Assessing Patents Based on Their Structural Significance in Patent Citation Network

Max Kuan
National Taiwan University of Science and Technology

Outlines

- **Background**
- Methodology
- Empirical Result
- Conclusion

Need for assessing patents

- Patent analysts often have to deal with thousands or tens of thousands of patents
  - Manually reviewing these patents is obviously a daunting task
- Analysts often seek to identify the most representative ones
  - so that their intended analytic work may be conducted in a more efficient and orderly manner

Representative of what?

- Influence/impact
  - E.g., citation count, generality
- Technology diversity
  - E.g., patent scope, originality
- Scientific linkage
  - E.g., # of academic articles cited
- # of claims
- …

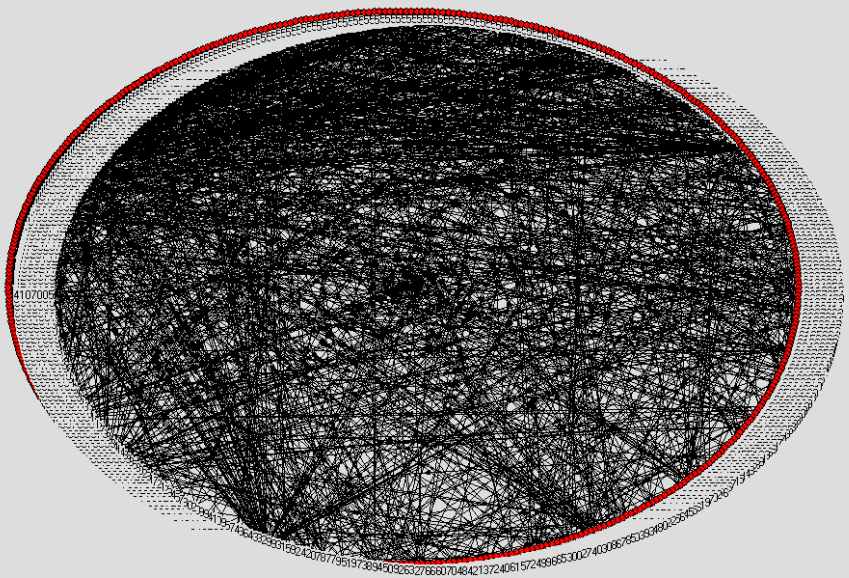
In this work, we intend to identify patents playing more representative roles in disseminating knowledge

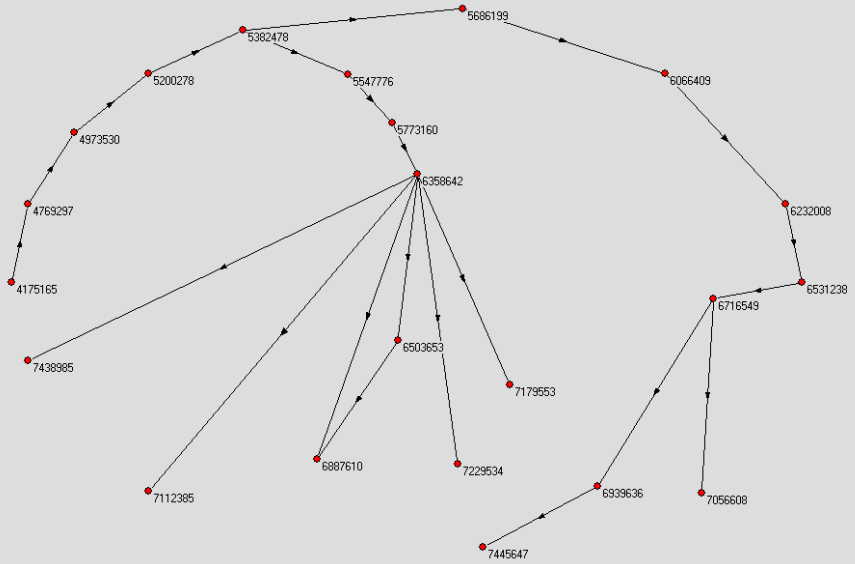
Citation≈Knowledge flow

Main path analysis

- A methodology borrowed from social network analysis
- Originally aimed to determine the major development trajectory in a scientific field
- Has been applied to patents in detecting technological changes, and finding trajectories of technological development

An Example

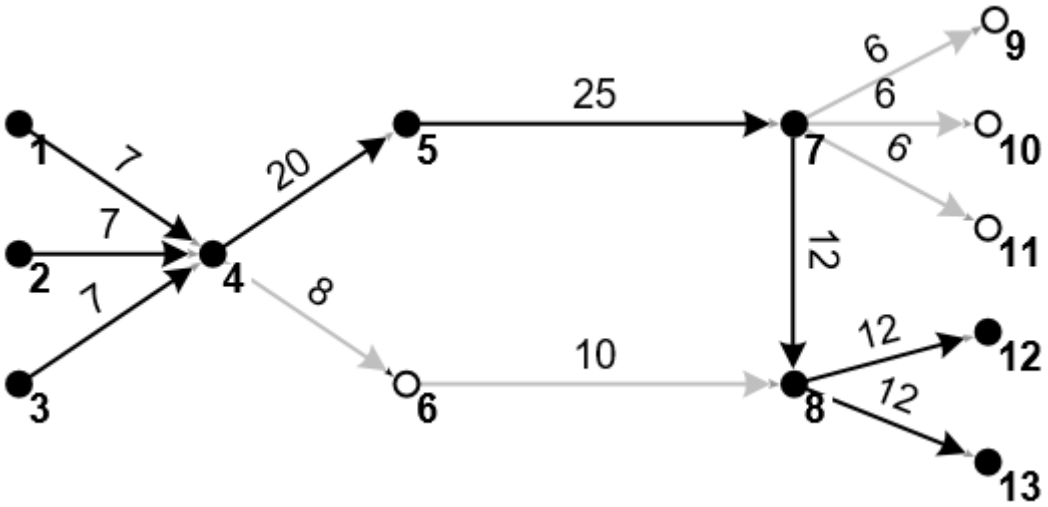




Analysis procedure

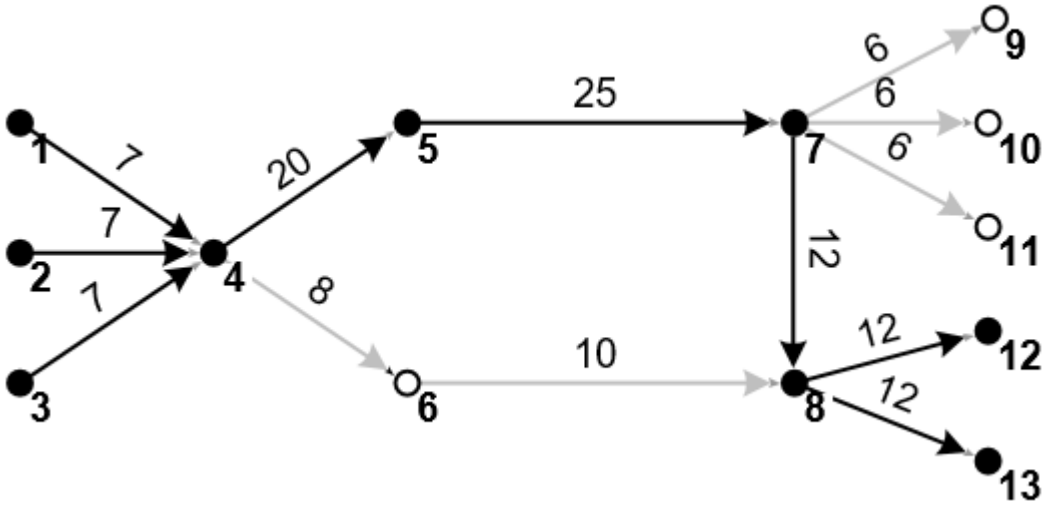
- (1) a citation network is constructed
- (2) a weight for each arc is assigned based on its traversal count
  - There are various algorithms
- (3) a series of connected arcs across the network is determined as a representative trajectory
  - There are various methods

Assigning weight



- Search Path Link Count (SPLC) algorithm
  - Weight of 5→7=25, as each of five preceding nodes (1 to 5) traverses the arc five times to reach the sink nodes (9 to 11)
  - Weight of 8→12=12, as four of its preceding nodes (1 to 4) traverses the arc twice (one following the arc 6→8 and the other following the arc 5→7), and the other four preceding nodes (5 to 8) traverse the arc once, to reach the sink node 12

Determining main path



- Global Search method
  - Selects one of the paths from source to sink nodes having the greatest combined weight
  - The main path in the network involves the source nodes 1 to 3, the intermediate nodes 4, 5, 7, 8, and the sink nodes 12, 13, and the combined weight along the path is 76 (=7+20+25+12+12)

Outlines

- Background
- ➢ **Methodology**
- Empirical Result
- Conclusion

Arc's structural connectivity

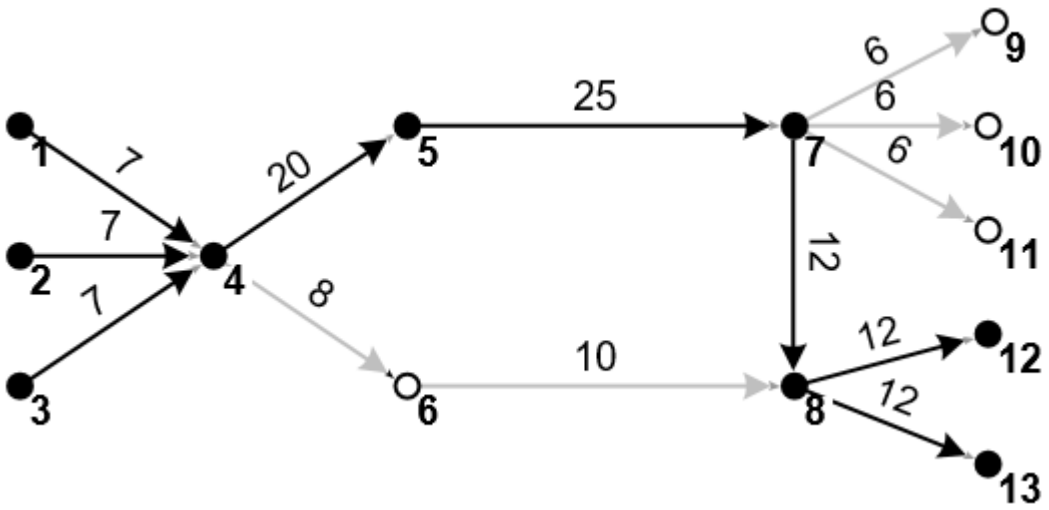
- An arc would have a greater weight if the arc has greater structural connectivity
  - meaning that it can be reached from more preceding nodes and/or it may lead to more succeeding nodes
- The arc may be deemed as more important in disseminating knowledge

Arc weight to node weight

- Each node is assigned a *pivotal* value related to the weights of its incident and/or outgoing arcs
  - A node's *pivotal* value reflects its significance in accordance with its structural position within the citation network.
    - A node would have a greater pivotal value if it is traversed more frequently from its proceeding nodes to succeeding nodes
- This pivotal value then may be considered as a measure regarding how *pivotal* a patent is in terms of its role in disseminating knowledge or evolving technology

Arc weight to node weight

	1	2	3	4	5	6	7	8	9	10	11	12	13
Incident	0	0	0	21	20	8	25	22	6	6	6	12	12
Outgoing	7	7	7	28	25	10	30	24	0	0	0	0	0
Combined	7	7	7	49	45	18	55	46	6	6	6	12	12



Pivotal value

- This study chooses to use the total weights of outgoing arcs
  - Combined Incident weight < Combined outgoing weight
    - So, combined incident+outgoing weight has duplicated information content
    - So, just use combined incident or outgoing weight
  - The sink nodes (i.e., nodes 9-13) are automatically filtered
    - This is convenient and reasonable as the corresponding patents are not yet cited and their importance in terms of knowledge dissemination or technology evolution is yet to determine

Benefit of pivotal value

- Allow analysts to differentiate the patents on the main path
- May identify some snub patents that are left out by the main path

Outlines

- Background
- Methodology
- ➢ **Empirical Result**
- Conclusion

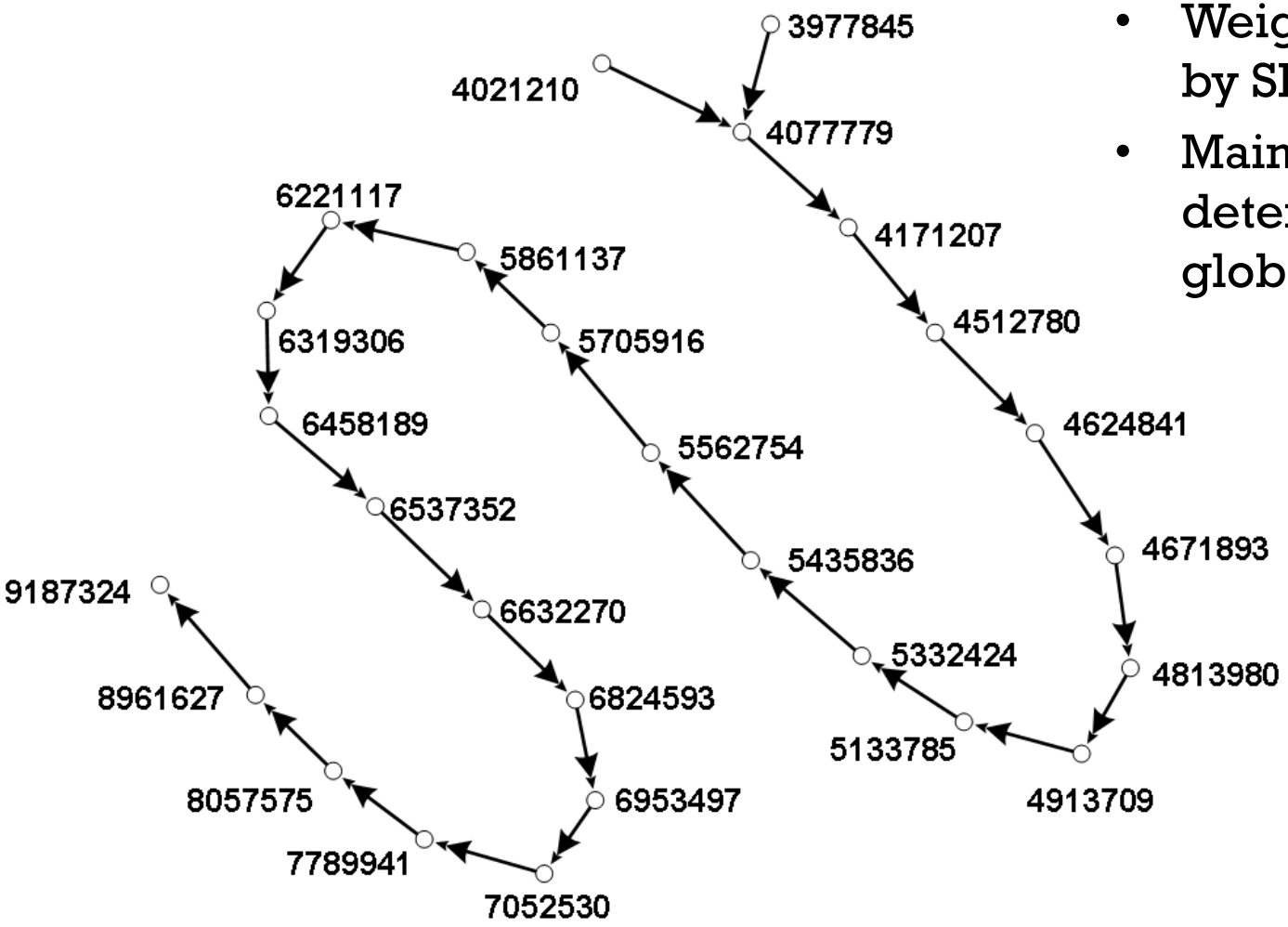
Question

- (1) Are the patents having the greatest pivotal values (or, *pivotal* patents) always the ones present on the main path
  - Are there some pivotal patents actually missed by the MPA?
- (2) Are the pivotal patents also the ones with the largest citation counts?

Empirical data

- The field of carbon dioxide capture, storage, recovery, delivery, and regeneration
- A total of 34,083 US utility patents issued between 1976/1/1 and 2017/3/31 by the United States Patent and Trademark Office database
  - there are 29,838 patents citing or being cited by at least one other patent
  - Total 155,076 citations
  - Each patent is cited 4.55 times on the average

Main path



- Weight assigned by SPLC
- Main path determined by global search

Global Main Path	Top Pivotalness	Top Citations
3977845	3977845	3986849
4021210	3986849	4094777
4077779	4021210	4101631
4171207	4171207	4310440
4512780	4466946	4440871
4624841	4578089	4522894
4671893	4701187	4567029
4813980	4705541	4711645
4913709	4737167	5057483
5133785	4914218	5354547
5332424	4952219	5409522
5435836	5073356	5458857
5562754	5096470	5505766
5705916	5133785	5597771
5861137	5234472	5858314
6221117	5248322	5861137
6319306	5294246	5938800
6458189	5326550	5997594
6537352	5415682	6063161
6632270	5417742	6077620
6824593	5435836	6221117
6953497	5503658	6333016
7052530	5571309	6719828
7789941	5669958	6878358
8057575	5779768	6890497
8961627	6231644	6953494
9187324	6610124	7132090

Local Main Path	Top Pivotalness	Top Citations
3977845	3977845	3986849
4021210	3986849	4094777
4077779	4021210	4101631
4171207	4171207	4310440
4512780	4466946	4440871
4705541	4578089	4474896
4790858	4701187	4522894
4913709	4705541	4567029
5133785	4737167	4711645
5332424	4914218	5057483
5435836	4952219	5354547
5562754	5073356	5409522
5705916	5096470	5458857
5861137	5133785	5505766
6221117	5234472	5597771
6319306	5248322	5858314
6458189	5294246	5861137
6537352	5326550	5938800
6632270	5415682	5997594
6824593	5417742	6063161
6953497	5435836	6077620
7052530	5503658	6221117
7789941	5562754	6333016
8057575	5571309	6719828
8961627	5669958	6878358
9102529	5779768	6890497
9187324	6231644	6953494
9556025	6610124	7132090

Patents appearing on the main path and also having the top pivotal values are shown with grey background

Those both on the main path and having top citation counts are in reversed color.

Observations

- Few patents on the main path are identified using the pivotal values and citation counts
  - Each has its merit and captures a separate aspect of a patent's quality

Observations

- As to citation count, it reflects a patent's "local" property, just between the patent and its immediate successors
  - They may lose their shine when they are evaluated in a broad context against other patents
- Patents on the main path are not necessarily the ones having the greatest citation counts or traversal counts
  - It is their "global" structural connectivity that matters

Observations

- Pivotal value seems to be a measure sitting in between
  - It is more "global" than the citation count
  - Yet it is not as "global" as the main path because it is still limited to those directly and indirectly preceding and succeeding it
    - Therefore the local search method's determining the main path is more consistent with the pivotal values
- The pivotalness value seems to favor the older patents
  - They are the ones accumulating a greater number of traversals

Outlines

- Background
- Methodology
- Empirical Result
- **Conclusion**

Summary

- Patents' pivotalness may be conveniently obtained as a side product to the traditional MPA using the readily available software Pajek
- Analysts may use the patents' pivotal values to identify a rather different set of representative patents
  - supplements to those identified using main path or citation count

Extension

- A more rigorous study on the correlation between patents' pivotal values and citation counts
- Incorporate the degree of similarity between the cited and the citing into the calculation of the pivotal value.
  - It is well known that not all citations are equal.
  - By considering the degree of similarity between the cited and citing patents when assigning weights to the citations, the pivotal value may be more accurate

Thank You