Interactive Visual Exploration of Longitudinal Historical Career Mobility Data

Yifang Wang, Hongye Liang, Xinhuan Shu, Jiachen Wang, Ke Xu, Zikun Deng, Cameron Campbell, Bijia Chen, Yingcai Wu, Huamin Qu

HKUST VisLab

The Hong Kong University of Science and Technology

Zhejiang University

Zhejiang Lab

NYU

Renmin University of China
Project Background

Quantitative Historical Datasets
CGED-Q
Career Mobility
• Digital Humanities
• History
• Sociology
• Demography
• ...

• Career
• Health
• Family
• Migration
• ...

Historical Quantitative Datasets
China Government Employee Database-Qing (CGED-Q)

- One-year project
- CGED-Q records the career trajectories of over 340,000 government officials in the bureaucracy of Qing China from 1760 to 1912

- Career Mobility: the study of career trajectories and the factors influencing them.

Social Ladders

Social Relationships / Social Inequality
Existing Career Mobility Analysis

Hout, 1983

<table>
<thead>
<tr>
<th>Father's Occupation</th>
<th>Son's Occupation</th>
<th>Nonmanual Upper</th>
<th>Nonmanual Lower</th>
<th>Manual Upper</th>
<th>Manual Lower</th>
<th>Farm</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inflow Percentages</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Upper Nonmanual</td>
<td>34.5</td>
<td>17.6</td>
<td>3.4</td>
<td>1.8</td>
<td>1.8</td>
<td>14.7</td>
<td></td>
</tr>
<tr>
<td>Lower Nonmanual</td>
<td>17.7</td>
<td>9.7</td>
<td>8.8</td>
<td>2.1</td>
<td>1.1</td>
<td>11.3</td>
<td></td>
</tr>
<tr>
<td>Upper Manual</td>
<td>19.5</td>
<td>21.9</td>
<td>32.6</td>
<td>4.3</td>
<td>4.8</td>
<td>20.5</td>
<td></td>
</tr>
<tr>
<td>Lower Manual</td>
<td>31.4</td>
<td>30.8</td>
<td>29.4</td>
<td>41.8</td>
<td>30.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Farm</td>
<td>10.0</td>
<td>12.0</td>
<td>16.8</td>
<td>20.2</td>
<td>80.9</td>
<td>23.4</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td></td>
<td>100.0</td>
</tr>
</tbody>
</table>

Jarvis and Song, 2017

Table 4. Log-Linear Model Fit Statistics

<table>
<thead>
<tr>
<th>Model</th>
<th>df</th>
<th>$L^2$</th>
<th>p-value</th>
<th>BIC</th>
<th>$\Delta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline Models</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0a: Origin + Destination + Period</td>
<td>27,872</td>
<td>244,675</td>
<td>&lt;.01</td>
<td>-61,245</td>
<td>09.0</td>
</tr>
<tr>
<td>1a: 0a + Origin + Period + Destination + Period</td>
<td>27,380</td>
<td>239,995</td>
<td>&lt;.01</td>
<td>-59,449</td>
<td>08.5</td>
</tr>
<tr>
<td>2a: 1a + Parameter-ABCD</td>
<td>27,282</td>
<td>27,431</td>
<td>.28</td>
<td>-271,052</td>
<td>18.2</td>
</tr>
<tr>
<td>Period-Varying Typological Mobility Models</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3a: 2a + Period + Parameter-A</td>
<td>26,992</td>
<td>24,874</td>
<td>.99</td>
<td>-270,328</td>
<td>14.2</td>
</tr>
<tr>
<td>4a: 2a + Period + Parameter-AB</td>
<td>26,958</td>
<td>24,723</td>
<td>.99</td>
<td>-270,096</td>
<td>14.1</td>
</tr>
<tr>
<td>5a: 2a + Period + Parameter-ABC</td>
<td>26,944</td>
<td>24,661</td>
<td>.99</td>
<td>-270,017</td>
<td>14.1</td>
</tr>
<tr>
<td>6a: 2a + Period + Parameter-ABCD</td>
<td>26,940</td>
<td>24,048</td>
<td>.99</td>
<td>-269,986</td>
<td>14.1</td>
</tr>
<tr>
<td>7a: 2a + Period + Parameter-APC</td>
<td>27,270</td>
<td>25,669</td>
<td>.99</td>
<td>-272,650</td>
<td>16.0</td>
</tr>
<tr>
<td>Women (N = 47,100)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Baseline Models</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0b: Origin + Destination + Period</td>
<td>27,872</td>
<td>174,593</td>
<td>&lt;.01</td>
<td>-126,434</td>
<td>06.6</td>
</tr>
<tr>
<td>1b: 0b + Origin + Period + Destination + Period</td>
<td>27,380</td>
<td>106,191</td>
<td>&lt;.01</td>
<td>-126,465</td>
<td>05.7</td>
</tr>
<tr>
<td>2b: 1b + Parameter-ABCD</td>
<td>27,282</td>
<td>19,403</td>
<td>.99</td>
<td>-274,074</td>
<td>15.2</td>
</tr>
<tr>
<td>Period-Varying Typological Mobility Models</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3b: 2b + Period + Parameter-A</td>
<td>26,992</td>
<td>18,287</td>
<td>.99</td>
<td>-272,193</td>
<td>12.3</td>
</tr>
<tr>
<td>4b: 2b + Period + Parameter-AB</td>
<td>26,958</td>
<td>18,154</td>
<td>.99</td>
<td>-271,939</td>
<td>12.2</td>
</tr>
<tr>
<td>5b: 2b + Period + Parameter-ABC</td>
<td>26,944</td>
<td>18,102</td>
<td>.99</td>
<td>-271,862</td>
<td>12.1</td>
</tr>
<tr>
<td>6b: 2b + Period + Parameter-ABCD</td>
<td>26,940</td>
<td>18,077</td>
<td>.99</td>
<td>-271,444</td>
<td>12.1</td>
</tr>
</tbody>
</table>

- Social scientists
- Group-level analysis
Existing Career Mobility Analysis

- Hypothesis-driven and ill-suited for exploration

- Social scientists
  Group-level analysis
Existing Career Mobility Analysis

**• Hypothesis-driven and ill-suited for exploration**

**Qu et al., 2016**

![Fig. 8. The multi-task and multi-feature prediction model](image)

**Xu et al., 2015**

![Graph](image)

**• Social scientists**
Group-level analysis

**• Data scientists**
Individual-level analysis
Existing Career Mobility Analysis

**Hypothesis-driven and ill-suited for exploration**

- Social scientists
  - Group-level analysis

- Data scientists
  - Individual-level analysis

**Do not target career mobility -- emphasizes the characteristics of groups and more complicated social relationships**

- Qu et al., 2016

- Jarvis and Song, 2017

- Hout, 1983
Existing Career Data Visualization (1/2)

• Network summarization

Fung et al., 2016

• Sequence summarization

Guo et al., 2018

• Similarity comparison

Du et al., 2016
Existing Career Data Visualization (2/2)

- Multi-task analysis

Khulusi et al., 2019

Zhang and Wang, 2019
Existing Career Data Visualization (2/2)

- Support a limited number of career comparisons and lose the overall context
- Are limited by the short time range
- Multi-task analysis
Data Description
Data Description

- **Timestamp**: The year and season covered by the record
- **Name**: The official’s real name in the Qing dynasty
- **Unique ID**: A 12-character unique identifier of each official generated by experts
Data Description

- **Timestamp**: The year and season covered by the record
- **Name**: The official’s real name in the Qing dynasty
- **Unique ID**: A 12-character unique identifier of each official generated by experts
- **Birthplace**: The geographic origin of the official
- **Family Background**: A identity indicating whether the official was associated with the imperial lineage
- **Ethnicity**: Three types of officials are identified based on ethnicities: Manchu, Mongol, or Han
- **Exam Degree** (科举结果): The examination or purchased degree held by the official. Those with high examination degrees were political elites
- **Job Location**: The geographical location of the official’s current job
- **Job Department**: The department in the bureaucracy where the official works. We classified them into fifteen categories according to experts’ suggestions
- **Job Level** (品级): The administrative rank of the job in the bureaucratic hierarchy, represented by a number (ranging from 10 to 1 with 0.5 as a step).
Challenges (1/3)

- How to visualize a large volume of longitudinal career data with a complex data structure?

- Temporal
  - 1760 ~ 1912

- Network
  - Colleagues
  - Townsmen
  - Classmates

- Multi-attributes
  - Personal Info
  - Job Info
Challenges (2/3)

• How to visualize a large volume of longitudinal career data with a complex data structure?
• How to extract and highlight social groups and social relationships from this large dataset?

• Predefined Group vs. Latent Group
• Semi-automatic
Challenges (3/3)

- How to visualize a large volume of longitudinal career data with a complex data structure?
- How to extract and highlight social groups and social relationships from this large dataset?
- How to support multi-level mobility analysis and reasoning?

- Overall Level
- Group Level
- Individual Level
Tasks and System Design
Experts and Tasks

- Experts
- What are the general characteristics of career mobility?
- What special features do the groups with vertical movements have at different time periods?

Group
- What are the characteristics of different social groups?
- What is the mobility pattern for each group?

Individual
- What are the mobility characteristics for different individuals?
- How do the mobility patterns of each individual and his social relationships change over time?
Visual Design

Data Filtering
Visual Design

Group & Person Recommendation
Visual Design

Detailed Career Mobility Analysis
Visual Design

Encoding:

Promotion  Steady  Demotion

Promotion Rate:

\[ UVMR_{(t_i, t_j)} = \frac{|PromotionGroup|}{|All|} \]

Demotion Rate:

\[ DVMR_{(t_i, t_j)} = \frac{|DemotionGroup|}{|All|} \]
Visual Design

Promotion vs. Steady

Demotion vs. Steady
Visual Design

Birthplace

Job Dept.

Exam Degree

Ethnicity

POPULATION DISTRIBUTION

Zhejiang, Anhui, Hunan, Guangzhou, Fujian
Visual Design

Birthplace → Hunan

Job Dept. → Exam Degree

Ethnicity

POPULATION DISTRIBUTION

Zhejiang  Anhui  Hunan  Guangzhou  Fujian
Visual Design

Birthplace

Hunan

Job Dept. → Hanlin Academy

Elite

Exam Degree

Qiren

Ethnicity
Latent Group Detection (MinDL\textsuperscript{1})

\[ L(C) = \sum_{(P,G) \in C} ||P|| + \left( \alpha \sum_{(P,G) \in C} \sum_{s \in G} ||edits(s, P)|| \right) + \lambda ||C|| \]

Lowest Job Level | Group Pattern | Highest Job Level

Visual Design

Detailed Career Mobility Analysis (2/2)
Visual Design

Job-level Mode
Visual Design

Job-level Mode

Department Mode

• Color Scheme
Relative-time Mode: align the career starting year
Visual Design

• Flow Design

• Overall Flow
Visual Design

• Flow Design

• Group Subflow

• Overall Flow
Visual Design

- Flow Design
- Career Threads
- Group Subflow
- Overall Flow
Visual Design

• Flow Design
Visual Design

- Flow Design + Social Relations of the Key Person (KP)

Key Person (KP)
Related Persons
Evaluation
Case Study

- Invite four internal experts to freely explore the system
- Political Elites (i.e., 进士) from South and East China
In this video, we demonstrate a case study to explore the political elites from South and East China.
Expert Interview

- Invite four internal experts and four new experts

Advantages

- More efficient workflow to explore the dataset from different LODs with a user-friendly interface
- Advanced techniques (e.g., latent group and dynamic network) to find interesting insights
- Intuitive visual representations and interactions

Suggestions

- More detailed information
  - Add the departments with provinces
  - Add raw data table
- Replacing the context menu with a switch button
Conclusions

Contributions

- A hierarchical problem characterization of historical career mobility analysis
- A visual analytics system *CareerLens*
- A novel flow design with a multi-scale approach (overall mobility flow, group subflow, and individual career threads)
- Two case studies and a longitudinal investigation to evaluate the system and receive positive feedback

Future Works

- Open to the public
- Multi-attributes for latent group detection
- Improvement of scalability of flow design
- Correlation of career mobility and different events
For More Information

- I’m currently looking for collaborators in both humanity and social science to work together! For more information, please visit: [wangyifang.top](http://wangyifang.top) or contact yifang.wang@connect.ust.hk

- The work is partially supported by Hong Kong Research Grants Council (RGC) General Research Fund (GRF) grant 16213317, National Natural Science Foundation of China (62072400), Zhejiang Provincial Natural Science Foundation (LR18F020001), and the 100 Talents Program of Zhejiang University. Construction of the CGED-Q was supported by Hong Kong RGC GRF 16600017
Interactive Visual Exploration of Longitudinal Historical Career Mobility Data

Yifang Wang$^{1,2}$ Hongye Liang$^{2,3}$ Xinhuan Shu$^{1}$ Jiachen Wang$^{2,3}$ Ke Xu$^{4}$
Zikun Deng$^{2,3}$ Cameron Campbell$^{1}$ Bijia Chen$^{5}$ Yingcai Wu$^{2,3}$ Huamin Qu$^{1}$

I’m currently looking for collaborators in both social science and humanities to work together!

For more information, please visit: wangyifang.top or contact yifang.wang@connect.ust.hk