On Analyzing Job Hop Behavior and Talent Flow Networks
Motivation

Job hopping is very common.
- People aged 18 to 35 tenure with their current employer: about 13 months.
  [US Department of Labor]

In the past and now, we conduct surveys to collect the relevant data.
- Time consuming
- Costly
Online Professional Network

With OPN becoming popular, job hopping insights be acquired through a data science approach.

OPN offers:
(a) Individual level,
(b) Historical, and
(c) Network data

OPN data is:
- Noisy
- Incomplete
Objectives

Data science study job hop behavior and talent flow.

- To analyze job hop behavior with job attributes.
- To analyze job hop behavior with job promotion/demotion.
- To determine interesting jobs using job hop behavior.
Outline of Presentation

Dataset

Job & Organization Hop Graphs

Data Science Analysis
- Distribution Analysis of Graph Metrics
- External Hop Analysis
- Promotion and Demotion Analysis
- Network Analysis

Conclusion
Dataset

Data extracted from an Online Professional Network:
- User profiles
- Organization profiles

Profiles from a developed city (covering largely the PMET job market)

Profiles are public

Profiles are active (with least non-empty education or skills fields)

<table>
<thead>
<tr>
<th>Statistics</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of user profiles</td>
<td>2,574,502</td>
</tr>
<tr>
<td>No. of active user profiles</td>
<td>490,200</td>
</tr>
<tr>
<td>No. of organizations</td>
<td>145,524</td>
</tr>
<tr>
<td>No. of industries</td>
<td>147</td>
</tr>
</tbody>
</table>
User Level Hop Behavior: External vs Internal Hops

**Hop:** A transition from one job to another with non-overlapping time period.
- Missing month info
- Missing end date info
- Multiple jobs at one time

**External hop**
- A hop which involves different organizations.
- Job title may or may not change

**Internal hop**
- A hop which involves the same organization.
- Noise removal: internal hops that involve same job title in same organization.
Job Hop Graph

Node: \((t, i)\)
- \(t\) denotes job title, \(i\) denotes industry
- Each node has a minimum support of 10 users who ever hold the job

Edge: \((t, i) \rightarrow (t', i')\)
- If there is at least one user moving from job \((t, i)\) to another job \((t', i')\)
- Edge weight: # users who ever make the move

Diagram:
- Intern, IT
- Sys Analyst, IT
- Mgt Trainee, Banking
- Project Mgr, IT
Metrics of Job Hop Graph

Work experience of \((t, i)\): Amount of working experience required for the job \((t, i)\).
\[
wk_{\text{exp}}(t, i) = \text{Avg}_{(p,c)} wk_{\text{exp}}(p, t, c, i)
\]
\[
wk_{\text{exp}}(p, t, c, i) = \text{start}_\text{date}(p, t, c, i) - \text{grad}_\text{date}(p)
\]
- **High \(wk_{\text{exp}}\) jobs:** “Professor”, “Managing Director”, “CEO”
- **Low \(wk_{\text{exp}}\) jobs:** “Intern”, “Teaching Assistant”

Job age of \((t, i)\): Duration from the start of \((t, i)\) until the current date.
\[
\text{job}_\text{age}(t, i) = \text{Avg}_{(p,c)} \text{job}_\text{age}(p, t, c, i)
\]
\[
\text{job}_\text{age}(p, t, c, i) = \text{current}_\text{date} - \text{start}_\text{date}(p, t, c, i)
\]
- **Ancient jobs:** “Director”, “System Engineer”, “Division Manager”
- **Young jobs:** “Data Scientist”, “Media Analyst”
Other Metrics

External hop fraction of some user group $g$: Fraction of users of group $g$ moving to different organizations.

$$\%_{\text{external\_hop}}(g) = \frac{|P^g_{c \rightarrow c'}|}{|P^g_{c \rightarrow c'}| + |P^g_{c \rightarrow c}|}$$

$g$: users sharing the same working experience, or job age, or skill count.

Job level of a job $t$ in a company $c$:

$$\text{job\_level}(t, c) = \frac{1}{|P_{t, c}|} \sum_{p \in P_{t, c}} \text{wk\_exp}(p, t, c, i)$$

Job level gain of a user $p$:

$$\text{level\_gain}((p, (t, c) \rightarrow (t', c'))) = \text{job\_level}(t, c) - \text{job\_level}(t', c')$$
Network Metrics of Job Hop Graph

- **In-degree centrality of a job** \((t, i)\)
  - Unweighted

- **Out-degree centrality of a job** \((t, i)\)
  - Unweighted

- **Pagerank centrality of a job** \((t, i)\)
  - Weighted transition
  - Transition probability \(\propto \frac{\text{outedge weight of } (t,i)\rightarrow (t',i')}{\text{total outedge weight of } (t,i)}\)
  - A measure of global competitiveness
Organization Hop Graph

Node: Company $c$
- $c$ has a minimum support of 10 users

Edge: $c \rightarrow c'$
- If there is at least one user moving from company $c$ to company $c'$
- Edge weight: # users who ever make the move
Network Metrics of Job Hop Graph

- In-degree centrality of a company $c$

- Out-degree centrality of a company $c$

- Pagerank centrality of a company $c$
  - Weighted transition
  - Transition probability $\propto \frac{\text{out-edge weight of } c \rightarrow c'}{\text{total out-edge weight of } c}$
  - A measure of global competitiveness
Jobs of work experience of 5 years or less attract more people.

Very few people who have worked for over 20 years.

Most jobs have been established for 1 year or more.

Very few jobs have been established for more than 20 years.
External Hop Analysis: Users of different work experience/job age

- For all job age groups under 20 years: Younger people are less loyal to their companies than older ones.

- For job age groups above 20 years: Older people (15 years or more) have larger proportion of external hops.
External Hop Analysis: Users of different work experience/skill count

- Within each skill count group, younger people with less work experiences are more likely to move out to other organizations—except again for the very seasoned people with work experience of 25 years.

- No clear association between the external hop fraction and # skills.

Less skills users  More skills users
Promotion and Demotion Analysis

\[ \text{level_gain}(p, (t, c) \rightarrow (t', c')) = \text{job_level}(t, c) - \text{job_level}(t', c') \]

Promotion = +ve level gain
Demotion = −ve level gain

To focus on known jobs, we require both source and target jobs of each hop to have minimum support of 10.

<table>
<thead>
<tr>
<th></th>
<th>Promotion</th>
<th>Demotion</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>External</td>
<td>4,813</td>
<td>1,720</td>
<td>6,533</td>
</tr>
<tr>
<td>Internal</td>
<td>3,891</td>
<td>387</td>
<td>4,278</td>
</tr>
<tr>
<td>Total</td>
<td>8,704</td>
<td>2,107</td>
<td>10,811</td>
</tr>
</tbody>
</table>

\[ \text{Prob}(\text{promotion}) > \text{Prob}(\text{demotion}) \]
\[ \text{Prob}(\text{promotion}|\text{int hop}) = 91\% > \text{Prob}(\text{promotion}|\text{ext hop}) = 74\% \]
Promotion and Demotion Analysis

External hops may work better or worst.
When can one get promoted?

Promotion hops most commonly happen after a person works for 1–2 years.
## Network Analysis

A giant component exists for job graph and organization graph.

<table>
<thead>
<tr>
<th>Metric</th>
<th>Job graph</th>
<th>Organization graph</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Basic</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of nodes</td>
<td>27,451</td>
<td>6,139</td>
</tr>
<tr>
<td>No. of edges</td>
<td>93,283</td>
<td>173,993</td>
</tr>
<tr>
<td>Sparsity of adjacency matrix</td>
<td>0.01%</td>
<td>0.46%</td>
</tr>
<tr>
<td><strong>Strongly connected component</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of SCCs</td>
<td>15,455</td>
<td>415</td>
</tr>
<tr>
<td>Size of the largest SCC</td>
<td>11,950 (43.53%)</td>
<td>5,725 (93.26%)</td>
</tr>
<tr>
<td>Size of the 2nd-largest SCC</td>
<td>4 (0.01%)</td>
<td>1 (0.02%)</td>
</tr>
<tr>
<td><strong>Weakly connected component</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No. of WCCs</td>
<td>882</td>
<td>15</td>
</tr>
<tr>
<td>Size of the largest WCC</td>
<td>25,747 (93.79%)</td>
<td>6,125 (99.77%)</td>
</tr>
<tr>
<td>Size of the 2nd-largest WCC</td>
<td>13 (0.05%)</td>
<td>1 (0.02%)</td>
</tr>
</tbody>
</table>

SCC = strongly-connected component, WCC = weakly-connected component
Centrality Distributions

Long tail centrality distributions
Jobs: Indegree vs Pagerank
Data science approach to job hop behaviour using OPN data.

Issues: (a) data availability, (b) data bias, (c) data noise

Future Works:
- Resume and Job Post data
- More countries and cities
- Job title normalization, e.g., “Research Manager” vs “Manager, Research”
- Job recommendation/search: When to change job? What job to change?
Thank you

Personalized Participatory Nation

School of Information Systems
Singapore Management University
80 Stamford Road
Singapore 178902
Tel: 65 6808 5227

larc.smu.edu.sg  facebook.com/larc.cmu.smu  twitter.com/larc_cmu_smu