NANYANG TECHNOLOGICAL UNIVERSITY

SINGAPORE

CZ4071
Assignment 1
Names/Mat.No

Dongwoo Han/U1722048E
Deng Ziyang/U1923770A
Min Thant Ko/U1821227E
Ng Teng Ann/U1821521F

Table of Contents

Project Description

SCSE of NTU has been growing positively both on reputation and number of publications. In order to understand the evolution, collaboration, reputation of SCSE as well as provide further insights, we constructed a co-authorship network where two nodes(authors) are connected if they have at least one publication together.

In this project, we take a list of SCSE faculty members (Faculty.xlsx) and top venues (Top.xlsx) as input, scrape the publication data from DBLP computer science bibliography website. We then constructed the co-authorship network using many python library such as Networkx, Matplotlib and PandasGUI.

GUI Initialization Steps

1. Open command prompt inside the directory that contains the project
2. Run “pip install -r requirements.txt”
    a. Require rights to install Python Libraries on the computer
3. Run project.py
4. Click “Load File”
5. Direct to “Faculty.xlsx”
6. Click “Load File” again
7. Direct to “Top.xlsx”
8. Click “Initialize Resources” and wait for completion
    a. May take 5~20 minutes
9. Click “Go to Main”

Note: Some features will take some time to load. Please wait patiently when the application freezes after click of a button.

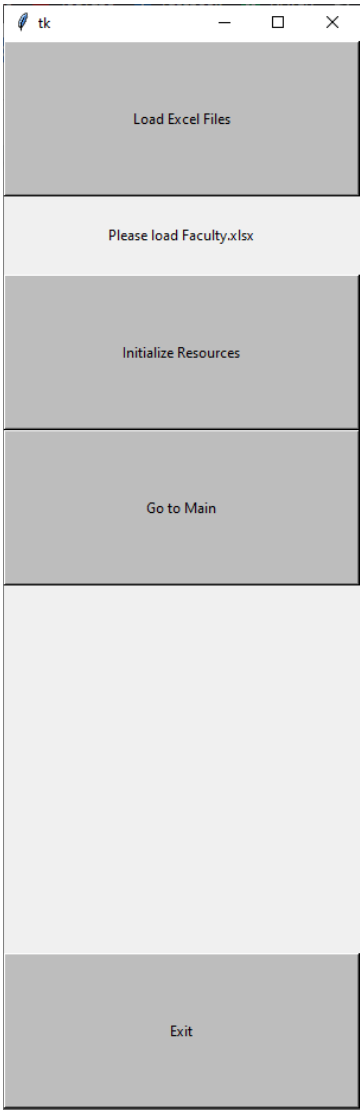


Figure 1: GUI of Initialization

Addressing of the Optional Challenge

The namesake problem is not an issue in our program. As long as the Faculty’s DBLP address is correctly inputted, our program is able to mine all publications and coauthors. Through using the Request library in Python, the program retrieves the returned link of the DBLP address containing the pid of that individual, for example https://dblp.org/pid/83/3179.html. Through using the returned link but not the input DBLP address, it makes sure that the link is valid. Also, there are some inputs that contain the name, rather than the pid, for example, https://dblp.org/pers/o/Oh:Hong_Lye.html. The returned link will be converted to https://dblp.org/pid/224/9431.html, which contains a pid that is unique to this person. The program also replaces all domains to dblp.org as some input have legacy links. With the swap of html to xml, it is able to directly download the xml of individual professors and store them in a folder called faculty_xml. This will contain all information of each professor’s publication and coauthors without having any namesake problems.

In this program, it takes all the links in the Faculty.xlsx and create all the relevant XML files. The XML files are basically the resources to compute different relationships for this project.

Note: The faculty_xml folder will appear after the initialization step in the GUI.

```
r = requests.get(link)
redirect_link = r.url
xml_link = redirect_link.replace("html","xml")
if("dblp.uni-trier.de" in xml_link):
    xml_new_link = xml_link.replace("dblp.uni-trier.de","dblp.org")
else:
    xml_new_link = xml_link
f.write(xml_new_link+'\n')
print(xml_new_link)
```

Figure 2: Code Snippet of XML preparation

Network Property GUI

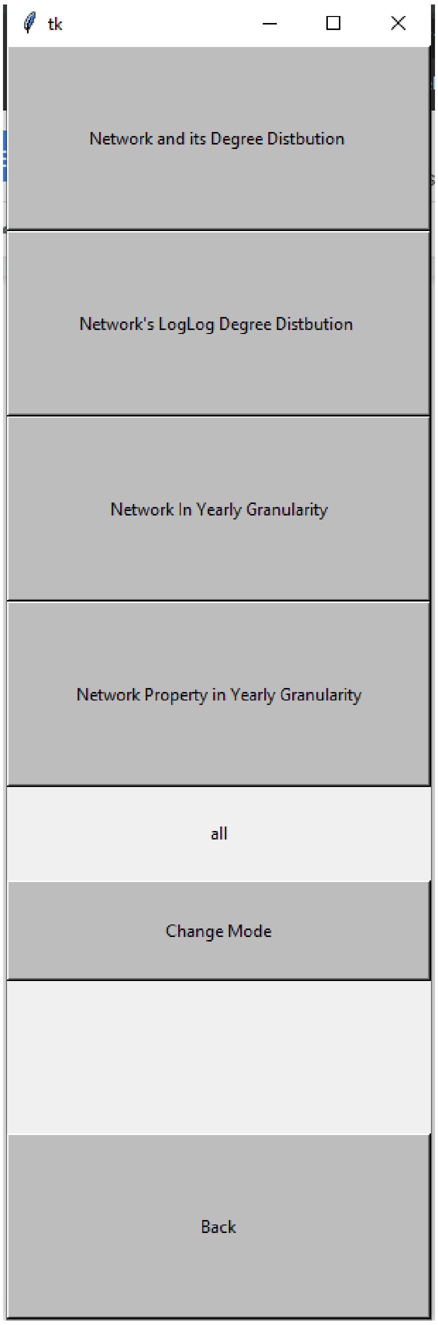


Figure 3: GUI of Network property

This section of the GUI provides an overview of network property of NTU SCSE network.

The 'Network and its Degree Distribution' button displays a bar graph of the network's degree distribution, and top-right is a plot of the network.

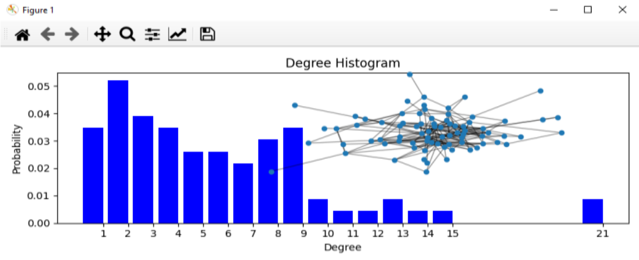


Figure 4: GUI of Degree Distrbution

The 'Network's LogLog Degree Distribution' button displays a scatter point graph of the network's degree distribution. Different from the first button, the graph's x and y axis are replaced with their natural logarithm.

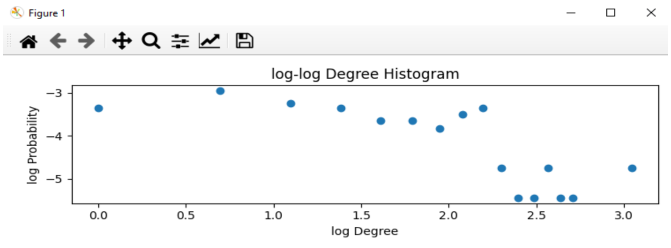


Figure 5: GUI of LogLog Degree Distribution

The 'Network in Yearly Granularity' button displays a collection of plots of the network from 2000 to 2021. This may take some time to complete.

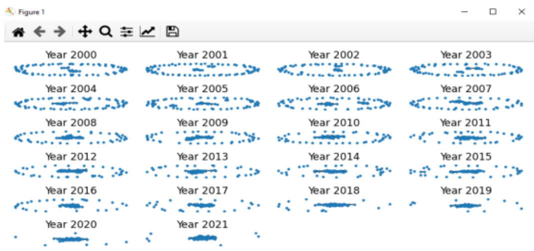


Figure 6: GUI of Yearly Network

The 'Network Property in Yearly Granularity' button displays a pandas DataFrame in PandasGUI with network property from year 2021 to year 2000. This process may take some time to complete. Details on how to use PandasGUI is below.

Since the faculty network before 2021 is not necessarily connected, we provided alternate solutions only considering connected components or giant components. Below the four buttons the text indicates the current mode, and by clicking "Change Mode" the mode will be switched between "all", "connected" and "giant". "all" mode will keep tracking on all current faculty members; "connected" mode will filter out unconnected nodes; "giant" mode will only leave the giant connected component.

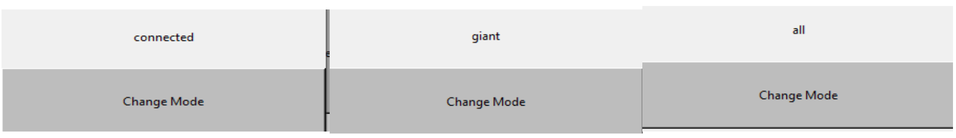


Figure 7: GUI of Mode Options

Collaboration GUI

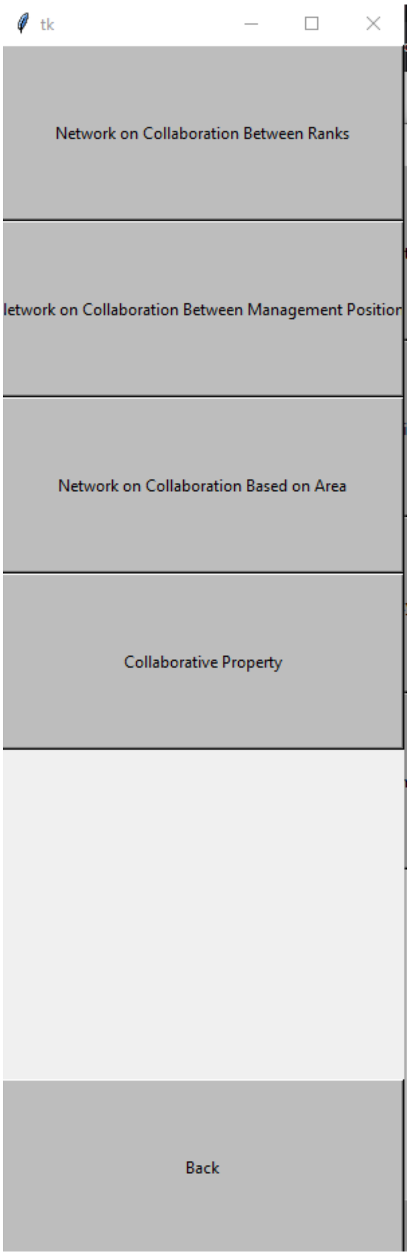


Figure 8: GUI of Collaboration GUI

This section of the GUI focuses on different collaboration undertaken within NTU SCSE faculty. Each button will respectively display the network using matplotlib regarding the below properties:

1. Rank Collaboration
2. Management Collaboration
3. Betweenness Collaboration

Collaborative Property GUI

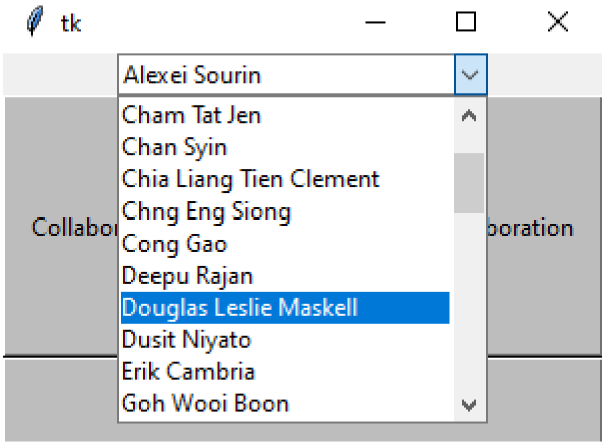


Figure 9: GUI of Drop down list of Faculty

This GUI contains a dropdown list at which the user can choose an individual from the inputted Faculty.xlsx. With the selected individual, it is possible to see their trend from 2000 until now. The options are below:

1. Collaborative Property on Number of Collaboration
2. Collaborative Property on Different Ranks
3. Collaborative Property among Management Positions
4. Collaborative Property on Area of Interest

Excellency GUI

This section of the GUI shows the centralities of the NTU SCSE network and its relationship with the 'Excellent' node.

The first three buttons display different types of centralities: Degree, Eigenvector and Betweenness and its heatmap is a network using matplotlib. As the colour gets closer to yellow, it indicates that it has a high centrality.

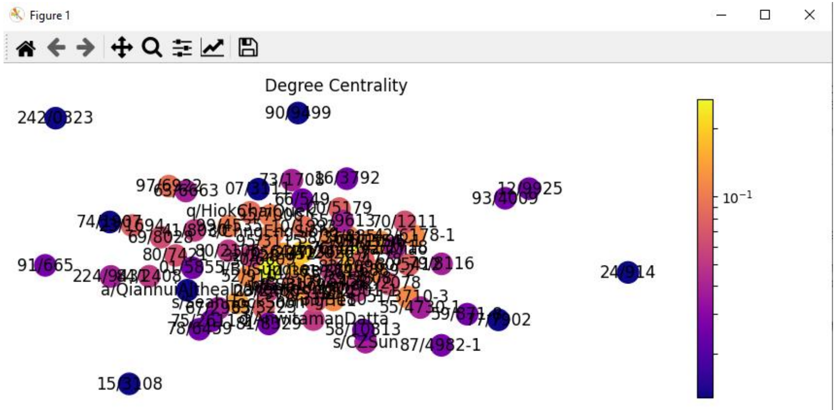


Figure 11: GUI of Ceentrality heatmap

The 'View DataFrame' button displays a Pandas DataFrame of centrality and 'Excellency' of each individual in NTU SCSE. It uses an external library called PandasGUI to perform this. It is interactive and can be filtered to the users' needs. Refer to below for more information on how to use PandasGUI.

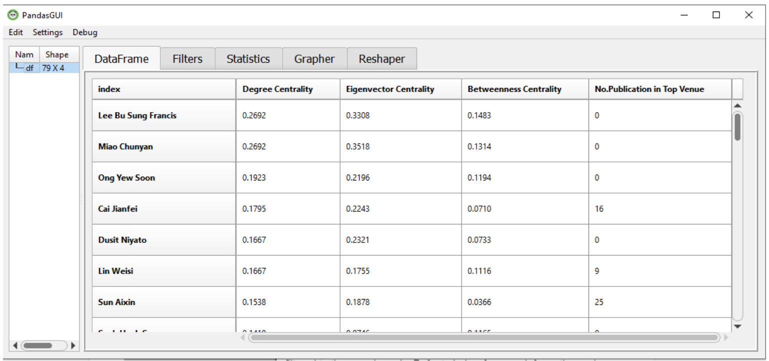


Figure 12: PandasGUI of Centrality and Excellancy

The 'Scatter Plot' will display a scatter plot using matplotlib to show the relationship between centrality and 'Excellency'.

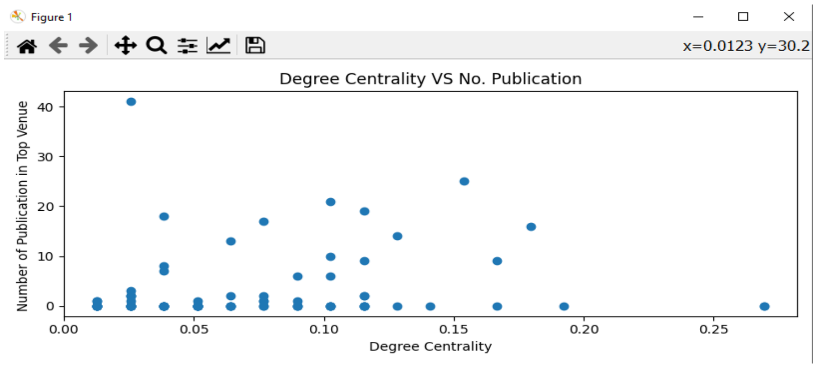


Figure 13:Scatter Plot GUI

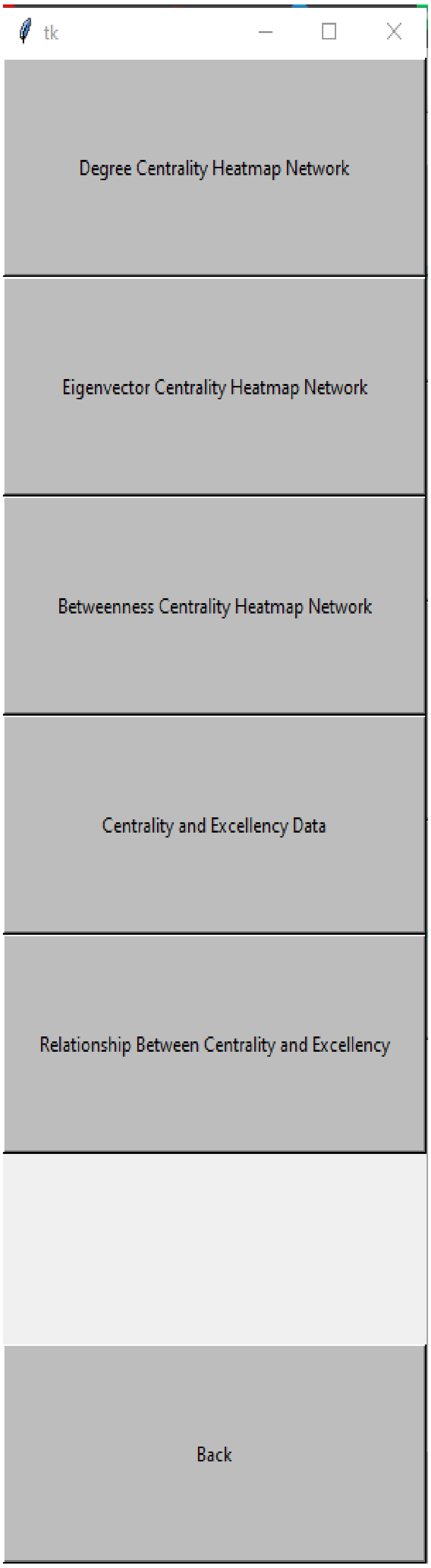


Figure 10: GUI of Excellency

Hiring New Member GUI

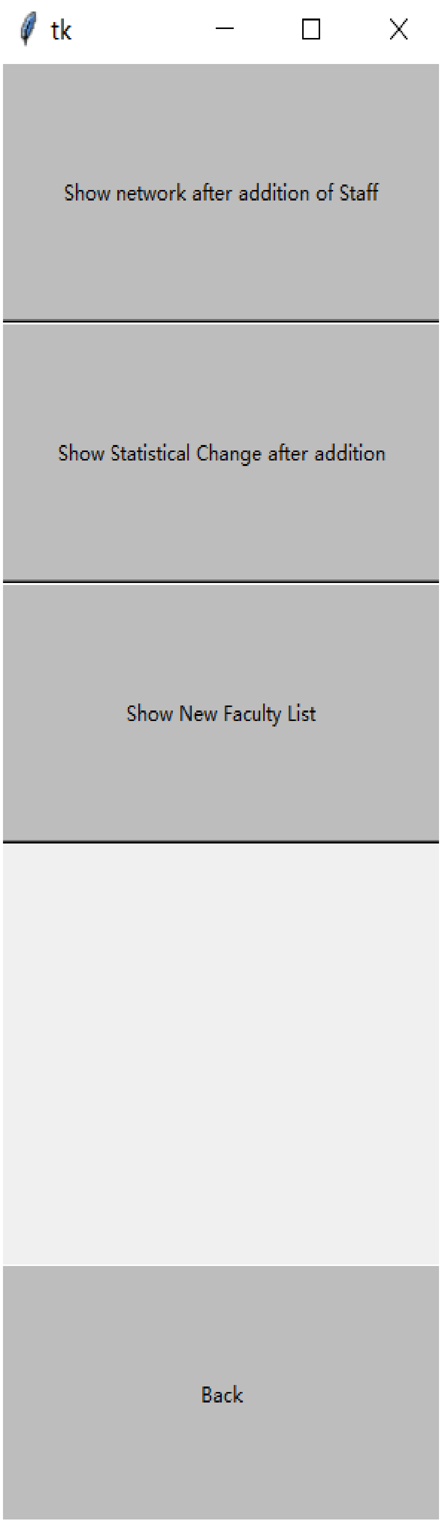


Figure 14: GUI of Recommendation

This section of the GUI is based on the assumption SCSE is about to hire a total of 1631 new members.

The “Show network after addition of Staff” displays the network graph after hiring the new members. In the graph each dot represents a node. In this graph, collaborations between current members are excluded.

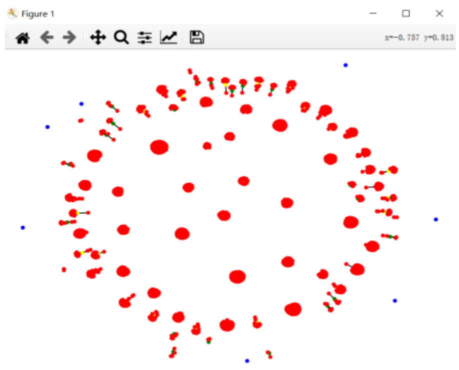
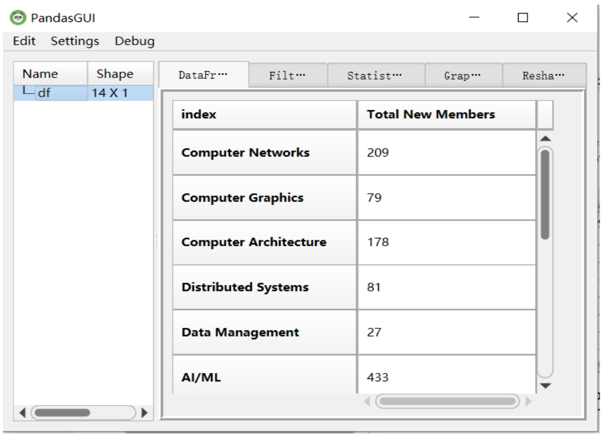


Figure 15: GUI of Network after adding staff

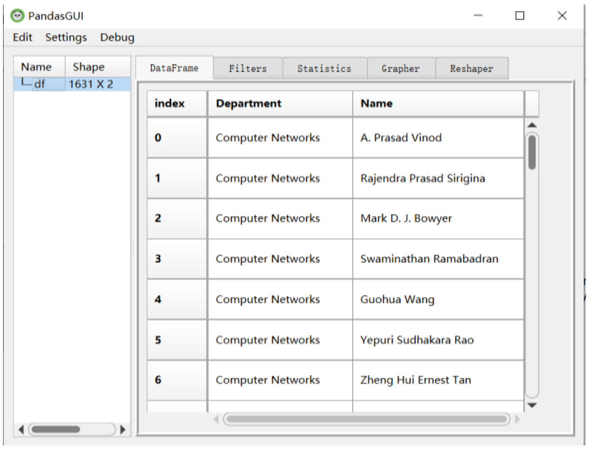
The “Show Statistical Change after addition” button displays a Pandas DataFrame showing the number of new members in each area.



index	Total New Members
Computer Networks	209
Computer Graphics	79
Computer Architecture	178
Distributed Systems	81
Data Management	27
AI/ML	433

Figure 16: PandasGUI of Statistical Change

The “Show New Faculty List” button displays a Pandas DataFrame showing a list of possible new members, and their area of study.

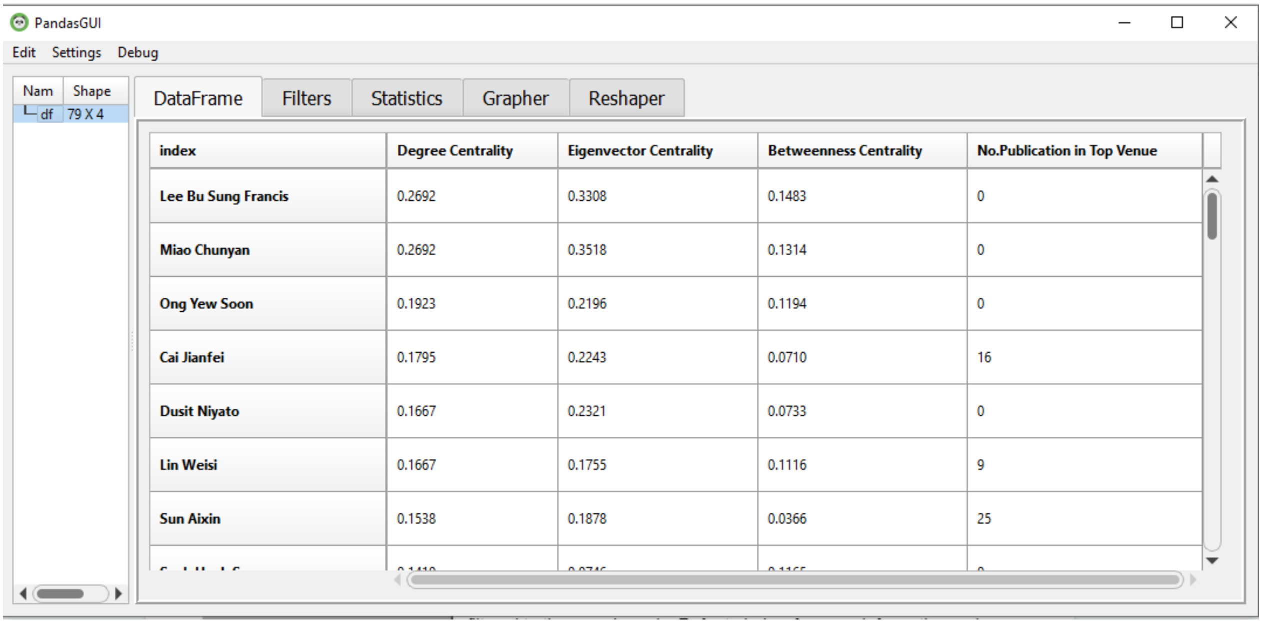


index	Department	Name
0	Computer Networks	A. Prasad Vinod
1	Computer Networks	Rajendra Prasad Sirigina
2	Computer Networks	Mark D. J. Bowyer
3	Computer Networks	Swaminathan Ramabadran
4	Computer Networks	Guohua Wang
5	Computer Networks	Yepuri Sudhakara Rao
6	Computer Networks	Zheng Hui Ernest Tan

Figure 17: PandasGUI of new Faculty

PandasGUI

PandasGUI is a library developed that easily displays Pandas DataFrame in an interactive way. It supports ordering, filtering and local editing of the DataFrame. The DataFrame can be easily exported as a csv with the click of the Export button under Edit. The degree of freedom that PandasGUI gives with the DataFrame is primarily why it was chosen to view our data.



index	Degree Centrality	Eigenvector Centrality	Betweenness Centrality	No.Publication in Top Venue
Lee Bu Sung Francis	0.2692	0.3308	0.1483	0
Miao Chunyan	0.2692	0.3518	0.1314	0
Ong Yew Soon	0.1923	0.2196	0.1194	0
Cai Jianfei	0.1795	0.2243	0.0710	16
Dusit Niyato	0.1667	0.2321	0.0733	0
Lin Weisi	0.1667	0.1755	0.1116	9
Sun Aixin	0.1538	0.1878	0.0366	25

Figure 18: PandasGUI Sample

The Grapher column allows a very interactive way to explore trends among the data. Each Column can be selected to create a graph from the selections in the below image.

For example, the Line graph was extensively used in our project to get a line graph on evolution of network properties in yearly granularity.

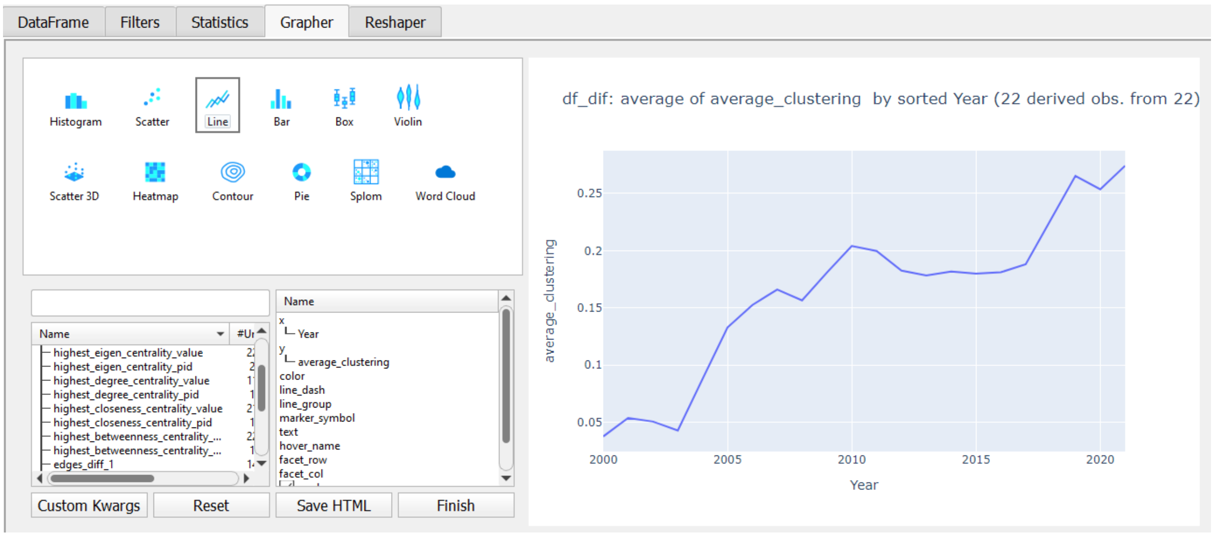


Figure 19: PandasGUI Grapher Sample

For more details on how to use PandasGUI, please refer their official website: https://github.com/adamerose/pandasgui

Analysis of SCSE Faculty Network

Network Property

Staff Size	Diameter	Average Distance	Average Clustering Coefficient	Average Degree
79	6	2.7643	0.2951	2.8861

Analysis

The Average Distance is larger than log(79)=1.897, which indicates that this network does not represent the small-world property.

The Average Clustering Coefficient is close to $\frac{1}{3}$. This indicates that there is a significant amount of modularity in this network. This has likely occurred due to staff from same areas of interest collaborating often or staff who collaborated in the past are likely to collaborate again. Detailed analysis is shown in the Collaboration section.

The Average Degree matches that of critical point between Scale-Free Network and Random Network where $< k > = 2.886 \approx \frac{lnN}{lnlnN} = 2.963$. Hence, it can be seen that the network may possess some properties from both.

Random Network

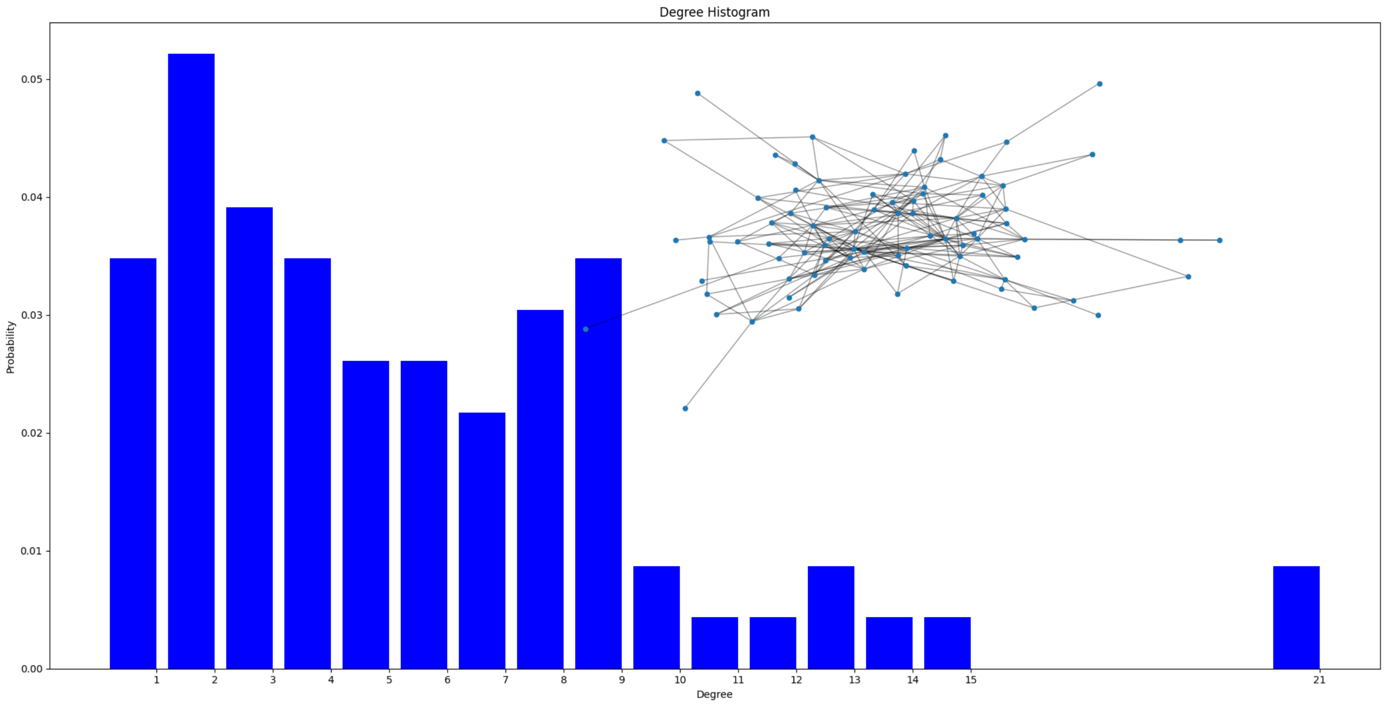


Figure 20: Network and Degree Distribution Histogram

The degree distribution of the Network is shown above. It can be seen that low degrees have a relatively high probability and high degrees have a relatively low probability. Hence, it does not follow the trends of Random network. Hence, it can be shown that the Network of collaboration among NTU SCSE is not a random network.

Scale-free Property(Preferential Attachment)

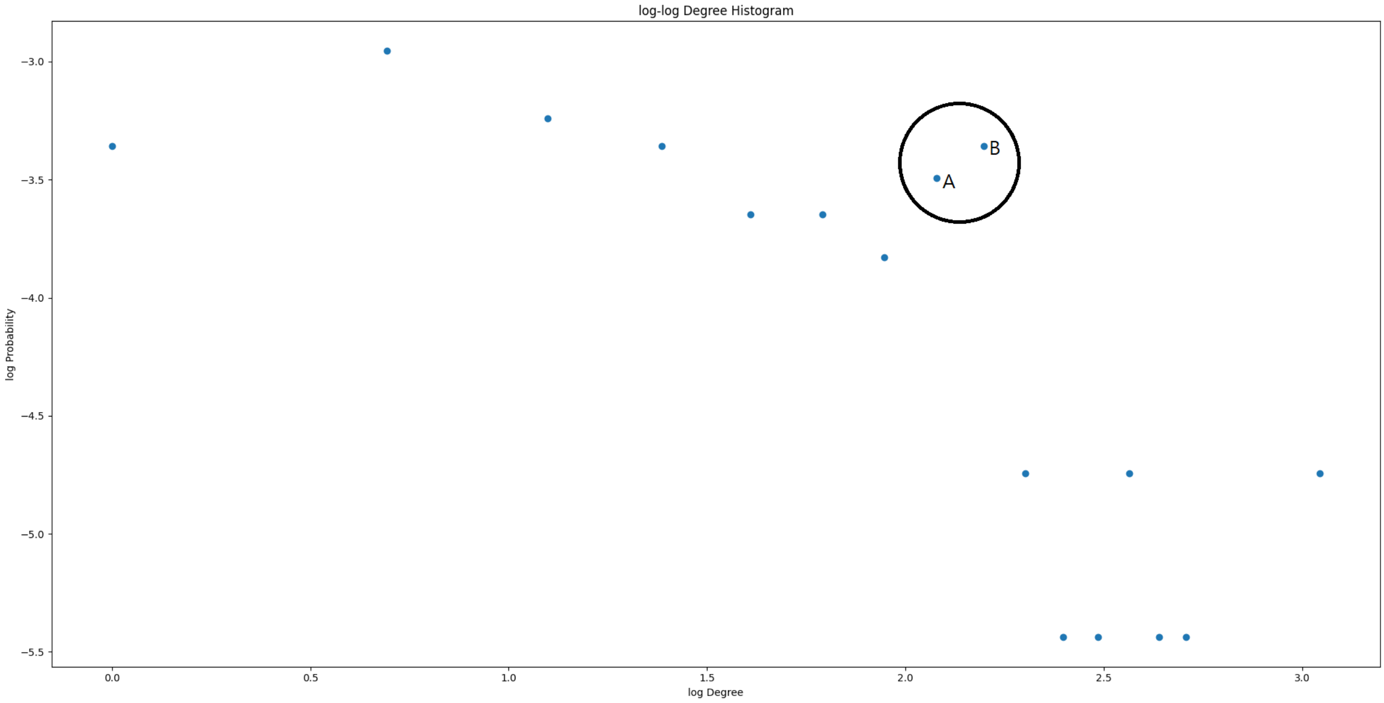


Figure 21: LogLog Degree Distribution Scatter Plot

Preferential Attachment is likely to happen in collaboration networks like this. People who have collaborated more often are likely to be preferred due to their experience. However, the reason why this is not necessarily dominant in this network is the following. The network consists of staff who have different areas of interest. This means that the number of collaborations cannot be prioritized over the area of interest; hence, staff with lower collaborations but same interest are more likely to be chosen for collaboration. Hence, this is why this network will not convey the full characteristic of preferential attachment. This explains a high degree distribution in point A and B (Degree 7-9), as there will be multiple clusters formed, each representing a certain area of interest.

Evolution of the Property

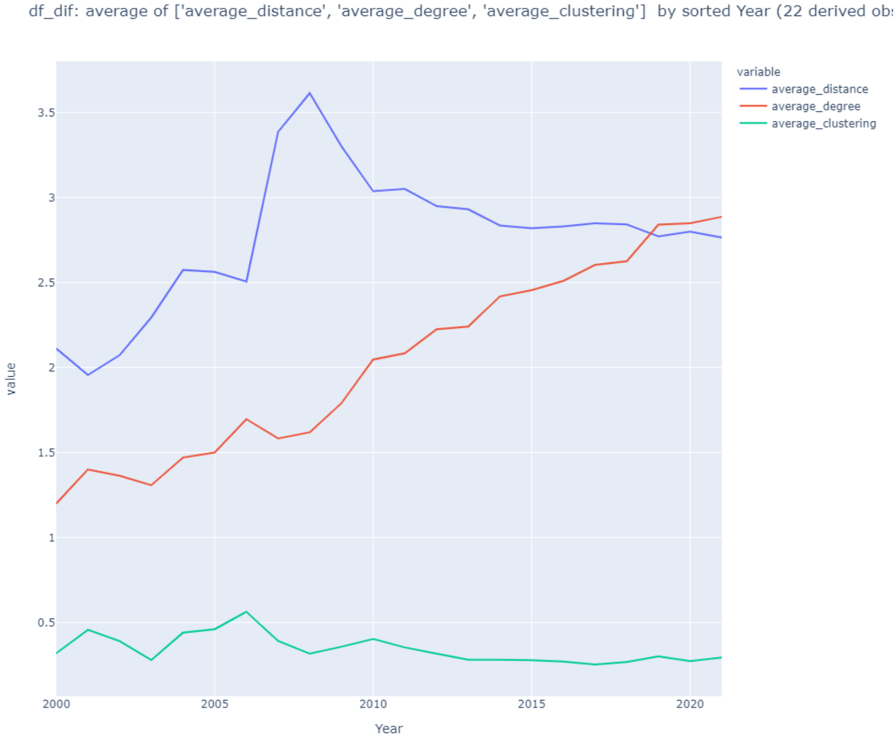


Figure 22: Line graph of Network Property Evolution

Each year, it can be seen that the average degree <k> increases in a linear manner. This indicates that the collaboration between NTU SCSE staff increased every year.

The evolution of average distance <d> has two phases, at which it increases and decreases after 2008. The average distance before 2008 increases as the number of nodes are relatively low and number of edges are low as well. However, as the size of faculty and collaboration increases and reaches a certain threshold, the average distance <d> decreases as addition of new staff and collaboration allows a shorter path.

The average clustering coefficient <C> shows a slight increase until 2008 and from then, a very slow decrease. It indicates that connection of each staff to other staff does not change to a great extent every year. This makes sense as all the staff belong to NTU SCSE so there will be some extent of connections. However, the average clustering coefficient will not increase with more people each year as the staff are divided into their area of interest. Instead of new neighbors connecting to every staff in the network, there will be a tendency to conduct research collaboration with staff sharing a common area of interest. As they have more people to collaborate within the area of interest, they are likely to collaborate less with people outside the area of interest. This explains why the average clustering coefficient <C> decreases in a very slow manner each year.

Limitation

The downside of this model is that it only takes information from the provided Excel file. Since the given faculty information does not represent all the staff who have worked, or is working in NTU SCSE, the model is not an accurate representation of how NTU SCSE has evolved. Currently, the model cannot distinguish if they were working in NTU at the time of their collaboration. It also does not include collaborations of ex-staff in NTU SCSE when they were working in NTU SCSE. If the data can include both staff who have worked and are working in SCSE and the date they entered or left NTU SCSE, it will be a better representation of how NTU SCSE has evolved.

Collaboration

Ranks

i. Network Graph

Figure 23 shows the network graph of collaboration between faculty of different ranks. The weights on the edges represents the total paper count of collaboration between different ranks.

The network graph shows that collaboration between Senior Lecturer and Lecturer does not exist. Assistant Professor, Associate Professor and Professor have all collaborated with one another and with Senior Lecturer and Lecturer.

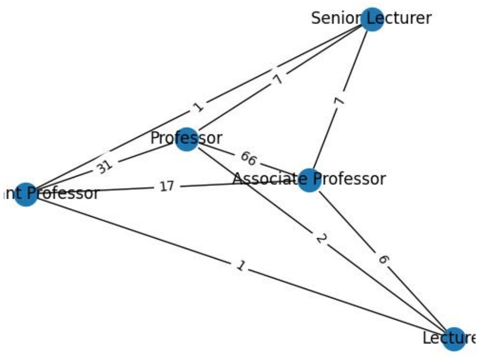


Figure 23: Network Graph - Collaboration between Faculty of Different Ranks

ii. Heatmap

Figure 24 shows a heatmap of the collaboration between different ranks. The numbers indicate the count of papers where the collaborations took place.

From the heatmap, it is shown that the Top 3 collaborations between different ranks are between:

1. Associate Professor and Professor (66 papers)
2. Assistant Professor and Professor (31 papers)
3. Associate Professor and Assistant Professor (17 papers)

Other collaborations between different ranks range from six to seven papers.

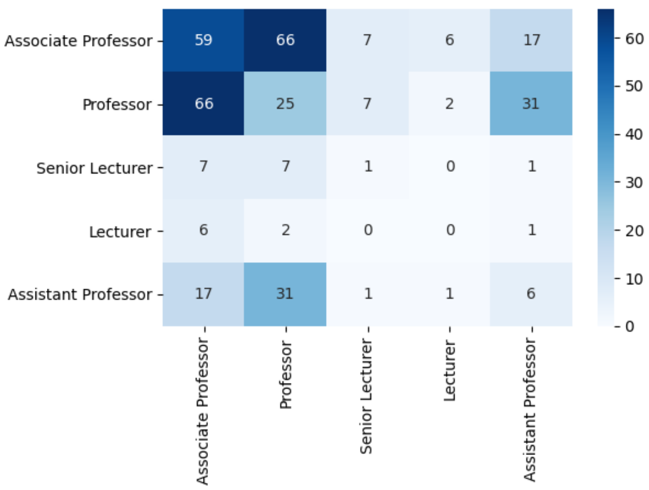
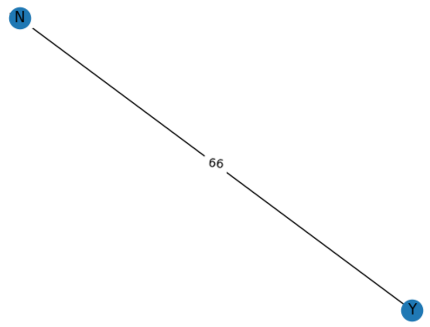


Figure 24: Heatmap - Collaboration between Faculty of Different Rank

Management Position

i. Network graph

Figure 25 shows the network graph of collaboration between faculty holding or held management positions and non-management faculty. The weights on the edges represent the total paper count of collaboration between management and non-management positions. There is only slightly less than one-third of the papers where this collaboration exists. (66 out of 229



papers)

Figure 25: Network Graph - Collaboration between Faculty Holding or Held Management Position and Non-management Faculty

Areas in Computer Science

i. Network Graph

Figure 26 shows the network graph of collaboration between different areas in computer science. Table 1 shows the respective degree of nodes, where each node represents an area in computer science. The higher the degree of the node, the darker it will be. In this case, a higher degree implies collaboration of the area with a wider variety of areas.

The Top 3 areas with most collaborations with other areas are:

1. AI/ML (13 collaborations)
2. Computer Networks, Cyber Security (10 collaborations)
3. Computer Graphics, Multimedia (nine collaborations)

AI/ML is the only area that collaborates with all other areas.

The weight on the edges indicates the count of papers where the collaborations took place.

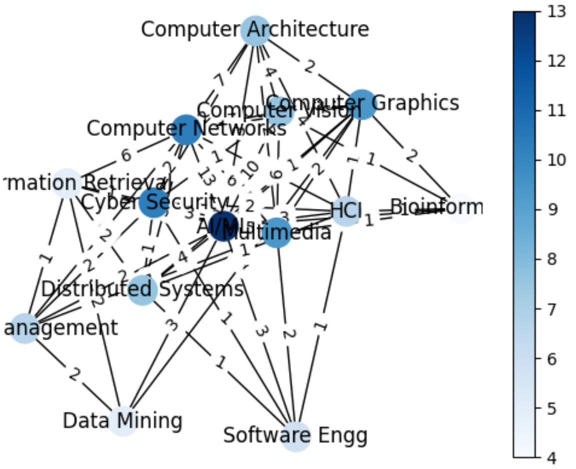


Figure 26: Network Graph - Collaboration between Different Areas in Computer Science

Table 1: Degree of Nodes

Area	Degree
Computer Networks	10
Computer Graphics	9
Computer Architecture	7
AI/ML	13
HCI	8
Cyber Security	10
Distributed Systems	7
Information Retrieval	6
Data Management	6
Data Mining	4
Computer Vision	7
Multimedia	9
Software Engg	5
Bioinformatics	5

ii. Heatmap

Figure 27 shows a heatmap of the collaboration between different areas in computer science. The numbers indicate the count of papers where the collaborations took place.

The Top 3 collaborations with most papers are between:

1. Computer Networks and AI/ML (13 papers)
2. AI/ML and Computer Vision (10 papers)
3. Computer Networks and Computer Architecture,
   Computer Networks and Computer Vision (seven papers)

However, the heatmap also suggests that collaborations are limited as the paper count for most ranges from one to three papers and a few others from four to six papers.

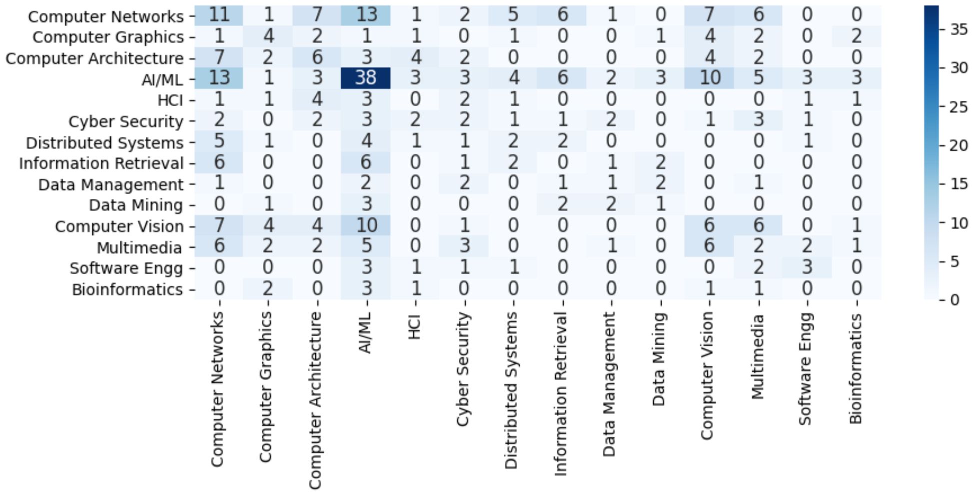


Figure 27: Heatmap - Collaboration between Different Areas in Computer Science

Collaborative Property

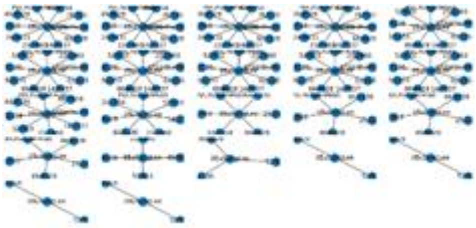
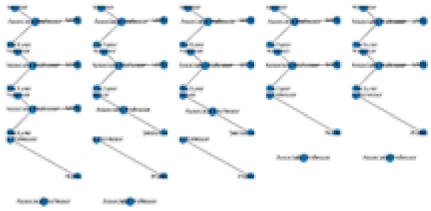
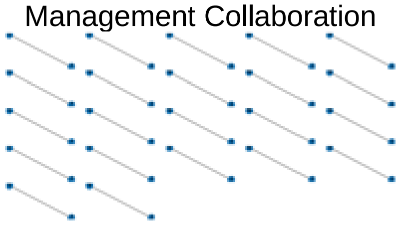
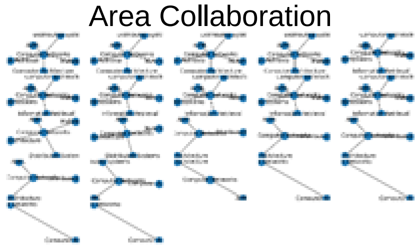
Collaborative properties include collaboration between faculty of different ranks, collaboration between management and non-management faculty and between different areas in computer science.

The set of faculty members used as input is considered based on the number of collaborations. Top 3 faculty members with the highest number of collaborations were considered. They are:

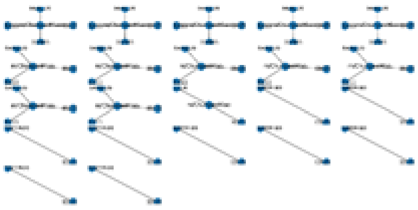
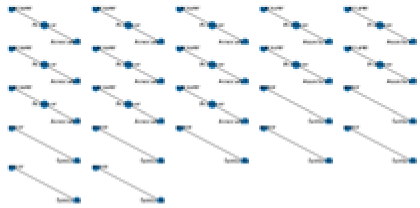
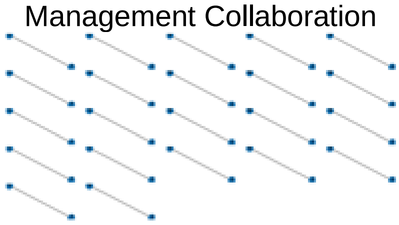
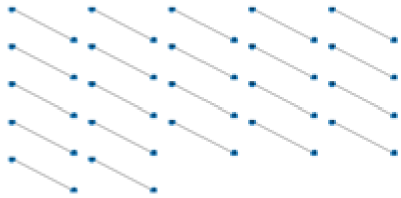
Faculty Member	No. of Collaborations	Position	Management ?	Area in Computer Science
Lee Bu Sung Francis	21	Associate Professor	N	Computer Networks
Miao Chunyan	21	Professor	Y	AI/ML
Ong Yew Soon	15	Professor	Y	AI/ML
Cai Jianfei	14	Professor	N	Computer Vision

The following network graphs shows the collaborative properties for the following faculty members from year 2000 to 2021, with the most recent year shown first.

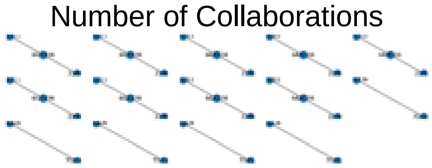
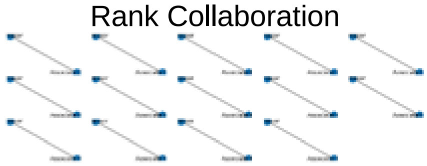
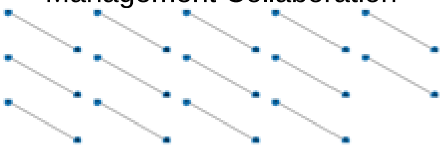
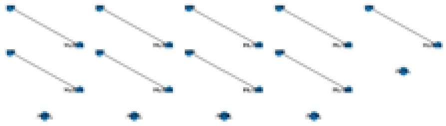
1. Lee Bu Sung Francis

Number of Collaborations  The number of collaborations has increased steadily over the years.	Rank Collaboration  The collaboration started off with only fellow associate professors, then with professors then with senior lecturers and lastly lecturers. He did not collaborate with any assistant professor.
Management Collaboration  He has started off his collaboration with faculty members holding or held management positions and the trend continued over the years.	Area Collaboration  He has started off his collaboration with faculty members in Computer Architecture, then AI/ML, Distributed Systems, Multimedia, and Information Retrieval.

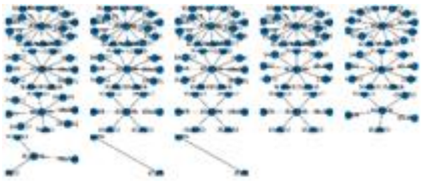
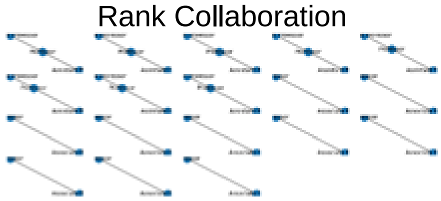
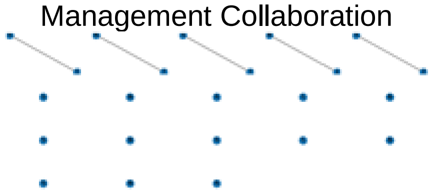
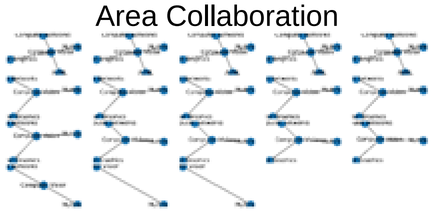
2. Miao Chunyan

<table>
<tr>
<td>Number of Collaborations

Although she is one of the two faculty members with the largest number of collaborations, her collaborations with other faculty members are limited. She did not collaborate with any new faculty members for 9 years (from 2000 to 2008), again for the next 7 years (from 2010 to 2016) and again for the next 5 years (from 2017 to 2021).</td>
<td>Rank Collaboration

Her collaboration started off with a senior lecturer and 10 years later, she collaborated with associate professors. She does not have any collaborations with assistant professors and lecturers.</td>
</tr>
<tr>
<td>Management Collaboration

She has started off her collaboration with non-management faculty and the trend continued over the years.</td>
<td>Area Collaboration

The only collaboration with a different area is with Multimedia. The remaining faculty members she had collaborated with in the first 17 years are all from the same field AI/ML.</td>
</tr>
</table>

3. Ong Yew Soon

Number of Collaborations  He did not have any collaborations from 2000 to 2002. He only collaborated with two other faculty members. This shows that he had multiple repeated collaborations with these two faculty members.	Rank Collaboration  He had only collaborated with associate professors from the time he started collaborating with fellow faculty members.
Management Collaboration  He has started off his collaboration with non-management faculty and the trend continued over the years.	Area Collaboration  The first faculty member he collaborated with is also in the AI/ML area. The other faculty member he collaborated with is in the Multimedia area.

4. Cai Jianfei

Number of Collaborations  He did not have any collaborations from 2000 to 2003. The number of collaborations he had has increased steadily over the years. He did not have any repeated collaboration with the same faculty member.	Rank Collaboration  His collaboration started off with an associate professor and 10 years later, he collaborated with assistant professors. He does not have any collaborations with senior lecturers and lecturers.
Management Collaboration  He only collaborated with fellow non-management faculty for 13 years from his first collaboration. He only started collaborating with faculty members holding or held management positions from year 2017.	Area Collaboration  He has started off his collaboration with Multimedia, then Computer Networks, Computer Graphics. Thereafter, he did not have collaboration with other areas for 10 years (from 2007 to 2016) before he had collaboration with faculty members in the AI/ML area.

Comparison of Central Nodes and Excellent Nodes

Central Node

In this analysis, three different centralities were considered: Degree, Eigenvector and Betweenness. Because each centrality score conveys a different meaning, all three were considered.

Degree centrality in this network helps find the staff who have worked with the largest number of staff in NTU SCSE department.

Eigenvector centrality in this network describes the staff with the largest influence inside the network.

Betweenness centrality in this network shows the staff with the largest influence in the flow of the system.

The network primarily describes how SCSE staff collaborated in publications. Hence, degree centrality is the most appropriate centrality to use to find the 'central' node.

Aim of Central Node vs Excellent Node

Excellent node refers to staff who have appeared frequently in the top venues in his/her area and the word 'excellency' will be interchangeable with the number of publications in top venues. Through comparing this value with degree centrality, below is the purpose of the comparison.

The degree centrality measures how often the individual in NTU SCSE has worked with staff inside NTU SCSE. It might be argued that the simplest way to measure the competency of a university is to view the percentage of the papers published by the university's staff that has published in the top venue. While this is also correct, by comparing the centrality and the excellency, it is possible to derive how a core member of the faculty performs. If the staff who published frequently in the top venue barely collaborates with staff in the same university, it is rather difficult to conclude that the staff represents the competency of the university he/she is in. Hence, by comparing this to the excellency, it is possible to get one of many indirect measures in how competent is NTU SCSE in the world of publications.

Result

Table 2: Top 10 Faculty in Degree Centrality

	Degree Centrality	Eigenvector Centrality	Betweenness Centrality	No.Publication in Top Venue
Lee Bu Sung Francis	0.25	0.330966432	0.127469043	0
Miao Chunyan	0.25	0.351284841	0.117665986	0
Ong Yew Soon	0.178571429	0.220371911	0.099716514	0
Cai Jianfei	0.166666667	0.224538212	0.060469534	16
Dusit Niyato	0.154761905	0.232157732	0.06374236	0
Lin Weisi	0.154761905	0.176008017	0.090011692	9
Sun Aixin	0.142857143	0.187479374	0.03232403	25
Seah Hock Soon	0.130952381	0.074579785	0.102478612	0
Chia Liang Tien Clement	0.119047619	0.171235176	0.019090286	14
Yu Han	0.119047619	0.189466258	0.014666337	0

Table 1 shows the Top 10 staff in terms of highest degree centrality. The person with the highest degree centrality is Lee Bu Sung, Francis and Miao Chunyan with the value of 0.25. However, excellency values are both 0. This indicates that the 'Excellent' node is not identical with the 'Central' Node.

It is extremely dangerous to deduce a conclusion from just a few data. Hence, the scatter plot A was drawn to see the overall relationship between degree centrality and the excellency. As the plot shows, there are no visible trends across the plot. Hence, it is possible to conclude that the degree centrality of the staff has no relationship with the number of publications in the top venue.

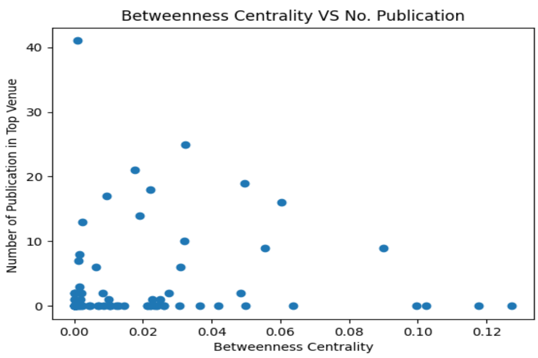


Figure 28: Scatter Plot between Betweenness Centrality and Number of publications in top venue

Limitations and Improvements

The network only accounts for NTU's SCSE faculty given in the list. It is possible that staff are connected through another staff that is not in the list. These individuals could be staff that previously worked in NTU or individuals who work closely with NTU SCSE but are not contracted under NTU SCSE. This can largely change the eigenvector and betweenness centrality of the network. However, this does not influence the degree centrality to a great extent.

Another limitation is that only the top venues for an individual's specialty were considered. It is possible that the staff has more than one specialization and published in different areas. Hence, the number of publications in top venues used in this analysis might not exactly reflect his/her capability to be published in the top venue.

Most importantly, the excellency of the node was determined by the number of publications of the top venues. To get an insight on the more detailed relationship between Centrality and Excellency, each venue could be given a measure that describes the prestigiousness of the venue and number of publications that the individual staff had published in each prestigiousness. This will allow us to get a more accurate prediction on how competent publications done by NTU SCSE is in the world.

Insight

The network is only limited to the staff in NTU SCSE, provided by the lab. Hence, the network centralities only convey information that is limited to our faculty. Hence, this insight is only an estimate with little information. Since staff with a high number of publications in the top venue are not the central nodes, it is possible to derive three predictions.

1. Staff who publish in the Top Venue often are not those who collaborate often with NTU SCSE staff.
2. Staff who collaborate often with NTU SCSE staff tend to not appear often in the Top Venue.
3. There is no relationship between the centrality and the excellency.

With the first two predictions, it is possible to deduce that publications released by collaboration of only/mostly NTU SCSE faculty are not likely to be published in the Top Venues. There might be different reasons and further research is required; however, below are some possible reasons:

1. The most appropriate publication venue for publications through collaboration of mainly NTU SCSE faculty are simply not in any of the top venues. Each venue has its own specialization and target audience; hence, the top venues might simply not fit well for NTU SCSE publications, given the resources and the area of interest in NTU SCSE.
2. Collaborations with other universities could be inevitable to produce publications in the Top Venue, possibly due to lack of resources within the University such as lack of super-computers or collectable data.

However, this does not measure NTU SCSE's ability in any accurate ways because collaboration between different universities is ubiquitous and there are many publications including staff from NTU SCSE that are in the top venue.

Hiring new members to SCSE

Justification

For hiring, the parameter we used is that an author must collaborate at least 4 times with any publisher in SCSE. The reason why we do so is because this shows that the author has sufficient experience working with SCSE members and through the publication process, they are more likely to have the knowledge of SCSE.

Using this, we added about 1600 new members to the graph. The graph nodes represent the SCSE members plus the newcomers and the edges represent which SCSE is related to the author.

This creates many distinct clusters without edges between them as we excluded the initial members relationship.

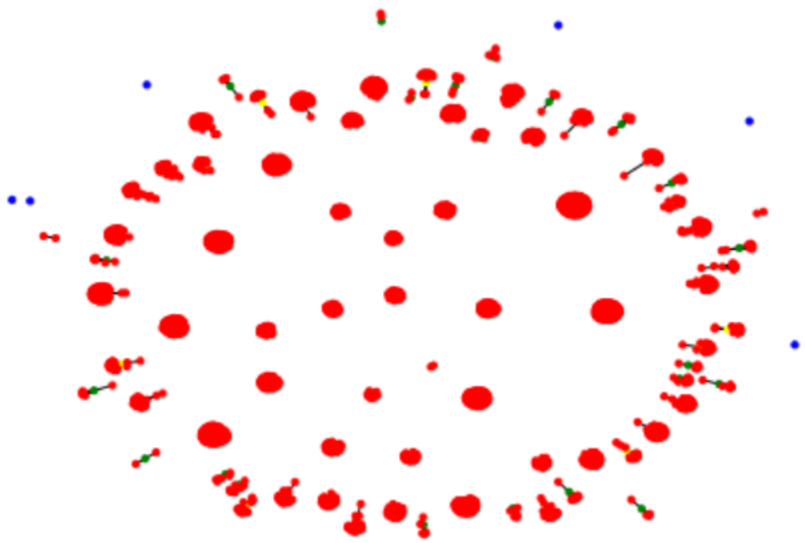


Figure 29: Network after the addition of the new members

Nodes	Betweeness Centrality
76/440	0.002826
51/3710-3	0.001855
14/3737	0.001560
83/6096	0.001010
01/5855	0.000936
83/3179	0.000900
c/ChngEngSiong	0.000796
80/7421	0.000796
99/4535	0.000763
h/YingHe1	0.000496
64/4136	0.000396
42/6178-1	0.000373
95/7006	0.000373
79/8116	0.000350
33/885	0.000350
33/3180	0.000328
m/ChunyanMiao	0.000328
l/BuSungLee	0.000287
62/2078	0.000248
24/914	0.000229

	Nodes	Degree Centrality
0	76/440	0.053444
1	51/3710-3	0.043349
2	14/3737	0.039786
3	83/6096	0.032067
4	01/5855	0.030879
5	83/3179	0.030285
6	c/ChngEngSiong	0.028504
7	80/7421	0.028504
8	99/4535	0.027910
9	h/YingHe1	0.022565
10	64/4136	0.020190
11	42/6178-1	0.019596
12	95/7006	0.019596
13	79/8116	0.019002
14	33/885	0.019002
15	33/3180	0.018409
16	m/ChunyanMiao	0.018409
17	l/BuSungLee	0.017221
18	62/2078	0.016033
19	24/914	0.015439

	Nodes	Eigenvector Centrality
0	76/440	7.071043e-01
1	95/1161-1	7.453550e-02
2	12/6277-4	7.453550e-02
3	50/2082	7.453550e-02
4	43/10100	7.453550e-02
5	r/RMAPRajatheva	7.453550e-02
6	18/6638	7.453550e-02
7	145/2071	7.453550e-02
8	174/9863	7.453550e-02
9	228/0442	7.453550e-02
10	173/5360	7.453550e-02
11	159/1492	7.453550e-02
12	37/1304-1	7.453550e-02
13	86/1369	7.453550e-02
14	83/514	7.453550e-02
15	249/5315	7.453550e-02
16	37/8852	7.453550e-02
17	98/153	7.453550e-02
18	82/6362-1	7.453550e-02
19	92/3836	7.453550e-02

Figure 30:Top nodes by its centralities

Limitation

However, this creates a rift in the increase of members as some departments have a more exponential increase than others. Therefore, this will create a situation where some departments are overstaffed while some departments will have less people.

This is an important issue to address as it will cause some departments to be lacking in manpower and will be unable to teach with the same efficiency as other departments with more members. The morale of the smaller departments will be lower also as they feel that they have more workload comparatively.

On the other hand, it is important to note that this limitation is preconceived by the idea that the initial members in the department are balanced. If the initial amount of members in the department is not balanced, this limitation might not be properly addressed.

Table 3: Change in numbers of members after addition

Department	Initial members	Final members	Percentage Increase
Computer Networks	9	196	2178%
Computer Graphics	4	78	1950%
Computer Architecture	11	177	1601%
Distributed Systems	4	81	2025%
Data Management	2	27	1350%
AI/ML	21	422	2001%
Computer Vision	12	218	1817%
Multimedia	3	114	3800%
Data Mining	2	35	1750%
HCI	3	35	1167%
Information Retrieval	2	38	1900%
Bioinformatics	3	35	1167%
Cyber Security	6	56	933%
Software Engg	3	88	2933%

Appendix

Network Evolution(Connected)

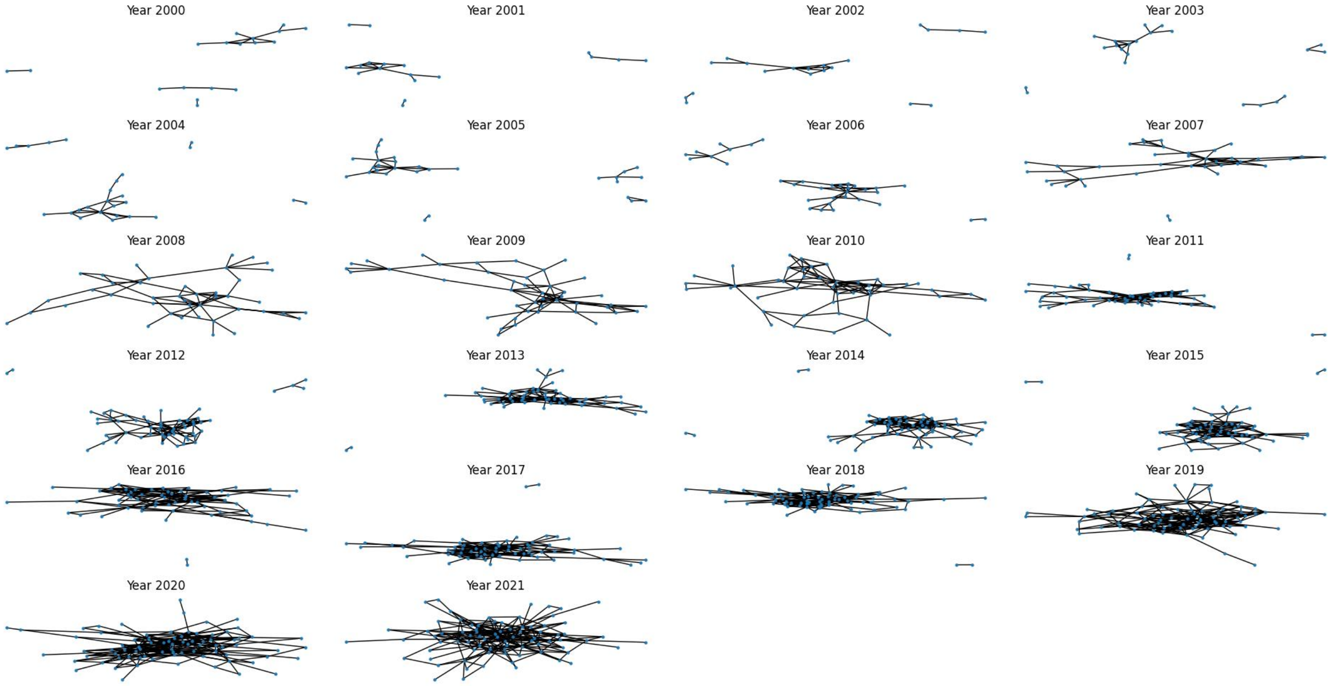


Figure 31: Connected Network Evolution

Network Evolution(Giant)

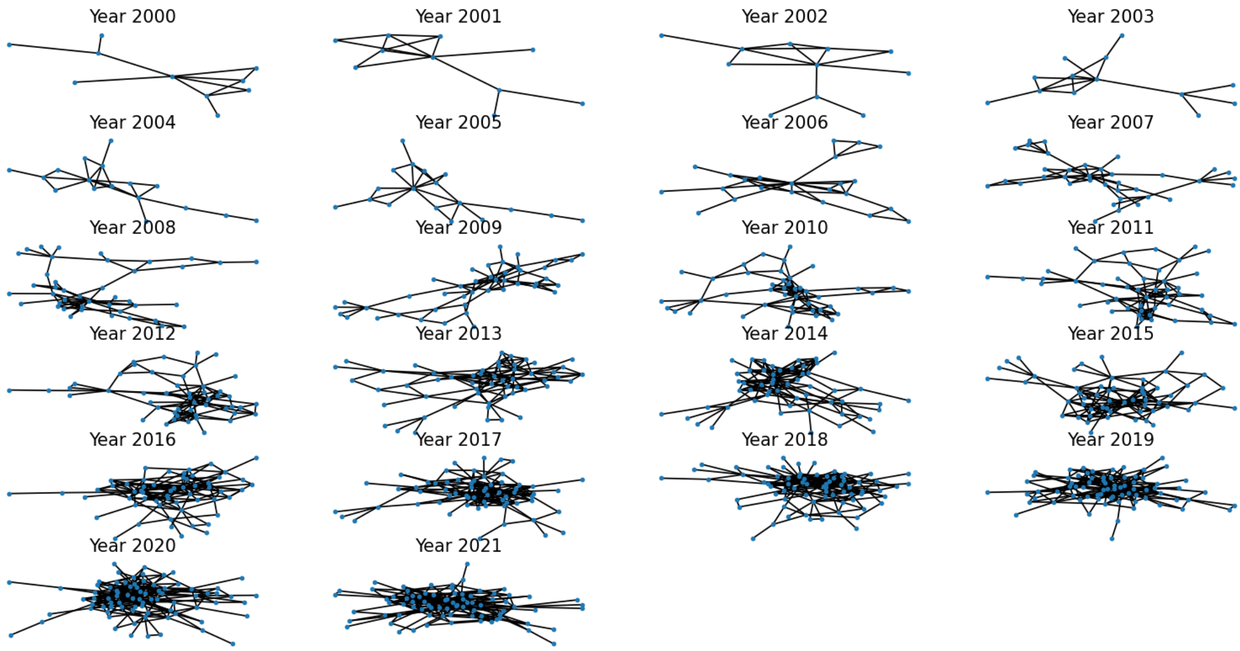


Figure 32: Giant Network Evolution

Degree Centrality Network

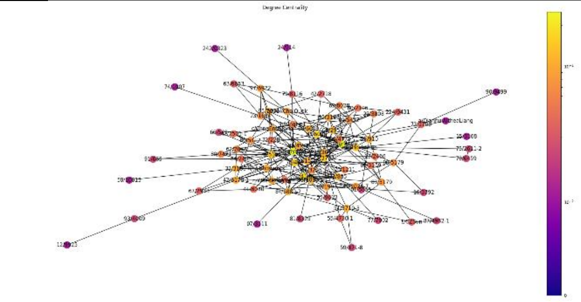


Figure 33: Degree Centrality with Edges

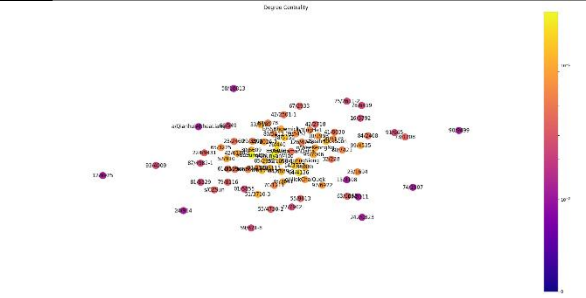


Figure 34: Degree Centrality Without Edges

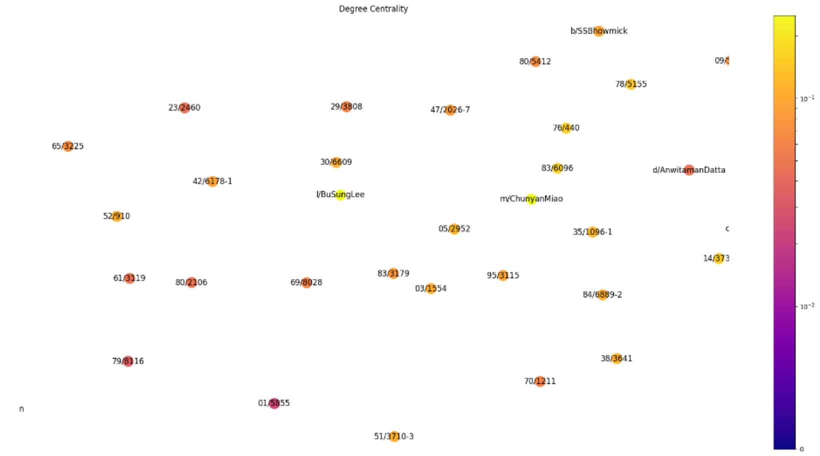


Figure 35: Degree Centrality Network. Zoomed to Centre

Centralities and Number of publications in top venue

Table 4: Top 10 Faculty in Degree Centrality

	Degree Centrality	Eigenvector Centrality	Betweenness Centrality	No.Publication in Top Venue
Lee Bu Sung Francis	0.25	0.330966432	0.127469043	0
Miao Chunyan	0.25	0.351284841	0.117665986	0
Ong Yew Soon	0.178571429	0.220371911	0.099716514	0
Cai Jianfei	0.166666667	0.224538212	0.060469534	16
Dusit Niyato	0.154761905	0.232157732	0.06374236	0
Lin Weisi	0.154761905	0.176008017	0.090011692	9
Sun Aixin	0.142857143	0.187479374	0.03232403	25
Seah Hock Soon	0.130952381	0.074579785	0.102478612	0
Chia Liang Tien Clement	0.119047619	0.171235176	0.019090286	14
Yu Han	0.119047619	0.189466258	0.014666337	0