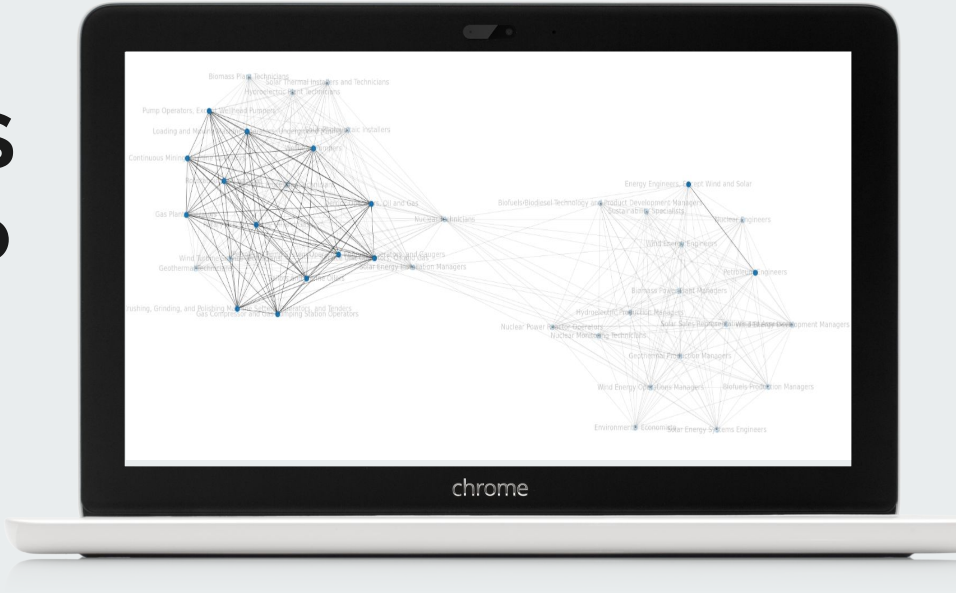
Skill Barriers to a “Just Transition” of US Fossil Fuel Workers to Emerging Green Jobs

Name: Xizhi Wu
Email: xiw183@pitt.edu



Data

O*NET Dataset: https://www.onetonline.org/find/descriptor/browse/2.A/2.A.1

Arm-Hand_Steadiness

Importance	Level	Job Zone	Code	Occupation
81	71	5	29-1022.00	Oral and Maxillofacial Surge
81	57	2	51-6051.00	Sewers, Hand
78	66	5	29-1021.00	Dentists, General
75	59	2	51-6093.00	Upholsterers
75	59	3	49-9064.00	Watch and Clock Repairers

Data

Resume Data: 39 occupations are picked, related to fossil fuel or green jobs

	S_Soc	D_Soc	Date	Counts
0	15-1252	15-1252	2017	7800
1	15-1252	15-1252	2016	7513
2	15-1252	15-1252	2015	7290
3	15-1252	15-1252	2018	6976
4	15-1252	15-1252	2014	6276
...	...	...	...	...
1149364	25-1081	39-6012	2010	1
1149365	25-1081	39-6012	2011	1
1149366	25-1081	39-6012	2014	1

'Rotary Drill Operators, Oil and Gas',

'Roustabouts, Oil and Gas',

'Service Unit Operators, Oil and Gas',

'Petroleum Engineers',

'Solar Photovoltaic Installers',

'Wind Energy Operations Managers',

'Hydroelectric Production Managers',

Method: Building a Skill Network *[1]*

$$\mathrm{rca}(j,s) = \frac{\mathrm{onet}(j,s) / \sum_{s'\in S} \mathrm{onet}(j,s')}{\sum_{j'\in J} \mathrm{onet}(j',s) / \sum_{j'\in J, s'\in S} \mathrm{onet}(j',s')}$$

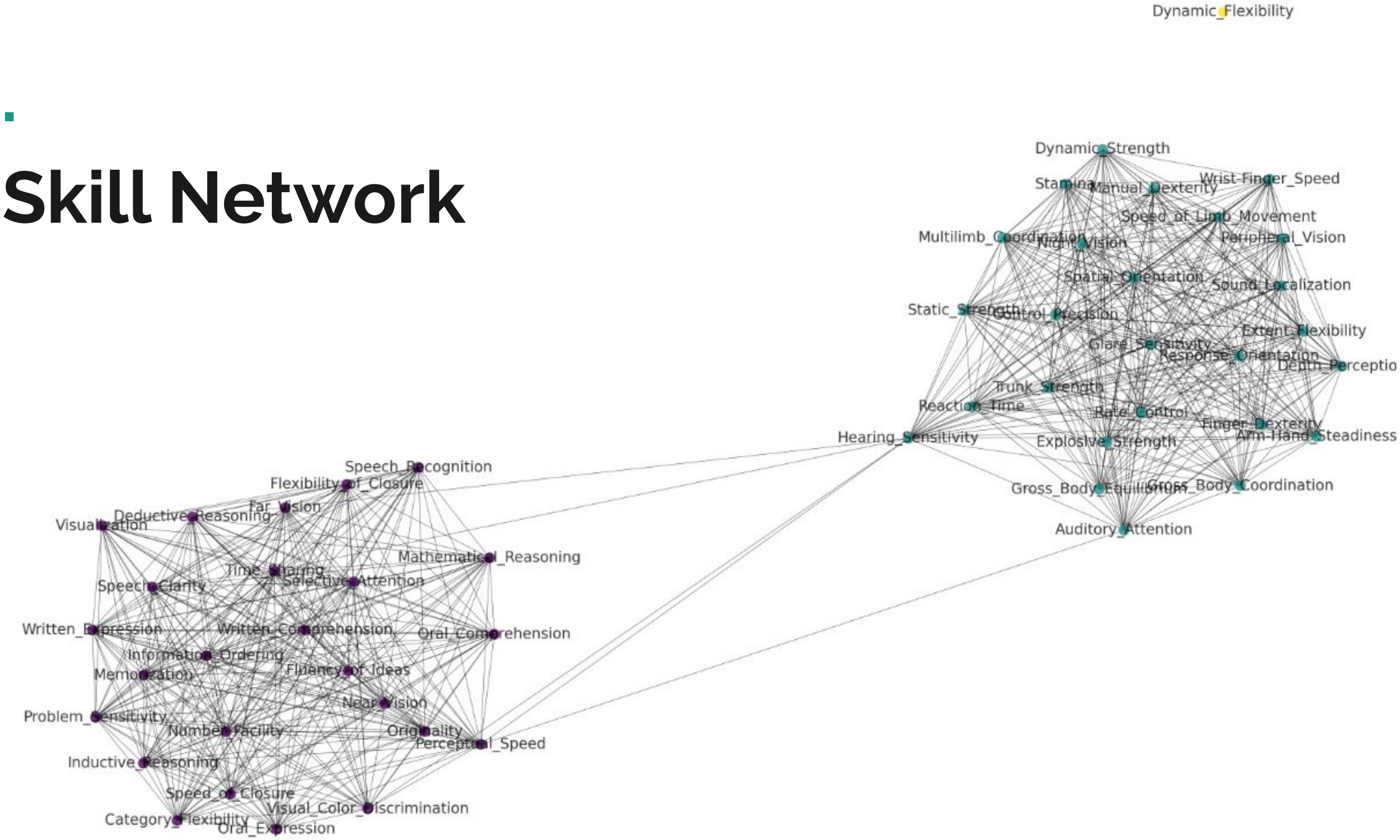
revealed comparative advantage (RCA)

if rca(j, s) > 1, $e(j, s) = 1$; else, $e(j, s) = 0$

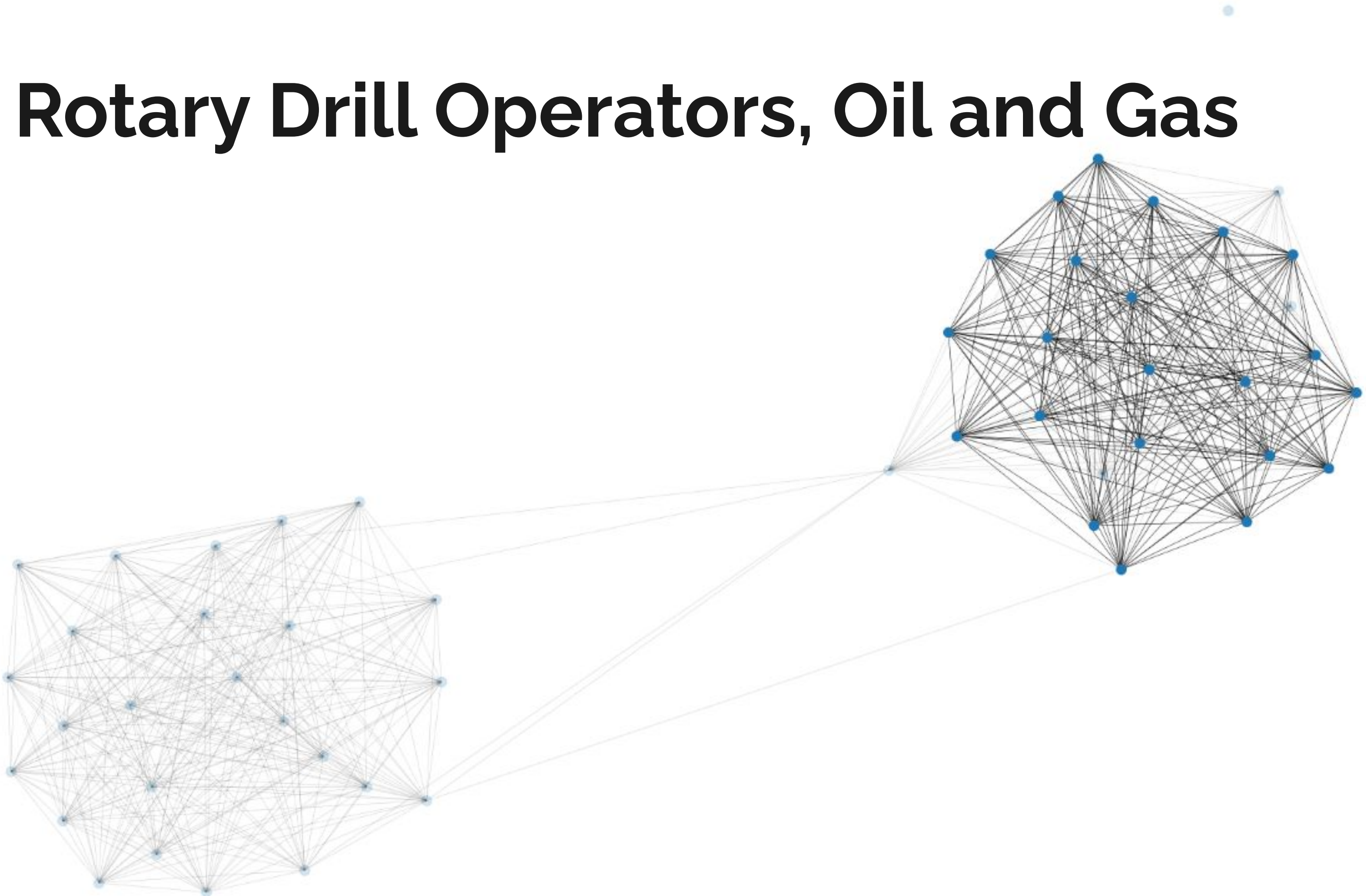
$$\theta(s,s') = \frac{\sum_{j\in J} e(j,s)\cdot e(j,s')}{\max\left(\sum_{j\in J} e(j,s), \sum_{j\in J} e(j,s')\right)}$$

θ(s,s') :The minimum of the conditional probabilities of a pair of skills being effectively used by the same occupation

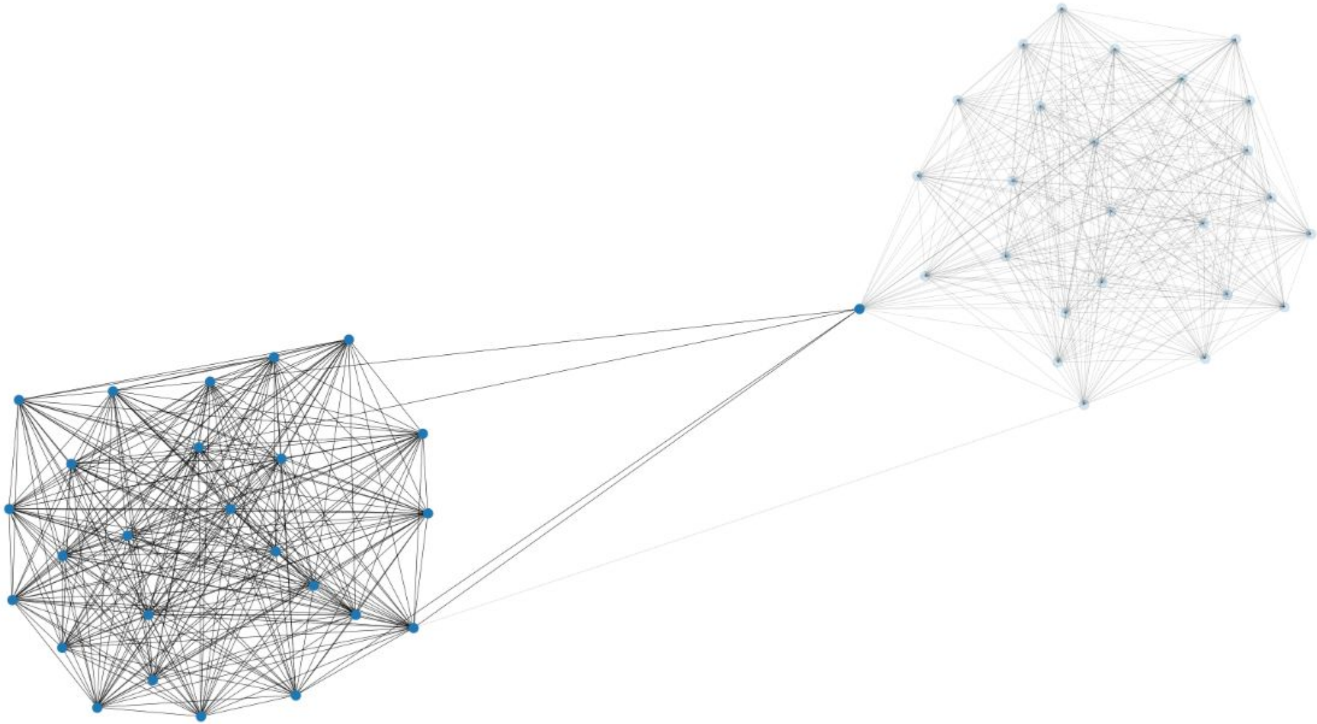
Skill Network



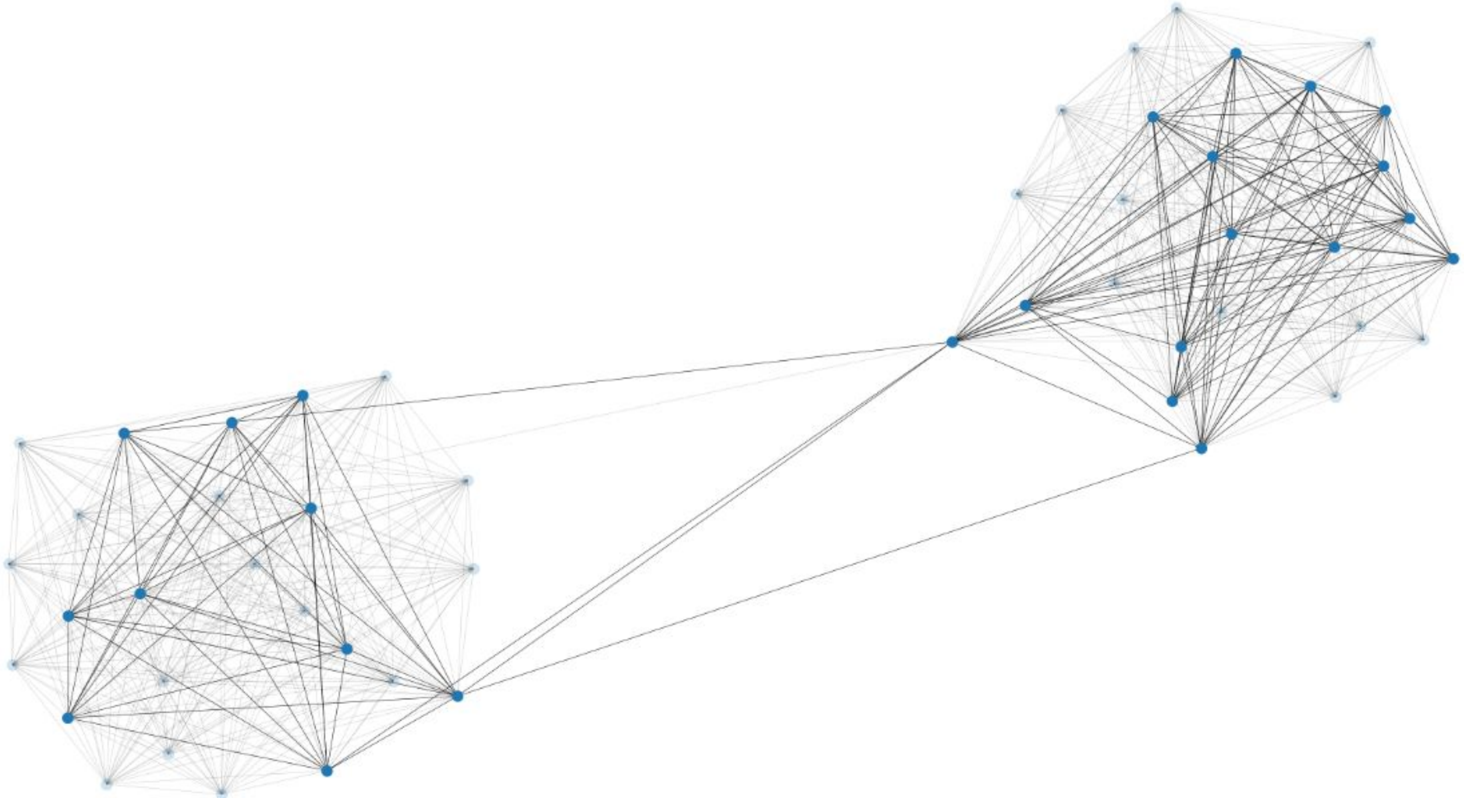
Rotary Drill Operators, Oil and Gas



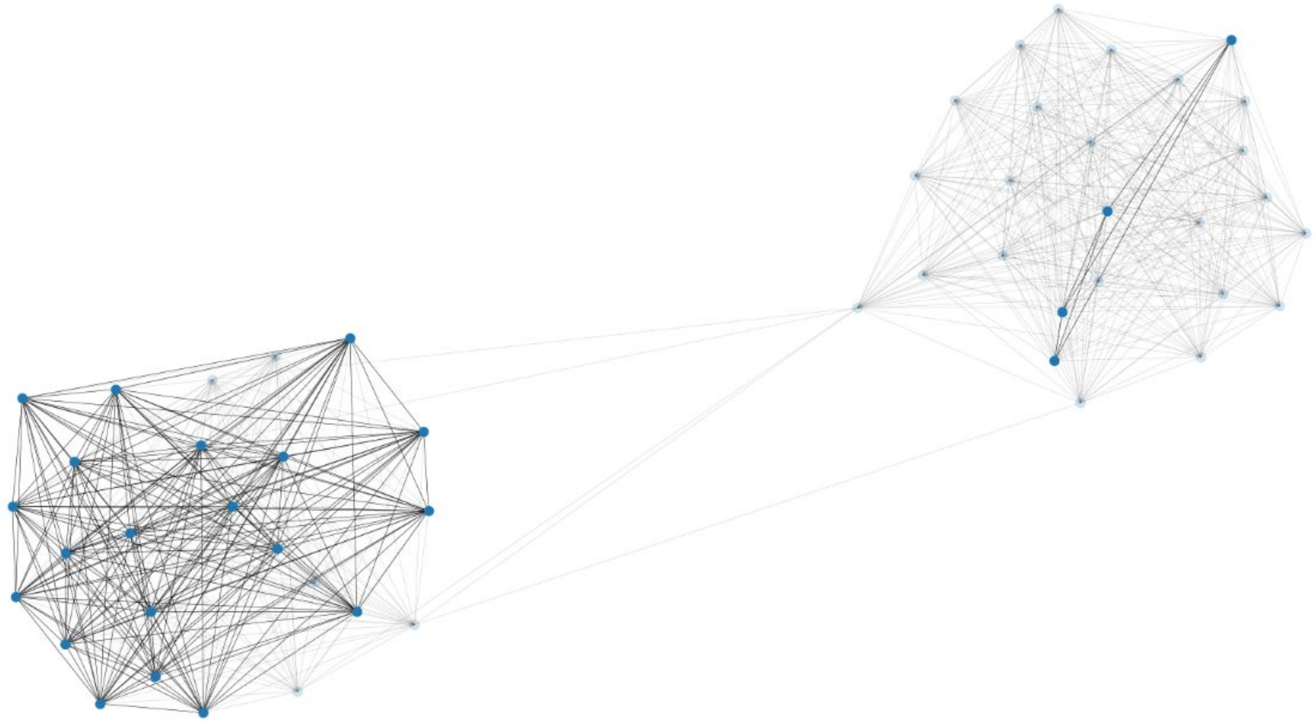
Petroleum Engineers



Nuclear Technicians



Solar Energy Systems Engineer



Method: Building a Occupation Network

$$\mathrm{rca}(j,s) = \frac{\mathrm{onet}(j,s) / \sum_{s'\in S} \mathrm{onet}(j,s')}{\sum_{j'\in J} \mathrm{onet}(j',s) / \sum_{j'\in J, s'\in S} \mathrm{onet}(j',s')}$$

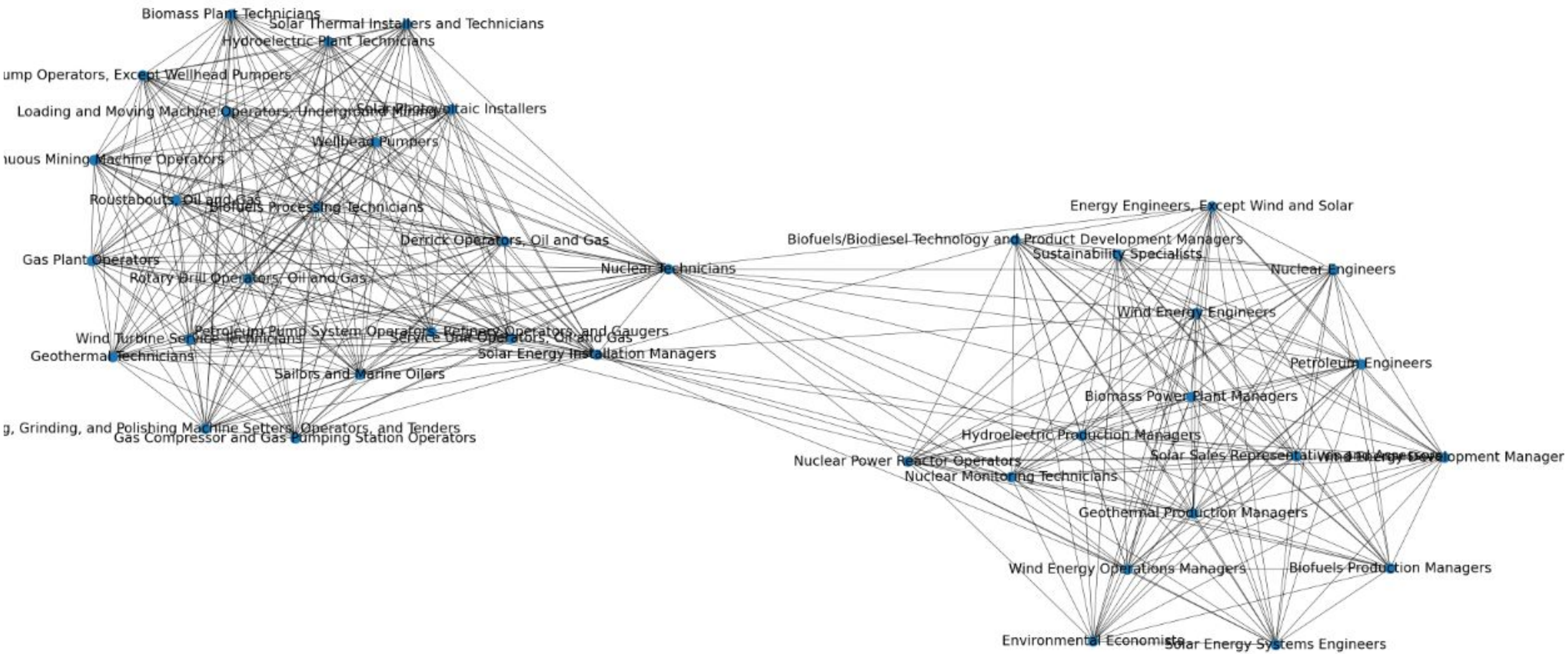
$$\Theta(j,j') = \frac{\sum_{s\in S} e(j,s) \times e(j',s)}{max(\sum_{s\in S} e(j,s), \sum_{s\in S} e(j',s))}$$

revealed comparative advantage (RCA)

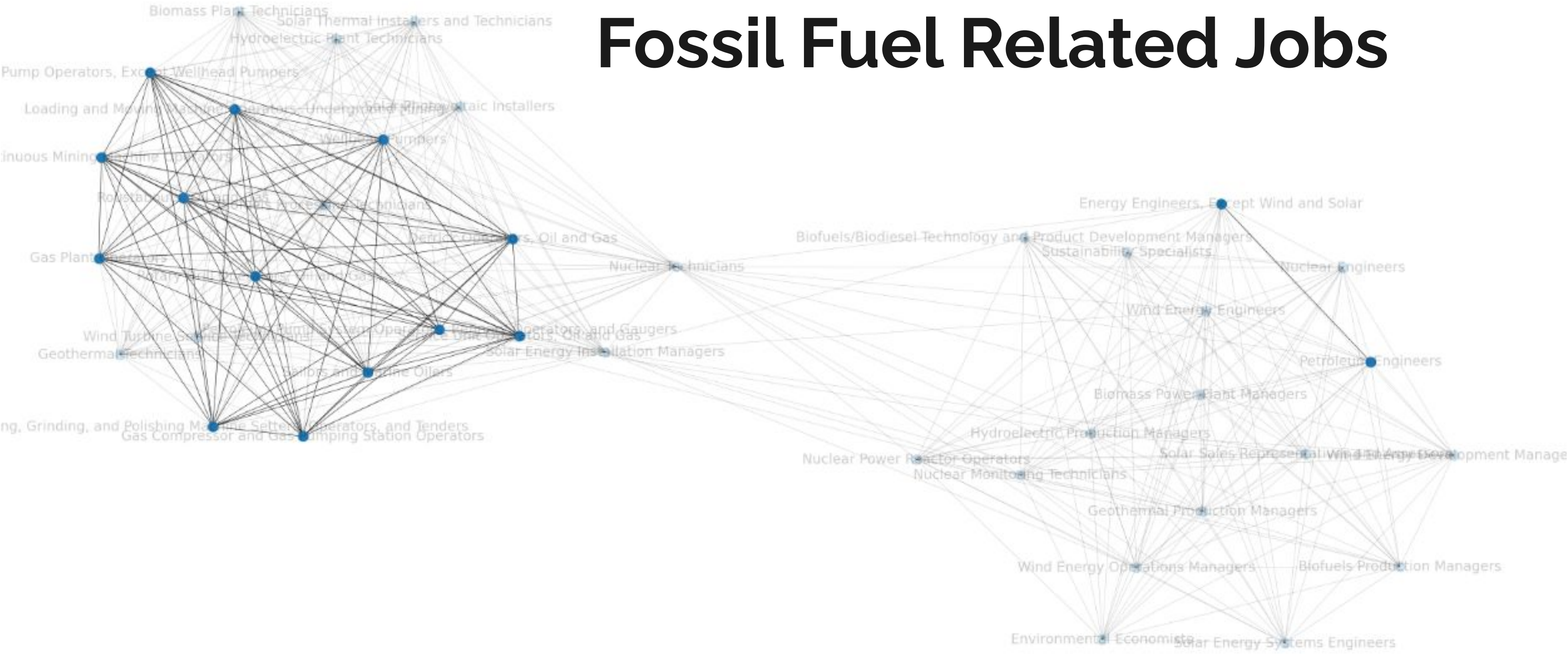
if rca(j, s) > 1, $e(j, s) = 1$; else, $e(j, s) = 0$

θ(j,j') : the minimum probability that a pair of jobs effectively used the same set of skills

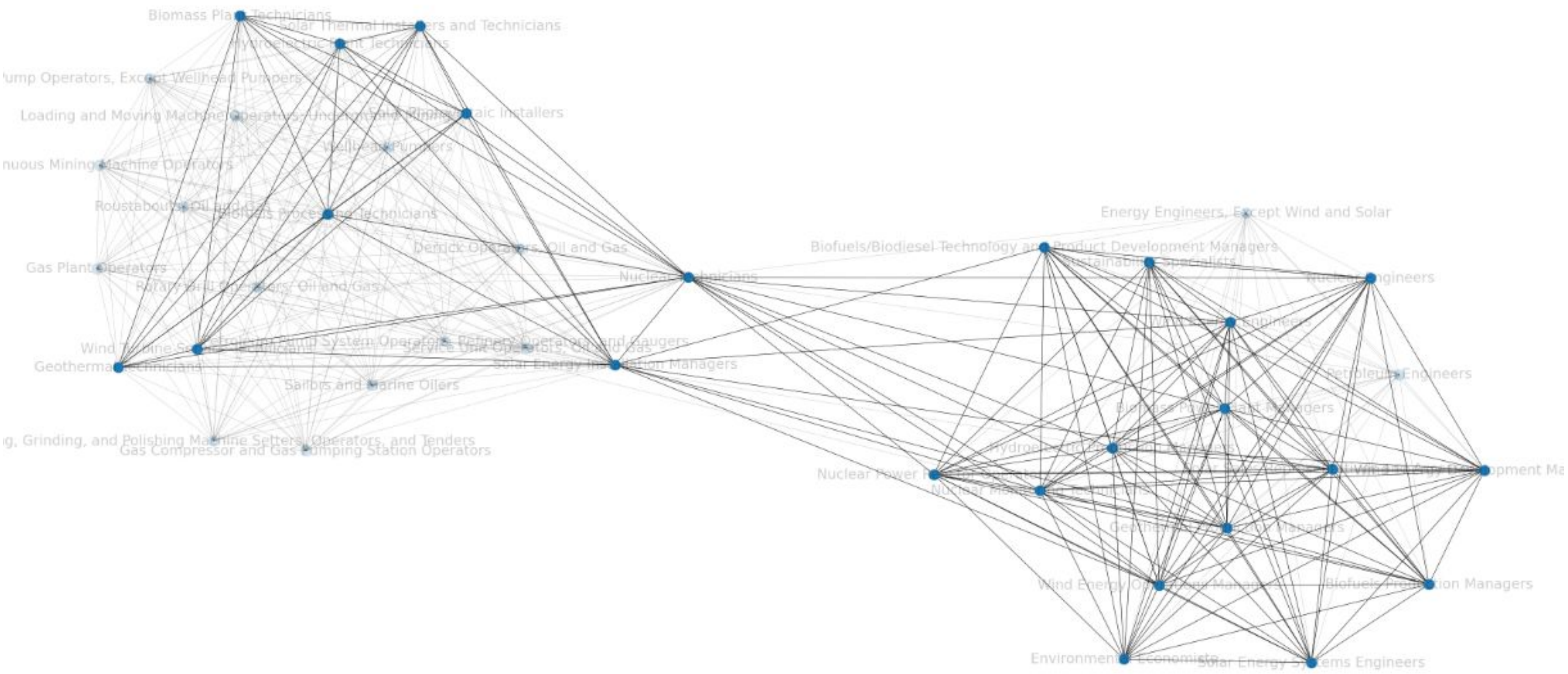
Occupations Network



Fossil Fuel Related Jobs



Green Jobs



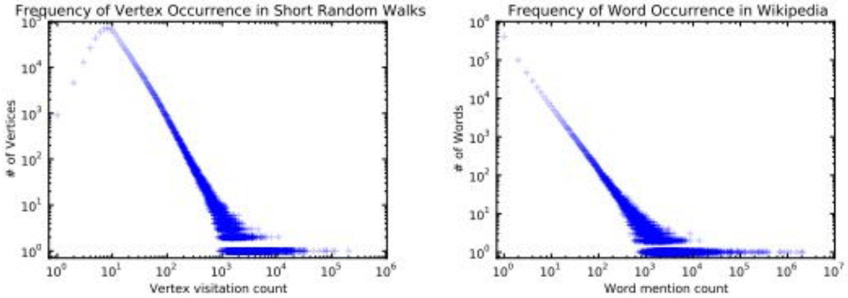
Biomass Plant Technicians
Solar Thermal Installers and Technicians
Hydroelectric Plant Technicians
Pump Operators, Except Wellhead Pumpers
Loading and Moving Machine Operators, Underground Mining
Solar Photovoltaic Installers
Wellhead Pumpers
Continuous Mining Machine Operators
Roustabouts, Oil and Gas
Biofuels Processing Technicians
Derrick Operators, Oil and Gas
Gas Plant Operators
Rotary Drill Operators, Oil and Gas
Nuclear Technicians
Wind Turbine Service Technicians
Petroleum Pump System Operators, Refinery Operators, and Gaugers
Service Unit Operators, Oil and Gas
Geothermal Technicians
Solar Energy Installation Managers
Sailors and Marine Oilers
Gas Compressor and Gas Pumping Station Operators
Energy Engineers, Except Wind and Solar
Biofuels/Biodiesel Technology and Product Development Managers
Sustainability Specialists
Nuclear Engineers
Petroleum Engineers
Biomass Power Plant Managers
Hydroelectric Production Managers
Nuclear Power Reactor Operators
Nuclear Monitoring Technicians
Geothermal Production Managers
Wind Energy Operations Managers
Biofuels Production Managers
Environmental Economists
Solar Energy Systems Engineers

Method: Node2Vec *[2]*

“An algorithmic framework for learning continuous feature representations for nodes in networks” – [2]

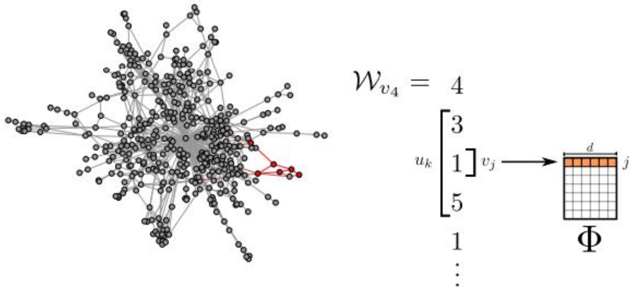
An mixture of both BFS and DFS random walk methods.

Method: Node2Vec *[3]*



(a) YouTube Social Graph (b) Wikipedia Article Text

Figure 2: The power-law distribution of vertices appearing in short random walks (2a) follows a power-law, much like the distribution of words in natural language (2b).

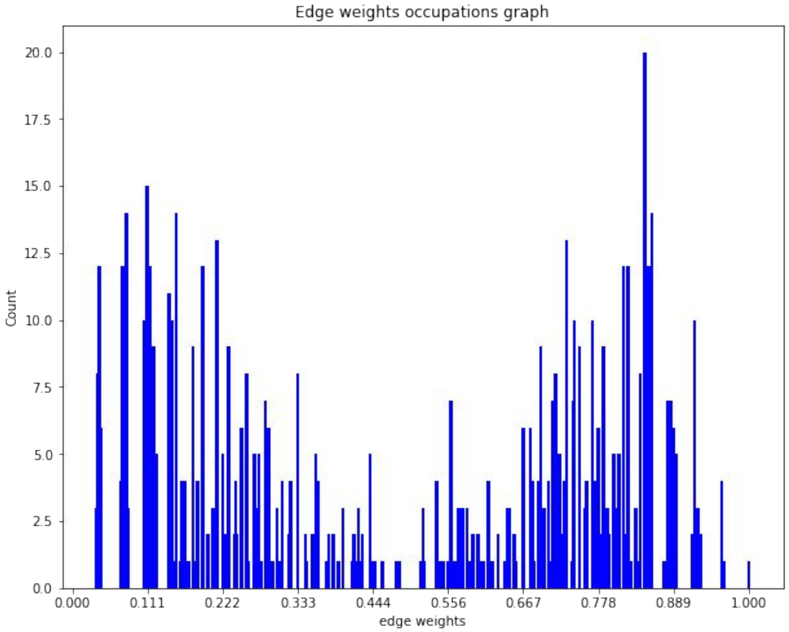


(a) Random walk generation. (b) Representation mapping.

Algorithm 2 SkipGram(Φ, $\mathcal{W}_{v_i}$, w)

for each $v_j \in \mathcal{W}_{v_i}$ **do**
  for each $u_k \in \mathcal{W}_{v_i}[j - w : j + w]$ **do**
    $J(\Phi) = -\log \Pr(u_k \mid \Phi(v_j))$
    $\Phi = \Phi - \alpha * \frac{\partial J}{\partial \Phi}$
  end for
end for

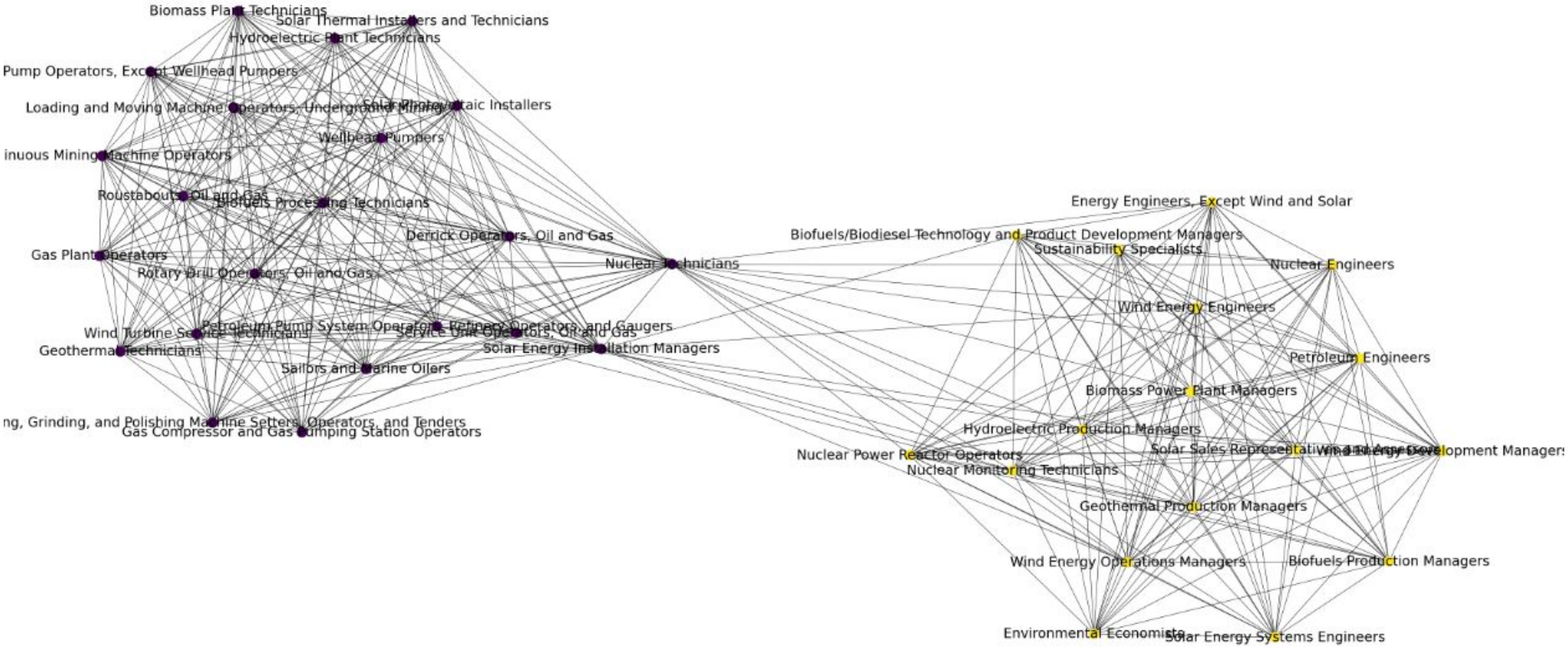
Weighted Node2Vec



Which explains our previous visualization on

occupations graph

Kmeans-clustering(C=2) on node embedding



Assumption: Job transitions will correlate to node embeddings.

If the similarity betweens embeddings of two node A and B
Is greater than the similarity betweens embeddings of two node C and D, the job transitions observed between job A and job B Should be also greater than job C and job D.
Is this true or false?

Assumption: Job transitions will correlate to node embeddings.

If the similarity betweens embeddings of two node A and B
Is greater than the similarity betweens embeddings of two node C and D, the job transitions observed between job A and job B Should be also greater than job C and job D.
Is this true or false?
False! Given our Resume Data, But why?

```
a=weighted_model.wv['Rotary Drill Operators, Oil and Gas']
```

```
[1059] b=weighted_model.wv['Gas Plant Operators']
```

```
score=cos_sim(a,b)
score
```

```
0.8896638
```

Rotary Drill Operators, Oil and Gas	Gas Plant Operators	3.0	0.889664

```
[1046] c=weighted_model.wv['Petroleum Engineers']
```

```
[1047] d=weighted_model.wv['Rotary Drill Operators, Oil and Gas']
```

```
score=cos_sim(c,d)
score
```

```
0.33061278
```

Petroleum Engineers	Rotary Drill Operators, Oil and Gas	141.0	0.330613

```
from scipy.stats.stats import pearsonr
print(pearsonr(similarity_score,count_transition))
```

```
(-0.15792960906773404, 0.24503478242875829)
```

Pearson Correlation Score: -0.1579...
P-value: 0.2450...

Possible explanations:

1. No standardization, did not normalize by job market sizes
2. Granularity issues: Resume Data's code for occupations only have 6-digits(15-1252 e.g.), but our actual occupations have 8-digits for a part of occupation. Already removed..
3. The scope of occupations considered are too small, with only fossil fuel and green jobs, maybe it will be safer to draw that correlation including all occupations

Conclusion and Future Works

1. Fossil fuel occupations heavily rely on sensory-physical skills, while green jobs are more relying on social-cognitive skills. Therefore, the barrier stem from the different skill set required by the two types of jobs.
2. However, required skills wouldn’t be the only perspective to examine the barriers, other factors such as market sizes could also be playing vital roles in the transitions.
3. Future works should be re-examined the correlation between number of transitions and similarities of skill set under a bigger picture, and considering all the occupations while picking only fossil fuel and green jobs as subjects. Normalizing factors such as market sizes of certain industries should also be investigated.

Thanks to everyone!

References

[1] Alabdulkareem, A., Frank, M. R., Sun, L., AlShebli, B., Hidalgo, C., & Rahwan, I. (2018). Unpacking the polarization of workplace skills. *Science advances*, *4*(7), eaao6030.

[2] Grover, Aditya, and Jure Leskovec. "node2vec: Scalable feature learning for networks." *Proceedings of the 22nd ACM SIGKDD international conference on Knowledge discovery and data mining*. 2016.

[3] Perozzi, Bryan, Rami Al-Rfou, and Steven Skiena. "Deepwalk: Online learning of social representations." Proceedings of the 20th ACMhttps://arxiv.org/pdf/1403.6652.pdf SIGKDD international conference on Knowledge discovery and data mining. 2014.