructed. All the occupations that exist in the dataset are regarded as nodes. The edges among these nodes represent the existence of career transition. Each edge is weighted with a standard of the number of transitions between two nodes. Figure.1 shows the general overview of the whole network. Due to the length of occupation titles that are too long to be shown clearly in the network, OCC codes are shown instead as node labels.

This network has 472 nodes and 8370 edges in total with density at roughly 0.038. The average clustering coefficient at 0.368 is considered moderate, which means there is a tendency for the neighbors of nodes to be connected with each other and form clusters. The negative degree assortativity coefficient suggests a moderate level of disassortative connection, meaning that nodes with different degrees tend to be connected to each other more frequently than expected by chance.

Table 1. General information about the network

Number of nodes	Number of edges	Density	Average Clustering Coefficient	Degree Assortativity Coefficient
472	8370	0.038	0.368	-0.214

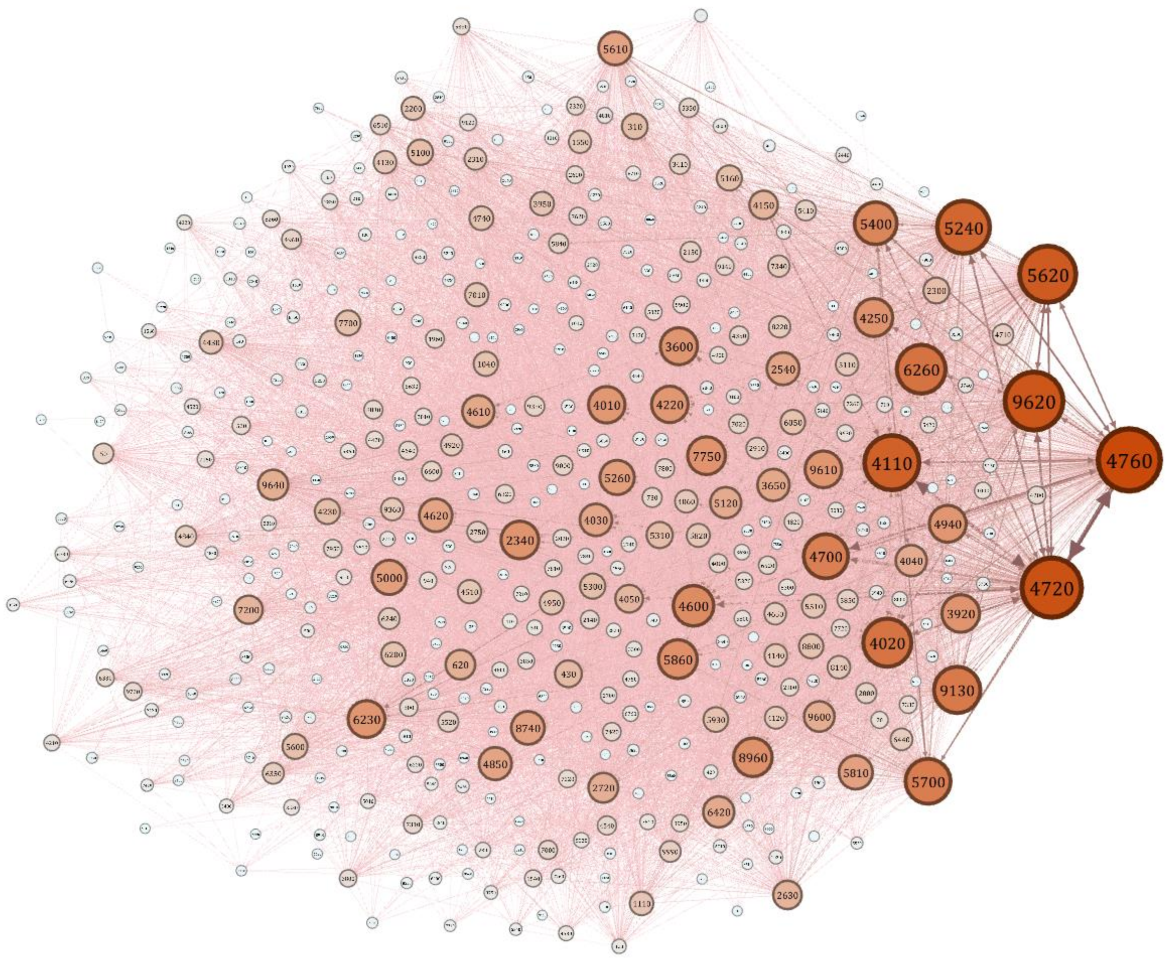

Figure 1 Overview of the whole network with OCC code as node labels (nodes sized and colored by "out-degree")

3.3 Turnover Rate

Here the turnover rate is defined as:

$$TR = \frac{sum(weights\ of\ all\ out - edges)}{sum(weights\ of\ all\ in - edges)}, \qquad sum(weights\ of\ all\ in - edges) \neq 0$$

This turnover rate describes the labor mobility for each occupation. If the value is greater than 1, it means the labor outflow in this occupation is greater than the inflow. If the value is less than 1, the flow is exactly the opposite. If it is equal to 1, it means that the occupation from the perspective of the whole labor market is relatively stable.

3.4 Community Detection

Community detection refers to the process of identifying cohesive groups of nodes within a network where nodes within the same community are more densely connected to each other than to nodes in other communities. The results of community detection typically reveal the underlying network structures. Within this study, the purpose of community detection is to answer whether there are clear clusters for career transitions. In other words, is there a tendency for career transitions to move from one occupation to another, making these occupations connected closer to each other?

The main methods for community detection in this study are the Louvain algorithm and the greedy algorithm aiming for modularity maximization. As shown in the Tabel 2, no algorithm produces significantly superior results, either for directed or undirected graphs. The highest modularity at 0.2136 suggests the presence of a meaningful community structure.

Table 2. Modularity from two algorithms applied to both directed and un-directed networks

Louvain Algorithm + Directed Network	**Louvain Algorithm + Un-directed Network**	**Greedy Algorithm + Directed Network**	**Greedy Algorithm + Un-directed Network**
0.2118	0.2133	0.2136	0.2130

4. Results

4.1 Basic network information

4.1.1 Educational background exaltation

When looking for a job as a graduate with limited work experience, educational background plays an important role to some extent. At the same time, for those who have already started their careers, it is also possible for them to improve their educational background with the aim of searching for a better position or a higher salary. Within the three years of the data tracking, 20.07 percent of respondents succeeded in pursuing educational background exaltation. 68.89 percent started with no degree, and of those people, 0.16 percent of them got a bachelor's degree in three years. 19.54 percent of those who ended up with a bachelor's degree started with an HS diploma or equivalent (no college degree), and 1.93 percent of them started with an associate degree (no bachelor's).

4.1.2 Wage change

After three years, 54.77 percent of all respondents finally got a raise. Among those people who improved their academic degree, 59.61 percent of them got a raise. All respondents got an average of a 3.36 percent increase in wage, and people who improved their academic degrees got an average of a 7.05 percent increase, which is more than twice times than the average.

4.1.3 Turnover rate

More than 58 percent of occupations that have a high turnover rate, in other words, outflow is more significant than inflow, are in the transportation and production industries.

The rates of the five occupations that have the highest centralities are also greater than 1, meaning that while there was a significant labor inflow to these occupations over the three years, more people left these jobs. This result conforms with the flexibility of these occupations.

Occupations having a lower turnover rate mainly concentrate in the office and administrative support industry and construction industry. Computer and information systems managers, lawyers, architects, medical and health services managers, and human resources managers are the top 5 occupations with the highest inflow rate. Within the three years, 21 transitions to and only 1 transition from computer and information systems managers occur.

4.2 Network Centrality

There are five measures for network centrality being used in this study, including in-degree, out-degree, closeness, betweenness, and Eigenvector centrality. When unemployed status and student status are taken into account, these two nodes have the highest centrality under all measures of centrality. This is reasonable because it is common that there will be a short gap between two different jobs when a career transition occurs and the time of two jobs does not fit seamlessly. Also, for those people who pursued educational background exaltation, there must be a period of student life between the different jobs. Thus, there are a large number of edges pointing in or out of the 'unemployed' and 'student' nodes, which contribute to making them the highest centrality nodes. In order to obtain clearer information on career transitions, these two nodes are excluded from the study.

Generally, five occupations, including (1) retail salespersons, (2) cashiers, (3) waiters and waitresses, (4) stock clerks and order fillers, (5) laborers and freight, stock, and material movers, hand, take up the top five positions in terms of all measures of in-degree, out-degree, and closeness in both inward and outward directions. This means that these occupations have a great deal of flexibility. There are no high requirements to meet to switch to or from these occupations. Particularly, for the betweenness centrality, the values of nodes are all less than 0.055, meaning that the node's position as a bridge or intermediary in facilitating career transition is low. On the other hand, all the occupations with the least centrality, such as (1) ship engineers, (2) subway, streetcar, and other rail transportation workers, and (3) fabric and apparel patternmakers, are mainly concentrated in the production and transportation industry.

Table 3. Five Occupations with Highest Centralities

	In-degree	**Out-degree**	**Closeness Centrality (inward)**	**Closeness Centrality (outward)**	**Betweenness Centrality**	**Eigenvector Centrality**
1	Retail salespersons	Laborers and freight, stock, and material movers, hand	Retail salespersons	Laborers and freight, stock, and material movers, hand	Retail salespersons	Cashiers
2	Cashiers	Retail salespersons	Cashiers	Retail salespersons	Laborers and freight, stock, and material movers, hand	Waiters and waitresses
3	Waiters and waitresses	Cashiers	Waiters and waitresses	Cashiers	Waiters and waitresses	Retail salespersons
4	Stock clerks and order fillers	Driver/sales workers and truck drivers	Customer service representatives	Driver/sales workers and truck drivers	Stock clerks and order fillers	Receptionists and information clerks
5	Customer service representatives	Waiters and waitresses	Stock clerks and order fillers	Waiters and waitresses	Cashiers	Customer service representatives

4.3 Community Detection

The result of community detection that gets the highest modularity contains six communities in total. Figure 2. displays the general distribution of nodes in different communities. The occupations in the community containing only two nodes are (1) Shoe and leather workers and repairers and (2) shoe machine operators and tenders. Three occupations including (1) locomotive engineers and operators, (2) railroad conductors and yardmasters, and (3) subway, streetcar, and other rail transportation workers are in another community. The four remaining communities focus on construction, office and administrative support, production, and healthcare, respectively.

To demonstrate whether the career transition is related to the skills required, the K-means clustering algorithm is implemented to get the same number of clusters as in community detection. If the ways how these occupations are partitioned based on the community detection and K-means are highly similar to each other, it means that career transitions are relevant to the skills, and career transitions tend to occur between jobs requiring similar skills. As shown in Figure. 3, the nodes are more evenly distributed among communities than the communities coming from the greedy algorithm.

Normalized mutual information (NMI) is a measure used to evaluate the similarity between two partitions of a dataset. The value derived from these two partitions is 0.22. This denotes a moderate level of agreement between the two partitions. It indicates that there is a similarity between them at a lower level, but there are more differences in the way the clusters are defined. It is concluded that the career transition is not closely related to the skills required for occupations. Skills could be a factor that impacts the direction of job change, but it is not a significant one.

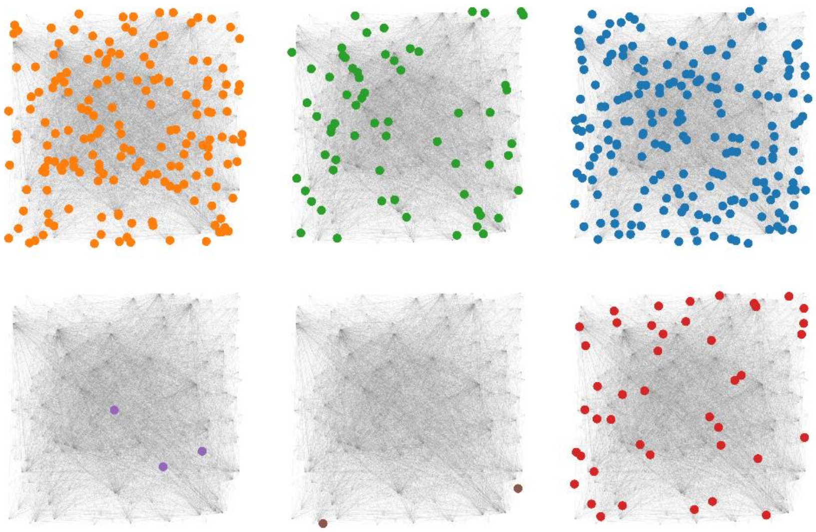

Figure 2 Distribution of the nodes in communities derived from the greedy algorithm

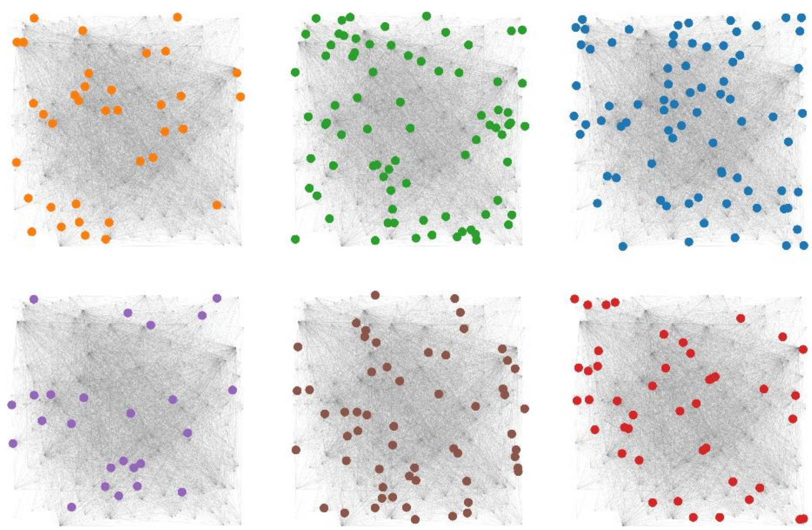

Figure 3 Distribution of the nodes in communities derived from K-means

5. Discussion

As the general metrics imply, the density of this career transition network is low, meaning that it is not easy to switch between any two occupations.

(1) Dredge, excavating, and loading machine operators and (2) chefs and head cooks are the occupations that have the highest and lowest outflow rates, respectively. Labor outflow from dredge, excavating, and loading machine operators is 8.5 times greater than labor

inflow. The majority of other occupations that have high outflow rates are in the transportation industry. However, some occupations requiring scientific or professional qualifications, such as computer and information systems managers, lawyers, and architects, have the highest inflow rate. The labor force entering these industries far exceeds the number of people leaving them. This is understandable, as embarking on these occupations is more demanding to be professional. They require a great deal of time and experience to reach the threshold for starting these professions at an entry level, and at the same time the income in these professions is relatively substantial than other professions, hence there is a tendency for people to remain in these professions or to prepare to enter them.

There is no such an occupation that is able to serve as an all-purpose position. It means none of the occupations demonstrated significant characteristics that would help with career transitions. Intuitively, people usually think of computer-related jobs as omnipotent because computer technologies are needed in many industries. However, in addition to the low labor outflow characteristic of computer-related professions, most of the career transitions achieved by virtue of knowledge related to computers, also occur in the computer field. Despite the fact that it allows for industry transitions, the occupational content is very similar. On the contrary, it is occupations such as retail salespersons and cashiers that connect a large number of other jobs. Because of their flexibility, these occupations are good choices for the transition between two careers.

If viewed at a distance from the other occupations, those occupations that are easy to get into are the core of the network, because they provide a considerable number of jobs. But in terms of the stability of the labor force in the industry, those occupations that require better professional skills or qualifications occupy an important position.

Career transitions show a pronounced homogeneity of categories in the domains of construction, office and administrative support, production, and healthcare. Due to the results of the community detection and the clustering results of the K-means made based on the importance of skills, the division of occupational categories did not show significant homogeneity with the division of skills. It implies that there are other factors that influence career change besides skills.

References

[1] Dishman, Lydia. 2021. "Is Now a Good Time to Change Careers? More Workers Are Feeling Good about It." Fast Company, February 23, 2021. https://www.fastcompany.com/90607167/is-now-a-good-time-to-change-careers-more-workers-are-feeling-good-about-it.

[2] "The Most Common Career Transitions for Recruiters | LinkedIn." n.d. https://www.linkedin.com/business/talent/blog/talent-strategy/most-common-career-transitions-for-recruiters.

[3] Nelson, Robin G, Julienne N Rutherford, Katie Hinde, and Kathryn BH Clancy. 2017. “Signaling Safety: Characterizing Fieldwork Experiences and Their Implications for Career Trajectories.” American Anthropologist 119 (4): 710–22.

[4] Cech, Erin A, and Mary Blair-Loy. 2019. “The Changing Career Trajectories of New Parents in STEM.” Proceedings of the National Academy of Sciences 116 (10): 4182–87.

[5] Bandura, Albert, Claudio Barbaranelli, Gian Vittorio Caprara, and Concetta Pastorelli. 2001. “Self-Efficacy Beliefs as Shapers of Children’s Aspirations and Career Trajectories.” Child Development 72 (1): 187–206.

[6] Xu, Huang, Zhiwen Yu, Hui Xiong, Bin Guo, and Hengshu Zhu. 2015. “Learning Career Mobility and Human Activity Patterns for Job Change Analysis.” In 2015 IEEE International Conference on Data Mining, 1057–62. https://doi.org/10.1109/ICDM.2015.122.

[7] Xu, Ye, Zang Li, Abhishek Gupta, Ahmet Bugdayci, and Anmol Bhasin. 2014. “Modeling Professional Similarity by Mining Professional Career Trajectories.” In Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 1945–54.

[8] Milovanović, Stevan, Zorica Bogdanović, Aleksandra Labus, Marijana Despotović-Zrakić, and Svetlana Mitrović. 2022. “Social Recruiting: An Application of Social Network Analysis for Preselection of Candidates.” Data Technologies and Applications 56 (4): 536–57.

[9] Toteva, Kostadinka, and Elissaveta Gourova. 2011. “Social Network Analysis in Professional E-Recruitment.” In Third International Conference on Software, Services and Semantic Technologies S3T 2011, 75–80. Springer.

[10] “Career Pathways Descriptive and Analytical Study Data.” n.d. DOL. https://www.dol.gov/agencies/oasp/evaluation/data-repo/career-pathways-da-data#apdu-5.

[11] US Census Bureau. 2023. “Census.Gov.” Census.Gov. July 10, 2023. https://www.census.gov/.