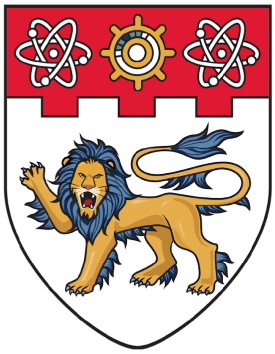

NANYANG TECHNOLOGICAL UNIVERSITY
SINGAPORE

CZ4071: NETWORK SCIENCE PROJECT 1

NETWORK SCIENCE-BASED ANALYSIS OF SCSE FACULTY COLLABORATION

TAN KIAT HWE (U1820766K)

FONG HAO WEI (U1821396H)

HUANG TIANCHENG (U1820154D)

KOH TAT YOU @ ARTHUR KOH (U1821604B)

Pre-processing

In order to obtain the data required from DBLP for the subsequent sections, we had to perform 8 steps of data pre-processing. These steps are outlined below as follow:

Step	Description
1	We first load the provided faculty details into a dataframe. Then, we choose the relevant columns required for scraping from DBLP.
2	Using the previously created faculty details dataframe, we iterate over the provided faculty DBLP pid links to obtain the .xml faculty records of each faculty member from DBLP.
3	Using the obtained 85 .xml faculty records, we use Beautiful Soup 4 to perform parsing and obtain the parsed data records.
4	Using the 85 parsed .xml faculty records, create a new dataframe from them.
5	Using the faculty details dataframe and dblp records dataframe, we perform NLP in order to account for the different representations of each faculty member's name. This is done in order to properly map each faculty member to their published work, as well as to derive how they were credited for their works. (E.g. Were they the first credited name?)
6	We then retrieve the non-SCSE authors from the list of authors who have collaborated with SCSE professors in the previously created DBLP records dataframe.
7	Using this retrieved list of non-SCSE authors, we pick 1000 faculty members of interest for eventual addition to the SCSE faculty graph network.
8	We then create 2 networks: the faculty collaboration network with 85 faculty members, as well as the augmented faculty collaboration network with 1085 faculty members. These networks will then be used in the subsequent questions for further network analysis.

Theoretical Basis:

Finding best fit for degree distribution:

In order to find out the best theoretical distribution to fit the data into, multiple pairwise Likelihood Ratio Tests (LRT) between 2 candidate distributions are conducted. R, which is the ratio of the two likelihoods, is computed. A positive R value indicates that the first candidate is more likely than the second and vice-versa. p, which is the probability of obtaining the value of R, is calculated as well. In our report, we will use a significance level, $\alpha = 0.05$. That is, if $p < 0.05$, the former candidate distribution is significantly more likely than the latter, i.e. not due to chance.

Question 1: What are the network properties of the SCSE faculty network?

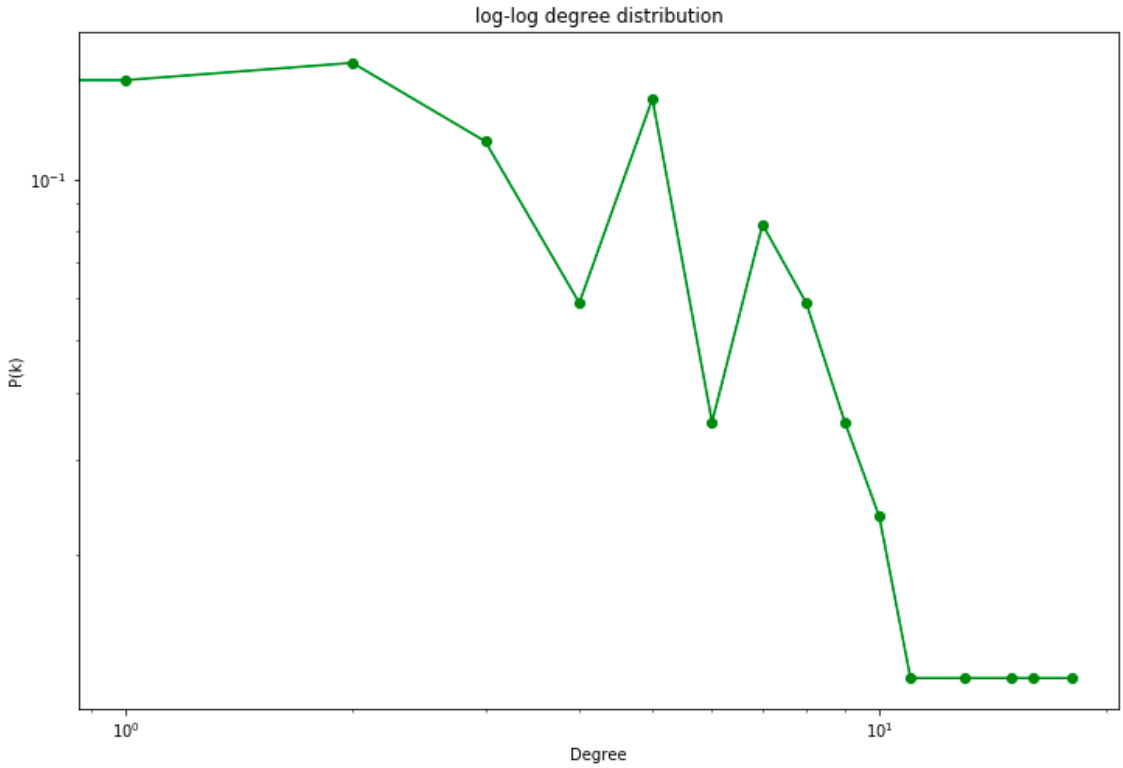

Figure 1.1: Log-Log Degree Distribution Graph

Property	Value
Size of Connected Component	79
Average Degree	4.45
Average Distance	2.95
Diameter	6
Average Clustering Coefficient	0.194

Centrality Measures

Rank	Betweenness Centrality	Eigenvector Centrality	Normalised Degree Centrality	Closeness Centrality
1	Ong Yew Soon 0.142	Miao Chunyan 0.368	Miao Chunyan 0.214	Lee Bu Sung Francis 0.436
2	Lee Bu Sung Francis 0.123	Lee Bu Sung Francis 0.314	Lee Bu Sung Francis 0.190	Miao Chunyan 0.436 (Tied 1st)
3	Miao Chunyan 0.117	Ong Yew Soon 0.272	Ong Yew Soon 0.178	Ong Yew Soon 0.428
4	Seah Hock Soon 0.0919	Cai Jianfei 0.232	Cai Jianfei 0.154	Cai Jianfei 0.409
5	Cai Jianfei 0.0837	Sun Aixin 0.222	Sun Aixin 0.130	Sun Aixin 0.395

Top 5 Faculty for various centrality measures

From the above table, we observe that there are 4 faculty members that are consistently in the top 5 positions based on the various measures. They are namely:

1) Lee Bu Sung Francis
2) Miao Chunyan
3) Ong Yew Soon
4) Cai Jianfei

From the Betweenness Centrality scores, we see that for Ong Yew Soon, roughly 14% of all geodesic paths between other faculty members go through him. It may be possible that he is the most important person when it comes to academic information flow.

Upon observation of the columns for Eigenvector Centrality and Normalised Degree Centrality, we notice that the ranking of the top 5 faculty members are exactly the same. This suggests that they are not just highly connected nodes but also connected to highly connected nodes. Knowing any one of them would put you at a very short network distance away from many professors.

Lastly, for the Closeness Centrality scores, for Lee Bu Sung Francis and Miao Chunyan, we see that they have the highest scores in this metric. This suggests that knowing them would make establishing professional connections with another faculty member faster on average.

The above analysis depends purely on network information, which may not fully capture the interaction dynamics of faculty members in real life.

Best Fitting Distribution

For the most recent faculty collaboration graph, 2021, the best parametric distribution is the Stretched Exponential with parameter estimates $\lambda = 0.283$, $\beta = 0.936$. The following table shows the LRT results.

Distribution	R	p
Power Law	2.77	0.00559
Truncated Power Law	0.722	0.470
Exponential	0.179	0.699
Positive Lognormal	3.39	0.000695

Likelihood of Stretched Exponential vs Others

From the table above, we see that the Stretched Exponential Fit is more likely than the other distributions. Specifically, it is significantly more likely than the Power Law and the Positive Lognormal Distributions.

Question 2: How has the network and its properties evolved since the year 2000?

Average Degree against Year

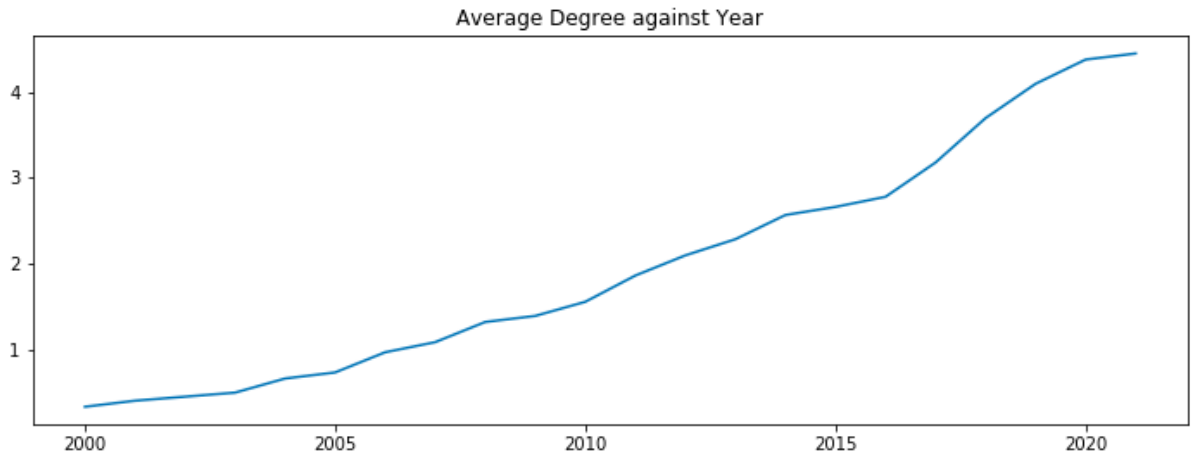

Figure 2.1: 'Average Degree' against 'Year' Graph

From the above figure, we observe that Average Degree has grown rather linearly over the years.

Average Clustering Coefficient against Year

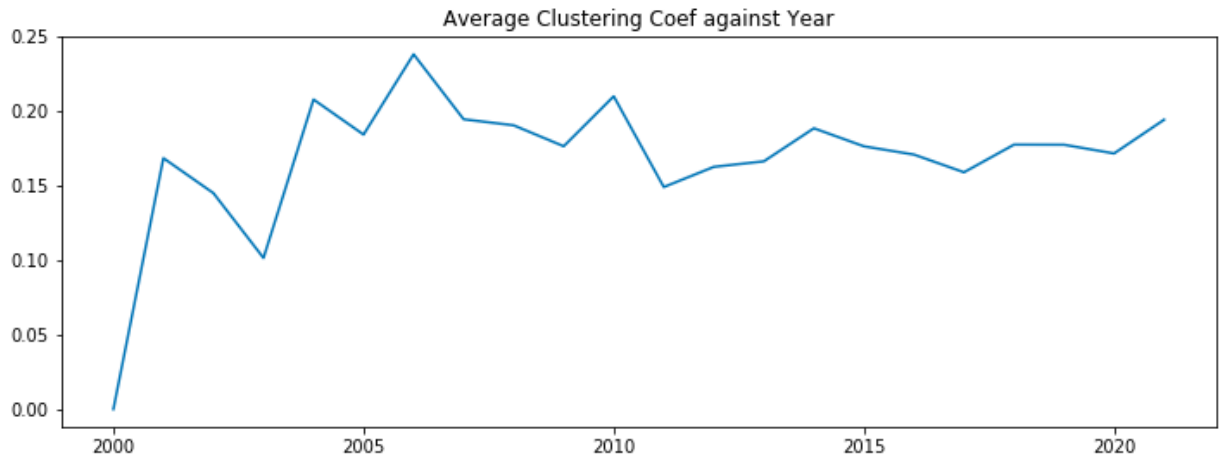

Figure 2.2: 'Average Clustering Coefficient' against 'Year' Graph

The average clustering coefficient increases from 2000 to 2006 before slowly decreasing to a plateau. This may suggest that faculty members may be recommending their co-authors less in recent years to their other co-authors, thus the probability that their co-authors publishing together becomes less.

Average Distance against Year

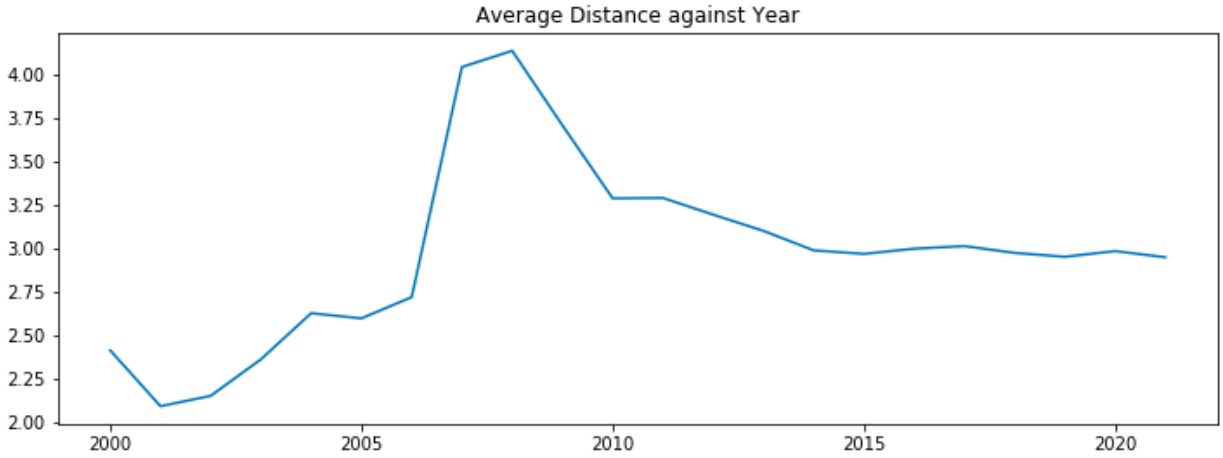

Figure 2.3: 'Average Distance' against 'Year' Graph

From this figure, we notice that average distance grew exponentially till 2008 and then subsequently drops to a plateau.

Size of Connected Component against Year

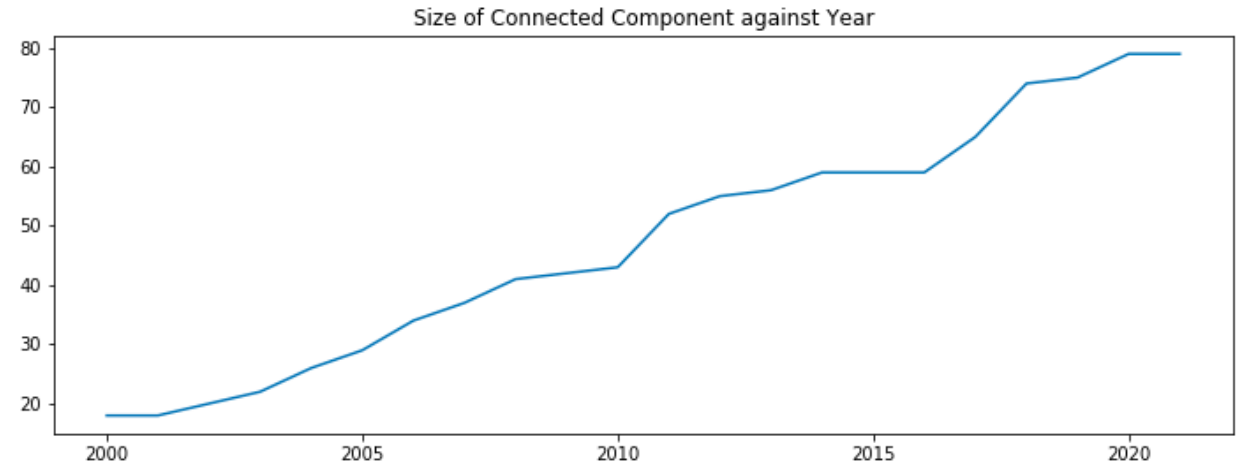

Figure 2.4: 'Connected Component Size' against 'Year' Graph

This seems to be a similar linear trend as Average Degree.

Diameter against Year

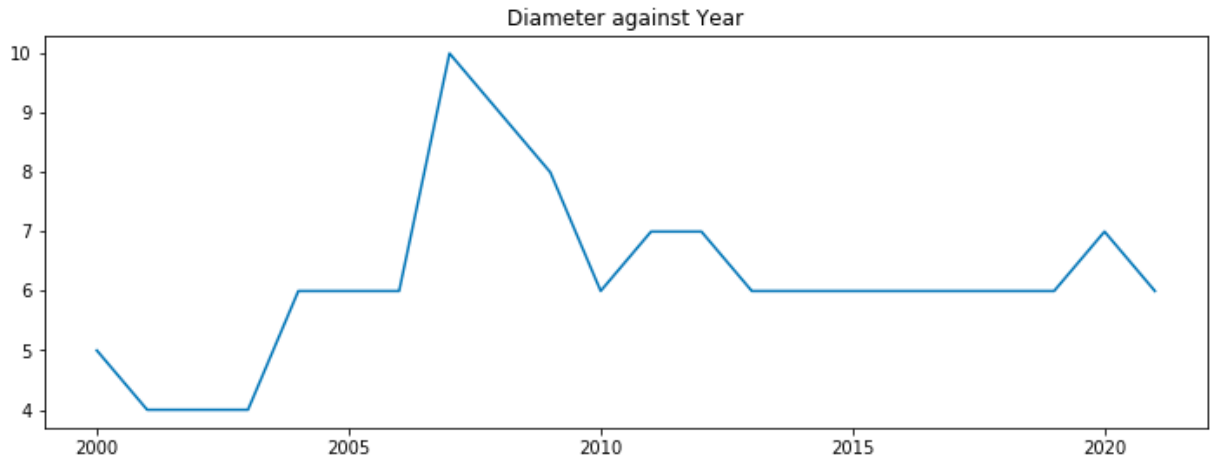

Figure 2.5: 'Diameter' against 'Year' Graph

This seems to follow a similar trend as Average Distance.

Best Fitting Distributions

Year	Distribution
2000	Power Law
2001	Power Law
2002	Power Law
2003	Power Law
2004	Truncated Power Law
2005	Power Law
2006	Truncated Power Law
2007	Truncated Power Law
2008	Truncated Power Law
2009	Truncated Power Law
2010	Truncated Power Law
2011	Truncated Power Law

2012	Truncated Power Law
2013	Truncated Power Law
2014	Truncated Power Law
2015	Truncated Power Law
2016	Truncated Power Law
2017	Truncated Power Law
2018	Truncated Power Law
2019	Stretched Exponential
2020	Stretched Exponential
2021	Stretched Exponential

The above table shows the best fitting distribution based on LRT. Full details on parameter estimates and significance testing can be found in our interface; you may explore these on our interface.

Centrality Measures

To prevent clutter in our report, the **Top 5 Faculty for Various Centrality Measures Per Year** can be found in our interface.

Collaboration

For questions 3 to 5, our group will be assessing intra-class and inter-class collaborations. On top of that, we also measure the rate of participation.

Intra-Class Collaboration

This measures the collaboration within a class. We measure the number of edges present against the theoretical maximum number of edges, i.e. $N(N-1)$ edges. Formally, $IntraCC = \frac{number\ of\ edges\ present}{N\ (N-1)}$.

Inter-Class Collaboration

This measures the collaboration between different classes. Similar to IntraCC, $InterCC = \frac{number\ of\ edges\ present}{Size\ of\ class{-}1 \times Size\ of\ class{-}2}$. The number of edges counted are those that are between a node of each class.

Rate of Participation

This measures the proportion of faculty members in the class that have at least an edge in the subgraph that is being assessed. Formally, $ROP = \frac{number\ of\ faculty\ members\ in\ the\ class\ that\ has\ at\ least\ an\ edge\ in\ the\ subgraph}{number\ of\ faculty\ members\ in\ the\ class}$

Visualisation of ROP, InterCC and IntraCC

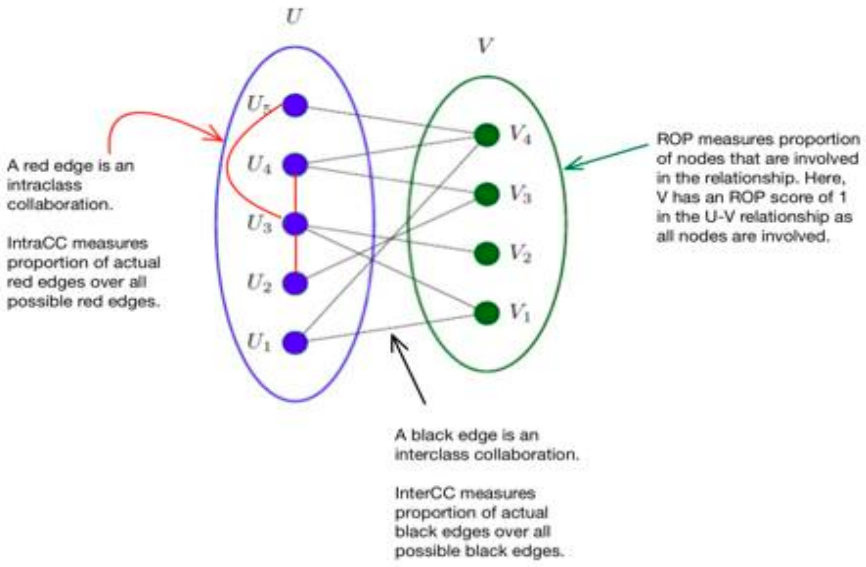

Figure 2.6: Conceptual visualization aid for ROP, InterCC and IntraCC

Question 3: Analyze the collaboration between faculty members of different ranks (e.g:, Professor vs Assistant Professor).

Inter CC

	Professor	Assistant Professor	Associate Professor	Senior Lecturer	Lecturer
Professor	0.096	-	-	-	-
Assistant Professor	-	0.001	-	-	-
Associate Professor	-	-	0.035	-	-
Senior Lecturer	-	-	-	0.400	-
Lecturer	-	-	-	-	-

Intra CC

	Professor	Assistant Professor	Associate Professor	Senior Lecturer	Lecturer
Professor	-	0.077	0.091	0.063	0.021
Assistant Professor	-	-	-	0.010	0.008
Associate Professor	-	0.018	-	0.032	0.018
Senior Lecturer	-	-	-	-	-
Lecturer	-	-	-	-	-

ROP

	Professor & Professor Collab	Professor & Assistant Professor Collab	Professor & Associate Professor Collab	Professor & Senior Lecturer Collab	Professor & Lecturer Collab	Assistant Professor & Assistant Professor Collab	Assistant Professor & Senior Lecturer Collab	Assistant Professor & Lecturer Collab	Associate Professor & Associate Professor Collab	Associate Professor & Assistant Professor Collab	Associate Professor & Senior Lecturer Collab	Associate Professor & Lecturer Collab	Senior Lecturer & Senior Lecturer Collab
Professor	0.813	0.688	1.00	0.250	0.125	-	-	-	-	-	-	-	-
Assistant Professor	-	0.762	0.834	-	-	0.429	0.048	0.048	-	0.476	-	-	-
Associate Professor	-	-	-	-	-	-	-	-	0.811	0.243	0.162	0.081	-
Senior Lecturer	-	-	-	0.600	-	-	0.200	-	-	-	0.600	-	-
Lecturer	-	-	-	-	0.167	-	-	0.217	-	-	-	0.667	0.050

Note: Lecturer-Lecturer and Lecturer-Senior Lecturer tables are not included as no collaboration exists between these two pairings.

Associate Professors seem to intensively seek collaborations with Professors

Looking at the involvement rate of 0.834 in the Professor-Associate Professor collaboration, which is only second to Professors' involvement rate in the same collaboration, it suggests that many of the Associate Professors are seeking mentorship from the Professors in order to secure a tenureship in the future. In this collaboration, it has the highest $InterCC$ out of all inter-class collaborations, further suggesting that working with a Professor in their Associate Professor days may be a good stepping stone to securing their tenureship.

Lecturers and Senior Lecturers are least popular

Across the board, ROP of class X in X-Lecturer or X-Senior Lecturer relationships are low. This is coherent with the knowledge that faculty in the teaching track are not seeking tenures. As such, the inclination for them to partake in research with the tenured or tenure seeking professors is low.

Professors are the most popular

Upon examination of Professor-X collaborations, the ROP of class X is highest as compared to Associate Professor-X or Assistant Professor-X.

Professor-Professor collaboration is most extensive

It has the highest $IntraCC$ score as compared to all other $IntraCC$ and $InterCC$ scores.

Overall

The observations and analyses above seem to point us to the conclusion that being a Professor in NTU requires intensive collaboration with all other ranks. We suspect that there is a colleague assessment that requires input from across the school's faculty members, thus, such collaborations could put them in favourable positions during such evaluations.

Question 4: Analyze the collaboration between faculty holding or held management positions and non-management faculty.

Management-Management

Metric	Score
$IntraCC$	0.143
ROP	0.857

Non-Management-Non-Management

Metric	Score
$IntraCC$	0.0220
ROP	0.846

Management-Non-Management

Metric	Score
$ROP\ (Management)$	1.0
$ROP\ (Non\ -\ Management)$	0.474
$InterCC$	0.0934

Management faculty members have all around highest involvement rate

Apart from the high ROPnumbers, Management also has high $IntraCC\ and\ InterCC$ numbers. This points towards a Management centric work style in NTU, where research usually has a Management faculty to manage the overall workflow.

Non-Management to Non-Management idea flow may be slow

While many of them are involved in at least 1 collaboration, the paltry $IntraCC$ value shows that the average distance between 2 Non-Management staff would be very large.

Overall

From the aforementioned points, it seems that Management faculty members are very important in information passing between faculty members. Their overall high involvement and their high $IntraCC\ and\ InterCC$ shows that they reduce average distance to improve overall information flow. Knowing that, it seems that SCSE may not suffer from poor information flow - a problem that plagues many organisations in their management structure.

Question 5: Analyze the collaboration between faculty of different areas in computer science (data management vs AI/ML).

Some classes do not have any collaboration with each other

Out of 105 possible combinations, there are 41 that do not show collaboration. The following class pairs do not have any collaboration with each other:

1) Computer Networks-Data Mining
2) Computer Networks-Bioinformatics
3) Computer Networks-Software Engineering
4) Computer Graphics-Data Management
5) Computer Graphics-HCI
6) Computer Graphics-Information Retrieval
7) Computer Graphics-Software Engineering
8) Computer Graphics-Cyber Security
9) Computer Architecture-Data Management
10) Computer Architecture-Distributed Systems
11) Computer architecture-Data Mining
12) Computer Architecture-Information Retrieval
13) Computer Architecture-Bioinformatics
14) Computer Architecture-Software Engineering
15) Distributed Systems-Data Management
16) Distributed Systems-Computer Vision
17) Distributed Systems-Multimedia
18) Distributed Systems-Data Mining
19) Distributed Systems-Bioinformatics
20) Data Management-Computer Vision
21) Data Management-HCI
22) Data Management-Bioinformatics
23) Data Management-Software Engg
24) Computer Vision-HCI
25) Computer Vision-Bioinformatics
26) Computer Vision-Software Engineering
27) Multimedia-Data Mining
28) Multimedia-HCI
29) Multimedia-Information Retrieval
30) Multimedia-Software Engineering
31) Data Mining-HCi
32) Data Mining-Bioinformatics
33) Data Mining-Cyber Security
34) Data Mining-Software Engineering
35) HCI-HCI
36) HCI-Information Retrieval

37) Information Retrieval-Information Retrieval
38) Information Retrieval-Bioinformatics
39) Information Retrieval-Software Engineering
40) Bioinformatics-Cyber Security
41) Bioinformatics-Software Engineering

From this list, we are able to identify areas that do not collaborate much with other fields. They are namely: HCI, Computer Networks, Data Management, Software Engineering, Bioinformatics and Information Retrieval. Areas like HCI and Software Engineering are very niche fields and collaboration could be very limited.

AI/ML is everywhere

Looking at the complement of this set, i.e. pairs that are not on the above list, we see the areas that collaboration exists. We notice that AI/ML is totally off the above list, which shows that it is a field that is seeking knowledge from other fields. It's extensiveness, as we know today, could be attributed to the involvement of AI/ML in almost every aspect in our lives.

AI/ML has highest involvements with Information Retrieval and Computer Vision

With Information Retrieval, 23.8% of AI/ML faculty members are involved in at least 1 collaboration. With Computer Vision, that number is 33.3%. This shows a strong coupling of AI/ML with the aforementioned 2 areas. For Information Retrieval in particular, 11.9% of all possible edges are observed ($InterCC = 0.119$) which is considerably high as most $InterCC\ and\ IntraCC$ scores are below 0.1. The extensiveness of collaboration of AI/ML with Information Retrieval is aligned with the industry's movement away from Statistical Based retrieval and into a more Natural Language Processing based Sentiment Analysis to meet the informational needs of users.

Cyber Security starts with Data Management

In this collaboration, the $InterCC$ score is 0.167 which shows a very extensive collaboration between the two domains. Coupled with the highest ROP score among all Cyber Security-X collaborations, we can see that there is a predisposition for this Cyber Security-Data Management collaborations.

BioInformatics require AI/ML

66.6% of BioInformatics faculty has worked with an AI/ML faculty member. With 4.76% of all possible edges, and AI/Ml having a very cardinality, it implies that Bioinformatics could use the expertise of AI/ML faculty for pattern analysis and statistical understanding.

Overall

The importance of the AI/ML department could not be more profound as seen from the collaboration analysis that we have conducted. The full scores can be found in the APPENDIX, but better visualised through our interface.

Question 6: Given a set of faculty members as input, define and track their collaborative properties over time. Think how to define the various collaborative properties.

Apart from the metrics that seen in the previous questions, we want to be able to identify certain key roles in a set of faculty members. This set of faculty members and their collaboration will be termed subgraph in this segment. Their collaboration can be visualised in the interface, which will help to show key information about their collaboration. On top of that, we could identify some potentially important nodes in the subgraph.

Key Information Passer

This person is defined to be the person with the highest betweenness centrality. Identifying this person would be key for a new faculty member to find out who he can talk to in order to find out about the developments that the set of faculty members are working on.

Leader

While not all subsets contain a Management staff, there could be leaders in every subset. This person could be identified by the node that has the highest closeness centrality.

Question 7: Are the central nodes of the network identical to excellence nodes when measured using network properties?

Data Cleaning

The provided input list of different area top venues are not identical to the format used in DBLP. As such, we had to scrape DBLP to find the equivalent tags used for these top venues' associated journals and conferences. After finding the corresponding tags used by DBLP, we used these tags to match with the key values associated with the previously scraped DBLP journal and conference publication records.

Spearman's Correlation

$\rho = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}$ where $\rho = Spearman's\ Rank\ Correlation\ Coefficient$, d_i is the difference in the two ranks for the i-th observation and n is the number of observations.

In this segment, we rank the excellence of a faculty member based on the number of publishings that he/she has in their area's top venue. We then compute the various Centrality measures and rank them based on the values. Finally, the Spearman Correlation Coefficient is derived between the excellence ranking and each of the 4 Centrality Measures.

Null Hypothesis: The ranks are uncorrelated.
Alternative Hypothesis: The ranks are correlated i.e. not due to chance.

Centrality Measure	ρ	p-value
Betweenness	0.162	0.138
Closeness	0.157	0.151
Eigenvector	0.147	0.180
Normalised Degree	0.163	0.136

Analysis

From the table above, we see that the coefficients are not very positive, indicating that the centrality measures do not tell the tale of whether a professor is an excellence node or not. From the p-values, which are all > 0.05, we cannot reject the null hypothesis that the correlations exist. As such, we say that the apparently positive correlation is due to chance.

For example, Lee Bu Sung Francis and Miao Chunyan, who consistently ranked top 5 in centrality measures, rank quite lowly in terms of number of top publications. In contrast, for Cai Jianfei and Sun Aixin, who rank rather highly for centrality measures, also rank highly in terms of number of top publications.

Question 8: Select at least 1000 co-authors (it is up to you to determine how to select them) of the faculty members as potential hires and add them to the network. Analyze how the network properties of the modified network differ from the original faculty network.

Property	Value (Original)	Value (After Hiring)
Size of Connected Component	79	1084
Average Degree	4.45	3.80
Average Distance	2.95	4.44
Diameter	6	10
Average Clustering Coefficient	0.194	0.378

Centrality Measures

Rank	Betweenness Centrality	Eigenvector Centrality	Normalised Degree Centrality	Closeness Centrality
1	Sourav S Bhowmick 0.248	Sourav S Bhowmick 0.587	Sourav S Bhowmick 0.155	Chunyan Miao 0.358
2	Chunyan Miao 0.232	Byron Choi 0.206	Chunyan Miao 0.0470	Sourav S Bhowmick 0.344
3	Anupam Chattopadhyay 0.0955	Huey-Eng Chua 0.131	Lin Weisi 0.0415	Siyuan Liu 0.321
4	Dusit Tao Niyato 0.0935	Xiaokui Xiao 0.129	Dusit Tao Niyato 0.0397	Steven C. H. Hoi 0.321
5	Lee Bu Sung Francis 0.0845	Sanjay Kumar Madria 0.126	Ong Yew Soon 0.0397	Lee Bu Sung Francis 0.309

Top 5 Faculty for various centrality measures after hiring 1000 faculty members

Size of Connected Component grew more than Average Distance and Diameter

Size of Connected Component increased 13.7 fold but Average Distance only increased 1.51 fold and Diameter increased 1.67 fold. The small world property is exhibited here in which the size of the network grows at a much faster rate than the Average Distance and Diameter.

Average Degree decreases

Average Degree decreases from 4.45 to 3.80. This suggests that many nodes that were added are not connected to each other, i.e. haven't collaborated with each other, except for with an SCSE faculty member.

Average Clustering Coefficient increases

Average Clustering Coefficient saw close to a 2 fold increase. As many papers published are between multiple people and for each paper, it is a complete subgraph if seen in isolation, therefore, the probability of a node's neighbours being connected would become higher.

Centrality leaderboard changed in composition

We see that the mainstays of the leaderboard in question 1 are no longer the constants in this segment. Some new additions also make the top 5 for some of the centrality measures.

Best Fitting Distribution

For the extended faculty graph, the best fit distribution is the Truncated Power Law. The table below shows the results of the LRT against the other distributions.

Distribution	R	p
Power Law	0.791	0.496
Stretched Exponential	5.58	2.45×10^{-8}
Exponential	8.58	9.59×10^{-18}
Positive Lognormal	9.35	9.36×10^{-21}

Likelihood of Truncated Power Law vs Others

Appendix

Proportion of domains linked up against Computer Network collaboration pairs

	Computer Network & Computer Networks Collab	Computer Network & Computer Graphics Collab	Computer Network & Computer Architecture Collab	Computer Network & Distributed Systems Collab	Computer Network & Data Management Collab	Computer Network & AI/ML Collab	Computer Network & Computer Vision Collab	Computer Network & Multimedia Collab	Computer Network & Data Mining Collab	Computer Network & HCI Collab	Computer Network & Information Retrieval Collab	Computer Network & Bioinformatics Collab	Computer Network & Cyber Security Collab	Computer Network & Software Engineering Collab
Computer Networks	1.00	0.11	0.33	0.11	0.11	0.56	0.44	0.44	0.00	0.11	0.56	0.00	0.11	0.00
Computer Graphics	-	0.25	-	-	-	-	-	-	-	-	-	-	-	-
Computer Architecture	-	-	0.27	-	-	-	-	-	-	-	-	-	-	-
Distributed Systems	-	-	-	0.75	-	-	-	-	-	-	-	-	-	-
Data Management	-	-	-	-	0.50	-	-	-	-	-	-	-	-	-
AI/ML	-	-	-	-	-	0.38	-	-	-	-	-	-	-	-
Computer Vision	-	-	-	-	-	-	0.25	-	-	-	-	-	-	-
Multimedia	-	-	-	-	-	-	-	1.00	-	-	-	-	-	-
Data Mining	-	-	-	-	-	-	-	-	0.00	-	-	-	-	-
HCI	-	-	-	-	-	-	-	-	-	0.33	-	-	-	-
Information Retrieval	-	-	-	-	-	-	-	-	-	-	1.00	-	-	-
Bioinformatics	-	-	-	-	-	-	-	-	-	-	-	0.00	-	-
Cyber Security	-	-	-	-	-	-	-	-	-	-	-	-	0.17	-
Software Engineering	-	-	-	-	-	-	-	-	-	-	-	-	-	0.00

Proportion of domains linked up against Computer Graphics collaboration pairs

	Computer Graphics & Computer Networks Collab	Computer Graphics & Computer Graphics Collab	Computer Graphics & Computer Architecture Collab	Computer Graphics & Distributed Systems Collab	Computer Graphics & Data Management Collab	Computer Graphics & AI/ML Collab	Computer Graphics & Computer Vision Collab	Computer Graphics & Multimedia Collab	Computer Graphics & Data Mining Collab	Computer Graphics & HCI Collab	Computer Graphics & Information Retrieval Collab	Computer Graphics & Bioinformatics Collab	Computer Graphics & Cyber Security Collab	Computer Graphics & Software Engineering Collab
Computer Networks	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Computer Graphics	-	1.0	0.25	0.25	0.00	0.25	0.50	0.25	0.25	0.00	0.00	0.50	0.00	0.00
Computer Architecture	-	-	0.09	-	-	-	-	-	-	-	-	-	-	-
Distributed Systems	-	-	-	0.25	-	-	-	-	-	-	-	-	-	-
Data Management	-	-	-	-	0.00	-	-	-	-	-	-	-	-	-
AI/ML	-	-	-	-	-	0.05	-	-	-	-	-	-	-	-
Computer Vision	-	-	-	-	-	-	0.17	-	-	-	-	-	-	-
Multimedia	-	-	-	-	-	-	-	0.33	-	-	-	-	-	-
Data Mining	-	-	-	-	-	-	-	-	0.50	-	-	-	-	-
HCI	-	-	-	-	-	-	-	-	-	0.00	-	-	-	-
Information Retrieval	-	-	-	-	-	-	-	-	-	-	0.00	-	-	-
Bioinformatics	-	-	-	-	-	-	-	-	-	-	-	0.67	-	-
Cyber Security	-	-	-	-	-	-	-	-	-	-	-	-	0.00	-
Software Engineering	-	-	-	-	-	-	-	-	-	-	-	-	-	0.00

Proportion of domains linked up against Computer Architecture collaboration pairs

	Computer Architecture & Computer Networks Collab	Computer Architecture & Computer Graphics Collab	Computer Architecture & Computer Architecture Collab	Computer Architecture & Distributed Systems Collab	Computer Architecture & Data Management Collab	Computer Architecture & AI/ML Collab	Computer Architecture & Computer Vision Collab	Computer Architecture & Multimedia Collab	Computer Architecture & Data Mining Collab	Computer Architecture & HCI Collab	Computer Architecture & Information Retrieval Collab	Computer Architecture & Bioinformatics Collab	Computer Architecture & Cyber Security Collab	Computer Architecture & Software Engineering Collab
Computer Networks	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Computer Graphics	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Computer Architecture	-	-	0.55	0.00	0.00	0.18	0.27	0.18	0.00	0.18	0.00	0.00	0.09	0.00
Distributed Systems	-	-	-	0.00	-	-	-	-	-	-	-	-	-	-
Data Management	-	-	-	-	0.00	-	-	-	-	-	-	-	-	-
AI/ML	-	-	-	-	-	0.14	-	-	-	-	-	-	-	-
Computer Vision	-	-	-	-	-	-	0.33	-	-	-	-	-	-	-
Multimedia	-	-	-	-	-	-	-	0.67	-	-	-	-	-	-
Data Mining	-	-	-	-	-	-	-	-	0.00	-	-	-	-	-
HCI	-	-	-	-	-	-	-	-	-	0.67	-	-	-	-
Information Retrieval	-	-	-	-	-	-	-	-	-	-	0.00	-	-	-
Bioinformatics	-	-	-	-	-	-	-	-	-	-	-	0.00	-	-
Cyber Security	-	-	-	-	-	-	-	-	-	-	-	-	0.33	-
Software Engineering	-	-	-	-	-	-	-	-	-	-	-	-	-	0.00

Proportion of domains linked up against Distributed Systems collaboration pairs

	Distributed Systems & Computer Networks Collab	Distributed Systems & Computer Graphics Collab	Distributed Systems & Computer Architecture Collab	Distributed Systems & Distributed Systems Collab	Distributed Systems & Data Management Collab	Distributed Systems & AI/ML Collab	Distributed Systems & Computer Vision Collab	Distributed Systems & Multimedia Collab	Distributed Systems & Data Mining Collab	Distributed Systems & HCI Collab	Distributed Systems & Information Retrieval Collab	Distributed Systems & Bioinformatics Collab	Distributed Systems & Cyber Security Collab	Distributed Systems & Software Engineering Collab
Computer Networks	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Computer Graphics	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Computer Architecture	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Distributed Systems	-	-	-	0.75	0.00	0.75	0.00	0.00	0.00	0.25	0.50	0.00	0.25	0.25
Data Management	-	-	-	-	0.00	-	-	-	-	-	-	-	-	-
AI/ML	-	-	-	-	-	0.19	-	-	-	-	-	-	-	-
Computer Vision	-	-	-	-	-	-	0.00	-	-	-	-	-	-	-
Multimedia	-	-	-	-	-	-	-	0.00	-	-	-	-	-	-
Data Mining	-	-	-	-	-	-	-	-	0.00	-	-	-	-	-
HCI	-	-	-	-	-	-	-	-	-	0.33	-	-	-	-
Information Retrieval	-	-	-	-	-	-	-	-	-	-	0.50	-	-	-
Bioinformatics	-	-	-	-	-	-	-	-	-	-	-	0.00	-	-
Cyber Security	-	-	-	-	-	-	-	-	-	-	-	-	0.17	-
Software Engineering	-	-	-	-	-	-	-	-	-	-	-	-	-	0.33

Proportion of domains linked up against <u>Data Management</u> collaboration pairs

	Data Management & Computer Networks Collab	Data Management & Computer Graphics Collab	Data Management & Computer Architecture Collab	Data Management & Distributed Systems Collab	Data Management & Data Management Collab	Data Management & AI/ML Collab	Data Management & Computer Vision Collab	Data Management & Multimedia Collab	Data Management & Data Mining Collab	Data Management & HCI Collab	Data Management & Information Retrieval Collab	Data Management & Bioinformatics Collab	Data Management & Cyber Security Collab	Data Management & Software Engineering Collab
Computer Networks	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Computer Graphics	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Computer Architecture	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Distributed Systems	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Data Management	-	-	-	-	1.00	0.50	0.00	0.50	1.00	0.00	0.50	0.00	1.00	0.00
AI/ML	-	-	-	-	-	0.05	-	-	-	-	-	-	-	-
Computer Vision	-	-	-	-	-	-	0.00	-	-	-	-	-	-	-
Multimedia	-	-	-	-	-	-	-	0.33	-	-	-	-	-	-
Data Mining	-	-	-	-	-	-	-	-	0.50	-	-	-	-	-
HCI	-	-	-	-	-	-	-	-	-	0.00	-	-	-	-
Information Retrieval	-	-	-	-	-	-	-	-	-	-	0.50	-	-	-
Bioinformatics	-	-	-	-	-	-	-	-	-	-	-	0.00	-	-
Cyber Security	-	-	-	-	-	-	-	-	-	-	-	-	0.33	-
Software Engineering	-	-	-	-	-	-	-	-	-	-	-	-	-	0.00

Proportion of domains linked up against AI/ML collaboration pairs

	AI/ML & Computer Networks Collab	AI/ML & Computer Graphics Collab	AI/ML & Computer Architecture Collab	AI/ML & Distributed Systems Collab	AI/ML & Data Management Collab	AI/ML & AI/ML Collab	AI/ML & Computer Vision Collab	AI/ML & Multimedia Collab	AI/ML & Data Mining Collab	AI/ML & HCI Collab	AI/ML & Information Retrieval Collab	AI/ML & Bioinformatics Collab	AI/ML & Cyber Security Collab	AI/ML & Software Engineering Collab
Computer Networks	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Computer Graphics	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Computer Architecture	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Distributed Systems	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Data Management	-	-	-	-	-	-	-	-	-	-	-	-	-	-
AI/ML	-	-	-	-	-	0.86	0.33	0.19	0.14	0.14	0.24	0.14	0.14	0.14
Computer Vision	-	-	-	-	-	-	0.50	-	-	-	-	-	-	-
Multimedia	-	-	-	-	-	-	-	0.67	-	-	-	-	-	-
Data Mining	-	-	-	-	-	-	-	-	0.50	-	-	-	-	-
HCI	-	-	-	-	-	-	-	-	-	0.33	-	-	-	-
Information Retrieval	-	-	-	-	-	-	-	-	-	-	1.00	-	-	-
Bioinformatics	-	-	-	-	-	-	-	-	-	-	-	0.67	-	-
Cyber Security	-	-	-	-	-	-	-	-	-	-	-	-	0.50	-
Software Engineering	-	-	-	-	-	-	-	-	-	-	-	-	-	0.33

Proportion of domains linked up against Computer Vision collaboration pairs

	Computer Vision & Computer Networks Collab	Computer Vision & Computer Graphics Collab	Computer Vision & Computer Architecture Collab	Computer Vision & Distributed Systems Collab	Computer Vision & Data Management Collab	Computer Vision & AI/ML Collab	Computer Vision & Computer Vision Collab	Computer Vision & Multimedia Collab	Computer Vision & Data Mining Collab	Computer Vision & HCI Collab	Computer Vision & Information Retrieval Collab	Computer Vision & Bioinformatics Collab	Computer Vision & Cyber Security Collab	Computer Vision & Software Engineering Collab
Computer Networks	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Computer Graphics	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Computer Architecture	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Distributed Systems	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Data Management	-	-	-	-	-	-	-	-	-	-	-	-	-	-
AI/ML	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Computer Vision	-	-	-	-	-	-	0.50	0.33	0.08	0.00	0.00	0.00	0.08	0.00
Multimedia	-	-	-	-	-	-	-	1.00	-	-	-	-	-	-
Data Mining	-	-	-	-	-	-	-	-	0.50	-	-	-	-	-
HCI	-	-	-	-	-	-	-	-	-	0.00	-	-	-	-
Information Retrieval	-	-	-	-	-	-	-	-	-	-	0.00	-	-	-
Bioinformatics	-	-	-	-	-	-	-	-	-	-	-	0.00	-	-
Cyber Security	-	-	-	-	-	-	-	-	-	-	-	-	0.16	-
Software Engineering	-	-	-	-	-	-	-	-	-	-	-	-	-	0.00

Proportion of domains linked up against Multimedia collaboration pairs

	Multimedia & Computer Networks Collab	Multimedia & Computer Graphics Collab	Multimedia & Computer Architecture Collab	Multimedia & Distributed Systems Collab	Multimedia & Data Management Collab	Multimedia & AI/ML Collab	Multimedia & Computer Vision Collab	Multimedia & Multimedia Collab	Multimedia & Data Mining Collab	Multimedia & HCI Collab	Multimedia & Information Retrieval Collab	Multimedia & Bioinformatics Collab	Multimedia & Cyber Security Collab	Multimedia & Software Engineering Collab
Computer Networks	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Computer Graphics	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Computer Architecture	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Distributed Systems	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Data Management	-	-	-	-	-	-	-	-	-	-	-	-	-	-
AI/ML	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Computer Vision	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Multimedia	-	-	-	-	-	-	-	1.00	0.00	0.00	0.00	0.33	0.67	0.00
Data Mining	-	-	-	-	-	-	-	-	0.00	-	-	-	-	-
HCI	-	-	-	-	-	-	-	-	-	0.00	-	-	-	-
Information Retrieval	-	-	-	-	-	-	-	-	-	-	0.00	-	-	-
Bioinformatics	-	-	-	-	-	-	-	-	-	-	-	0.33	-	-
Cyber Security	-	-	-	-	-	-	-	-	-	-	-	-	0.17	-
Software Engineering	-	-	-	-	-	-	-	-	-	-	-	-	-	0.00

Proportion of domains linked up against Data Mining collaboration pairs

	Data Mining & Computer Networks Collab	Data Mining & Computer Graphics Collab	Data Mining & Computer Architecture Collab	Data Mining & Distributed Systems Collab	Data Mining & Data Management Collab	Data Mining & AI/ML Collab	Data Mining & Computer Vision Collab	Data Mining & Multimedia Collab	Data Mining & Data Mining Collab	Data Mining & HCI Collab	Data Mining & Information Retrieval Collab	Data Mining & Bioinformatics Collab	Data Mining & Cyber Security Collab	Data Mining & Software Engineering Collab
Computer Networks	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Computer Graphics	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Computer Architecture	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Distributed Systems	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Data Management	-	-	-	-	-	-	-	-	-	-	-	-	-	-
AI/ML	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Computer Vision	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Multimedia	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Data Mining	-	-	-	-	-	-	-	-	1.00	0.00	0.50	0.00	0.00	0.00
HCI	-	-	-	-	-	-	-	-	-	0.00	-	-	-	-
Information Retrieval	-	-	-	-	-	-	-	-	-	-	1.00	-	-	-
Bioinformatics	-	-	-	-	-	-	-	-	-	-	-	0.00	-	-
Cyber Security	-	-	-	-	-	-	-	-	-	-	-	-	0.00	-
Software Engineering	-	-	-	-	-	-	-	-	-	-	-	-	-	0.00

Proportion of domains linked up against HCI collaboration pairs

	HCI & Computer Networks Collab	HCI & Computer Graphics Collab	HCI & Computer Architecture Collab	HCI & Distributed Systems Collab	HCI & Data Management Collab	HCI & AI/ML Collab	HCI & Computer Vision Collab	HCI & Multimedia Collab	HCI & Data Mining Collab	HCI & HCI Collab	HCI & Information Retrieval Collab	HCI & Bioinformatics Collab	HCI & Cyber Security Collab	HCI & Software Engineering Collab
Computer Networks	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Computer Graphics	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Computer Architecture	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Distributed Systems	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Data Management	-	-	-	-	-	-	-	-	-	-	-	-	-	-
AI/ML	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Computer Vision	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Multimedia	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Data Mining	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HCI	-	-	-	-	-	-	-	-	-	0.00	0.00	0.33	0.67	0.33
Information Retrieval	-	-	-	-	-	-	-	-	-	-	0.00	-	-	-
Bioinformatics	-	-	-	-	-	-	-	-	-	-	-	0.33	-	-
Cyber Security	-	-	-	-	-	-	-	-	-	-	-	-	0.33	-
Software Engineering	-	-	-	-	-	-	-	-	-	-	-	-	-	0.33

Proportion of domains linked up against Information Retrieval collaboration pairs

	Information Retrieval & Computer Networks Collab	Information Retrieval & Computer Graphics Collab	Information Retrieval & Computer Architecture Collab	Information Retrieval & Distributed Systems Collab	Information Retrieval & Data Management Collab	Information Retrieval & AI/ML Collab	Information Retrieval & Computer Vision Collab	Information Retrieval & Multimedia Collab	Information Retrieval & Data Mining Collab	Information Retrieval & HCI Collab	Information Retrieval & Information Retrieval Collab	Information Retrieval & Bioinformatics Collab	Information Retrieval & Cyber Security Collab	Information Retrieval & Software Engineering Collab
Computer Networks	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Computer Graphics	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Computer Architecture	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Distributed Systems	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Data Management	-	-	-	-	-	-	-	-	-	-	-	-	-	-
AI/ML	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Computer Vision	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Multimedia	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Data Mining	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HCI	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Information Retrieval	-	-	-	-	-	-	-	-	-	-	0.00	0.00	0.50	0.00
Bioinformatics	-	-	-	-	-	-	-	-	-	-	-	0.00	-	-
Cyber Security	-	-	-	-	-	-	-	-	-	-	-	-	0.17	-
Software Engineering	-	-	-	-	-	-	-	-	-	-	-	-	-	0.00

Proportion of domains linked up against <u>Bioinformatics</u> collaboration pairs

	Bioinformatics & Computer Networks Collab	Bioinformatics & Computer Graphics Collab	Bioinformatics & Computer Architecture Collab	Bioinformatics & Distributed Systems Collab	Bioinformatics & Data Management Collab	Bioinformatics & AI/ML Collab	Bioinformatics & Computer Vision Collab	Bioinformatics & Multimedia Collab	Bioinformatics & Data Mining Collab	Bioinformatics & HCI Collab	Bioinformatics & Information Retrieval Collab	Bioinformatics & Bioinformatics Collab	Bioinformatics & Cyber Security Collab	Bioinformatics & Software Engineering Collab
Computer Networks	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Computer Graphics	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Computer Architecture	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Distributed Systems	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Data Management	-	-	-	-	-	-	-	-	-	-	-	-	-	-
AI/ML	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Computer Vision	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Multimedia	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Data Mining	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HCI	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Information Retrieval	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Bioinformatics	-	-	-	-	-	-	-	-	-	-	-	0.00	0.00	0.00
Cyber Security	-	-	-	-	-	-	-	-	-	-	-	-	0.00	-
Software Engineering	-	-	-	-	-	-	-	-	-	-	-	-	-	0.00

Proportion of domains linked up against Cyber Security collaboration pairs

	Cyber Security & Computer Networks Collab	Cyber Security & Computer Graphics Collab	Cyber Security & Computer Architecture Collab	Cyber Security & Distributed Systems Collab	Cyber Security & Data Management Collab	Cyber Security & AI/ML Collab	Cyber Security & Computer Vision Collab	Cyber Security & Multimedia Collab	Cyber Security & Data Mining Collab	Cyber Security & HCI Collab	Cyber Security & Information Retrieval Collab	Cyber Security & Bioinformatics Collab	Cyber Security & Cyber Security Collab	Cyber Security & Software Engineering Collab
Computer Networks	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Computer Graphics	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Computer Architecture	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Distributed Systems	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Data Management	-	-	-	-	-	-	-	-	-	-	-	-	-	-
AI/ML	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Computer Vision	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Multimedia	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Data Mining	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HCI	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Information Retrieval	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Bioinformatics	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Cyber Security	-	-	-	-	-	-	-	-	-	-	-	-	0.50	0.17
Software Engineering	-	-	-	-	-	-	-	-	-	-	-	-	-	0.33

Proportion of domains linked up against Software Engineering collaboration pairs

	Software Engineering & Computer Networks Collab	Software Engineering & Computer Graphics Collab	Software Engineering & Computer Architecture Collab	Software Engineering & Distributed Systems Collab	Software Engineering & Data Management Collab	Software Engineering & AI/ML Collab	Software Engineering & Computer Vision Collab	Software Engineering & Multimedia Collab	Software Engineering & Data Mining Collab	Software Engineering & HCI Collab	Software Engineering & Information Retrieval Collab	Software Engineering & Bioinformatics Collab	Software Engineering & Cyber Security Collab	Software Engineering & Software Engineering Collab
Computer Networks	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Computer Graphics	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Computer Architecture	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Distributed Systems	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Data Management	-	-	-	-	-	-	-	-	-	-	-	-	-	-
AI/ML	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Computer Vision	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Multimedia	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Data Mining	-	-	-	-	-	-	-	-	-	-	-	-	-	-
HCI	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Information Retrieval	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Bioinformatics	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Cyber Security	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Software Engineering	-	-	-	-	-	-	-	-	-	-	-	-	-	1.00

Observed Possible Two-Domain Bigraph Edges within Collaboration Network

	Computer Networks	Computer Graphics	Computer Architecture	Distributed Systems	Data Management	AI/ML	Computer Vision	Multimedia	Data Mining	HCI	Information Retrieval	Bioinformatics	Cyber Security	Software Engineering
Computer Networks	0.15	0.02	0.03	0.08	0.05	0.05	0.03	0.19	0.00	0.04	0.28	0.00	0.02	0.00
Computer Graphics	-	0.33	0.02	0.06	0.00	0.01	0.04	0.08	0.13	0.00	0.00	0.17	0.00	0.00
Computer Architecture	-	-	0.05	0.00	0.00	0.01	0.03	0.06	0.00	0.06	0.00	0.00	0.03	0.00
Distributed Systems	-	-	-	0.17	0.00	0.05	0.00	0.00	0.00	0.08	0.25	0.00	0.04	0.08
Data Management	-	-	-	-	0.50	0.02	0.00	0.17	0.50	0.00	0.25	0.00	0.17	0.00
AI/ML	-	-	-	-	-	0.07	0.04	0.06	0.07	0.05	0.12	0.05	0.02	0.05
Computer Vision	-	-	-	-	-	-	0.04	0.17	0.04	0.00	0.00	0.00	0.01	0.00
Multimedia	-	-	-	-	-	-	-	0.33	0.00	0.00	0.00	0.11	0.11	0.00
Data Mining	-	-	-	-	-	-	-	-	0.50	0.00	0.50	0.00	0.00	0.00
HCI	-	-	-	-	-	-	-	-	-	0.00	0.00	0.11	0.11	0.11
Information Retrieval	-	-	-	-	-	-	-	-	-	-	0.00	0.00	0.08	0.00
Bioinformatics	-	-	-	-	-	-	-	-	-	-	-	0.00	0.00	0.00
Cyber Security	-	-	-	-	-	-	-	-	-	-	-	-	0.07	0.06
Software Engineering	-	-	-	-	-	-	-	-	-	-	-	-	-	0.33