



SOCIAL NETWORK ANALYSIS

Assignment 1 – CPIS483



Team Members:

Ruyan Nasser Alatawi	1505563
Aseel Hani Qutub	1506455
Munirah Hamad Alhassan	1506765

IBM Case Study

Retaining Valuable Employees

The HR analyst employee in IBM company wants to determine which factors keep employees at the company and which prompt others to leave to prevent the loss of good people.

The HR analyst collects the data in the spreadsheet about past and present employees, he is aiming to understand how this relates to workforce attrition and to make better decisions about employee retention. To achieve the aims, we will analyze the data using Gephi Software tool.

➤ Implementation

1. The spreadsheet contents

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
	Age	Attrition	BusinessTravel	DailyRate	Department	DistanceFromHome	Education	EducationField	EmployeeCount	EmployeeNumber	EnvironmentSatisfaction	Gender	HourlyRate	JobInvolvement	JobLevel	JobRole	JobSatisfaction	MaritalStatus	MonthlyIncome
1																			
2	41	Yes	Travel_Rare	1102	Sales	1	2	Life Sciences	1	1	2	Female	94	3	2	Sales Execut	4	Single	5993
3	49	No	Travel_Frequ	279	Research & I	8	1	Life Sciences	1	2	3	Male	61	2	2	Research Sci	2	Married	5130
4	37	Yes	Travel_Rare	1373	Research & I	2	2	Other	1	4	4	Male	92	2	1	Laboratory Ti	3	Single	2090
5	33	No	Travel_Frequ	1392	Research & I	3	4	Life Sciences	1	5	4	Female	56	3	1	Research Sci	3	Married	2909
6	27	No	Travel_Rare	591	Research & I	2	1	Medical	1	7	1	Male	40	3	1	Laboratory Ti	2	Married	3468
7	32	No	Travel_Frequ	1005	Research & I	2	2	Life Sciences	1	8	4	Male	79	3	1	Laboratory Ti	4	Single	3068
8	59	No	Travel_Rare	1324	Research & I	3	3	Medical	1	10	3	Female	81	4	1	Laboratory Ti	1	Married	2670
9	30	No	Travel_Rare	1356	Research & I	24	1	Life Sciences	1	11	4	Male	67	3	1	Laboratory Ti	3	Divorced	2693
10	38	No	Travel_Frequ	216	Research & I	23	3	Life Sciences	1	12	4	Male	44	2	3	Manufacturi	3	Single	9526
11	36	No	Travel_Rare	1299	Research & I	27	3	Medical	1	13	3	Male	94	3	2	Healthcare R	3	Married	5237
12	35	No	Travel_Rare	809	Research & I	16	3	Medical	1	14	1	Male	84	4	1	Laboratory Ti	2	Married	2426
13	29	No	Travel_Rare	153	Research & I	15	2	Life Sciences	1	15	4	Female	49	2	2	Laboratory Ti	3	Single	4193
14	31	No	Travel_Rare	670	Research & I	26	1	Life Sciences	1	16	1	Male	31	3	1	Research Sci	3	Divorced	2911
15	34	No	Travel_Rare	1346	Research & I	19	2	Medical	1	18	2	Male	93	3	1	Laboratory Ti	4	Divorced	2661
16	28	Yes	Travel_Rare	103	Research & I	24	3	Life Sciences	1	19	3	Male	50	2	1	Laboratory Ti	3	Single	2028
17	29	No	Travel_Rare	1389	Research & I	21	4	Life Sciences	1	20	2	Female	51	4	3	Manufacturi	1	Divorced	9980
18	32	No	Travel_Rare	334	Research & I	5	2	Life Sciences	1	21	1	Male	80	4	1	Research Sci	2	Divorced	3298
19	22	No	Non-Travel	1123	Research & I	16	2	Medical	1	22	4	Male	96	4	1	Laboratory Ti	4	Divorced	2935
20	53	No	Travel_Rare	1219	Sales	2	4	Life Sciences	1	23	1	Female	78	2	4	Manager	4	Married	15427
21	38	No	Travel_Rare	371	Research & I	2	3	Life Sciences	1	24	4	Male	45	3	1	Research Sci	4	Single	3944
22	24	No	Non-Travel	673	Research & I	11	2	Other	1	25	1	Female	96	4	2	Manufacturi	3	Divorced	4011
23	36	Yes	Travel_Rare	1218	Sales	9	4	Life Sciences	1	27	3	Male	82	2	1	Sales Repres	1	Single	3407
24	34	No	Travel_Rare	419	Research & I	7	4	Life Sciences	1	28	1	Female	53	3	3	Research Dir	2	Single	11994
25	21	No	Travel_Rare	391	Research & I	15	2	Life Sciences	1	30	3	Male	96	3	1	Research Sci	4	Single	1232
26	34	Yes	Travel_Rare	699	Research & I	6	1	Medical	1	31	2	Male	83	3	1	Research Sci	1	Single	2960
27	53	No	Travel_Rare	1282	Research & I	5	3	Other	1	32	3	Female	58	3	5	Manager	3	Divorced	19094
28	32	Yes	Travel_Frequ	1125	Research & I	16	1	Life Sciences	1	33	2	Female	72	1	1	Research Sci	1	Single	3919
29	42	No	Travel_Rare	691	Sales	8	4	Marketing	1	35	3	Male	48	3	2	Sales Execut	2	Married	6825
30	44	No	Travel_Rare	477	Research & I	7	4	Medical	1	36	1	Female	42	2	3	Healthcare R	4	Married	10248
31	46	No	Travel_Rare	705	Sales	2	4	Marketing	1	38	2	Female	83	3	5	Manager	1	Single	18947
32	33	No	Travel_Rare	924	Research & I	2	3	Medical	1	39	3	Male	78	3	1	Laboratory Ti	4	Single	2496
33	44	No	Travel_Rare	1459	Research & I	10	4	Other	1	40	4	Male	41	3	2	Healthcare R	4	Married	6465
34	30	No	Travel_Rare	125	Research & I	9	2	Medical	1	41	4	Male	83	2	1	Laboratory Ti	3	Single	2206
35	39	Yes	Travel_Rare	895	Sales	5	3	Technical De	1	42	4	Male	56	3	2	Sales Repres	4	Married	2086
36	24	Yes	Travel_Rare	813	Research & I	1	3	Medical	1	45	2	Male	61	3	1	Research Sci	4	Married	2293

Figure 1 Spreadsheet

T	U	V	W	X	Y	Z	AA	AB	AC	AD	AE	AF	AG	AH	AI	AJ
MonthlyRate	NumCompanies Worked	Over18	OverTime	PercentSalary Hike	Performance Rating	Relationship Satisfaction	StandardHours	StockOption Level	TotalWorking Years	TrainingTime LastYear	WorkLifeBalance	YearsAtCompany	YearsInCurrent Role	YearsSinceLast Promotion	YearsWithCurrent Manager	
19479	8	Y	Yes	11	3	1	80	0	8	0	1	6	4	0	5	
24907	1	Y	No	23	4	4	80	1	10	3	3	10	7	1	7	
2396	6	Y	Yes	15	3	2	80	0	7	3	3	0	0	0	0	
23159	1	Y	Yes	11	3	3	80	0	8	3	3	8	7	3	0	
16632	9	Y	No	12	3	4	80	1	6	3	3	2	2	2	2	
11864	0	Y	No	13	3	3	80	0	8	2	2	7	7	3	6	
9964	4	Y	Yes	20	4	1	80	3	12	3	2	1	0	0	0	
13335	1	Y	No	22	4	2	80	1	1	2	3	1	0	0	0	
8787	0	Y	No	21	4	2	80	0	10	2	3	9	7	1	8	
16577	6	Y	No	13	3	2	80	2	17	3	2	7	7	7	7	
16479	0	Y	No	13	3	3	80	1	6	5	3	5	4	0	3	
12682	0	Y	Yes	12	3	4	80	0	10	3	3	9	5	0	8	
15170	1	Y	No	17	3	4	80	1	5	1	2	5	2	4	3	
8758	0	Y	No	11	3	3	80	1	3	2	3	2	2	1	2	
12947	5	Y	Yes	14	3	2	80	0	6	4	3	4	2	0	3	
10195	1	Y	No	11	3	3	80	1	10	1	3	10	9	8	8	
15053	0	Y	Yes	12	3	4	80	2	7	5	2	6	2	0	5	
7324	1	Y	Yes	13	3	2	80	2	1	2	2	1	0	0	0	
22021	2	Y	No	16	3	3	80	0	31	3	3	25	8	3	7	
4306	5	Y	Yes	11	3	3	80	0	6	3	3	3	2	1	2	
8232	0	Y	No	18	3	4	80	1	5	5	2	4	2	1	3	
6986	7	Y	No	23	4	2	80	0	10	4	3	5	3	0	3	
21293	0	Y	No	11	3	3	80	0	13	4	3	12	6	2	11	
19281	1	Y	No	14	3	4	80	0	0	6	3	0	0	0	0	
17102	2	Y	No	11	3	3	80	0	8	2	3	4	2	1	3	
10735	4	Y	No	11	3	4	80	1	26	3	2	14	13	4	8	
4681	1	Y	Yes	22	4	2	80	0	10	5	3	10	2	6	7	
21173	0	Y	No	11	3	4	80	1	10	2	3	9	7	4	2	
2094	3	Y	No	14	3	4	80	1	24	4	3	22	6	5	17	
22822	3	Y	No	12	3	4	80	0	22	2	2	2	2	2	1	
6670	4	Y	No	11	3	4	80	0	7	3	3	1	1	0	0	
19121	2	Y	Yes	13	3	4	80	0	9	5	4	4	2	1	3	
16117	1	Y	No	13	3	1	80	0	10	5	3	10	0	1	8	
3335	3	Y	No	14	3	3	80	1	19	6	4	1	0	0	0	
3020	2	Y	Yes	16	3	1	80	1	6	2	2	2	0	2	0	

Figure 2 Spreadsheet

2. Import our dataset from Excel file

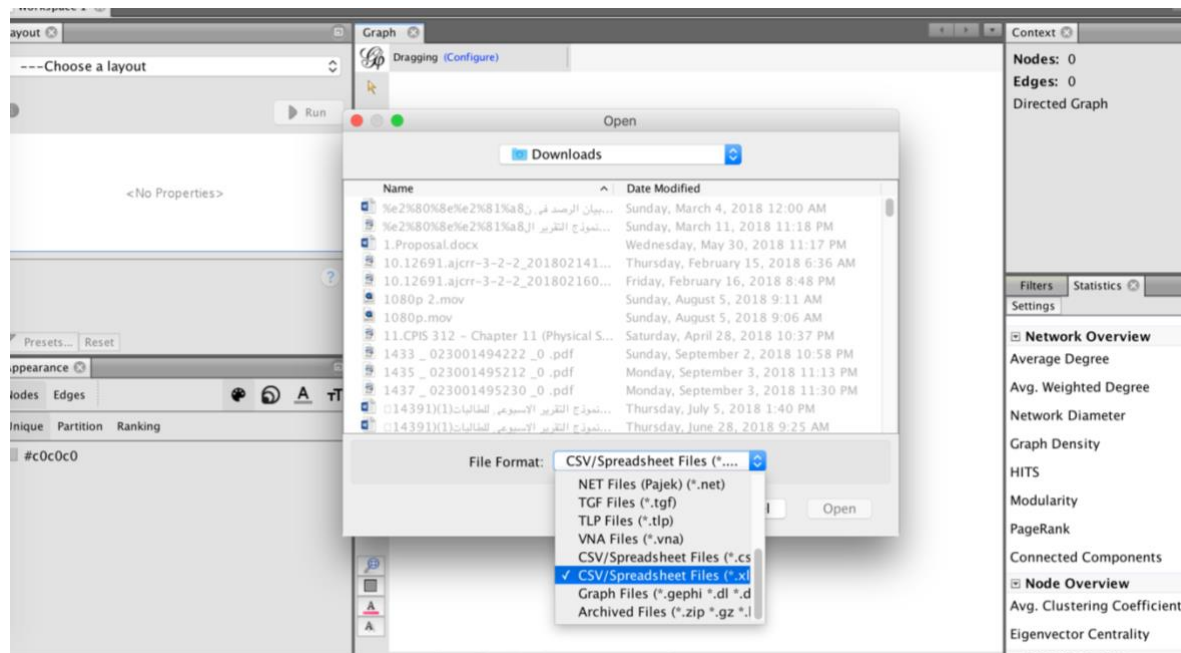


Figure 3 Import to Gephi

3. The network is in a mess so first thing to do is layout our network by applying Yifan Hu algorithm on the data.

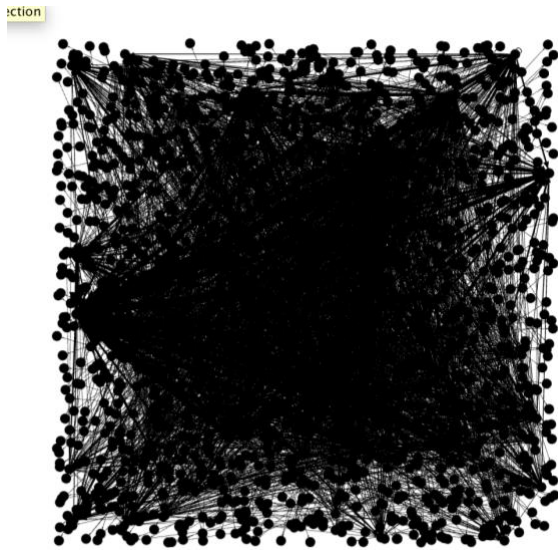


Figure 4 Before applying Yifan Algorithm

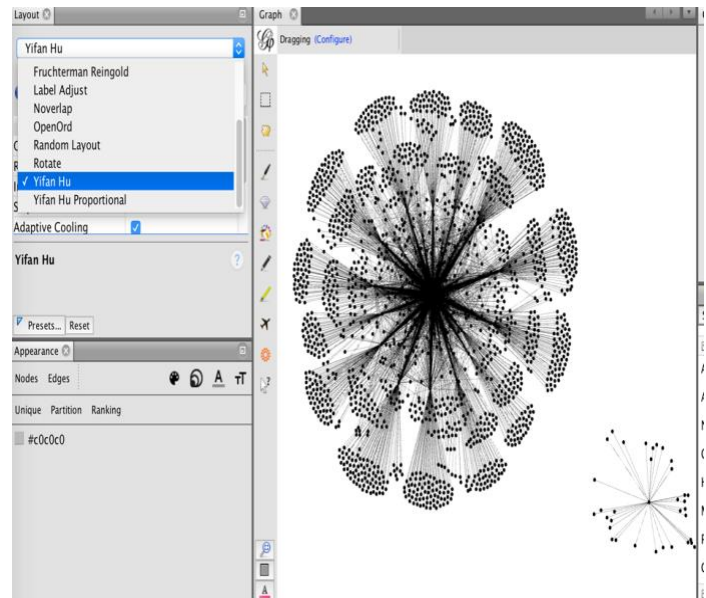


Figure 5 After applying Yifan Algorithm

4. Degree Range filter
To filter nodes that don't have connection out and keep only the important nodes (main cluster).

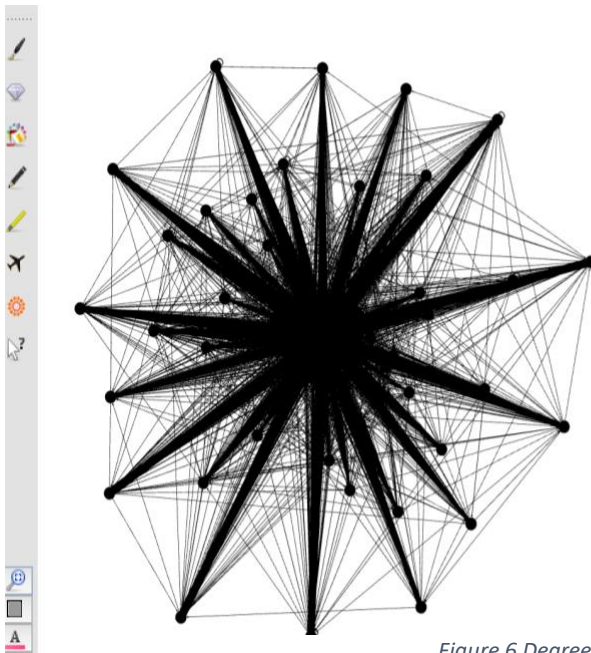
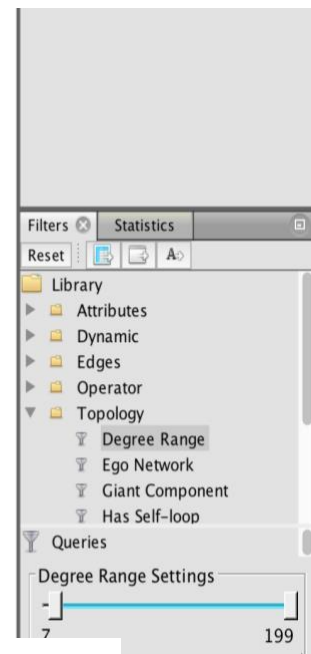


Figure 6 Degree Range Filter



5. Statistic

Network diameter: it's give us allots of centrality majors (Betweenness centrality, Closeness centrality, Eccentricity)

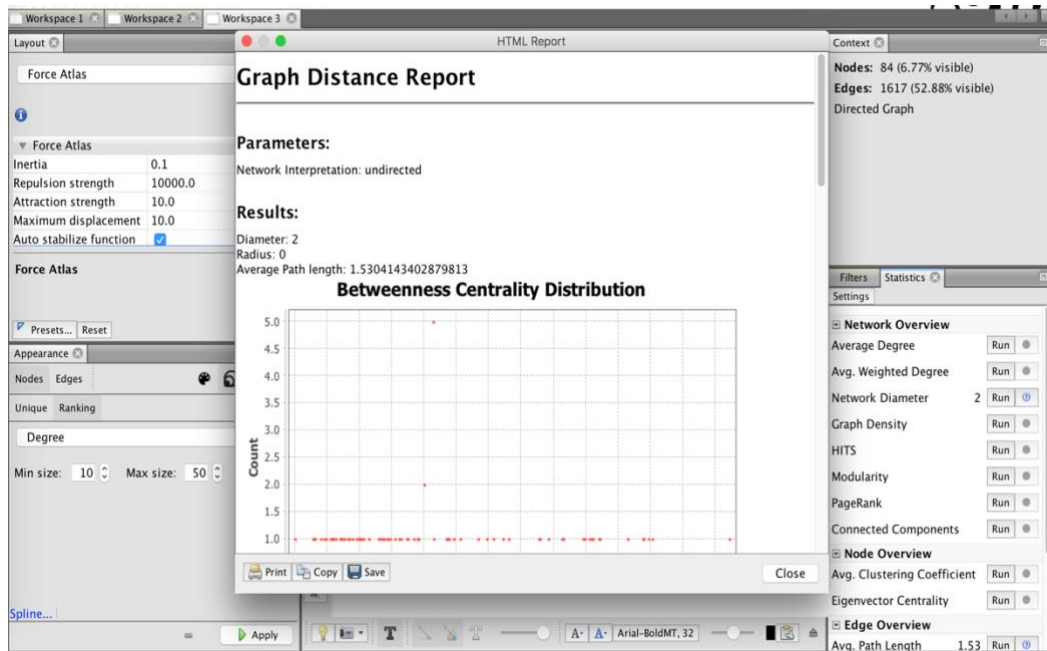


Figure 7 Graph report for Network Diameter

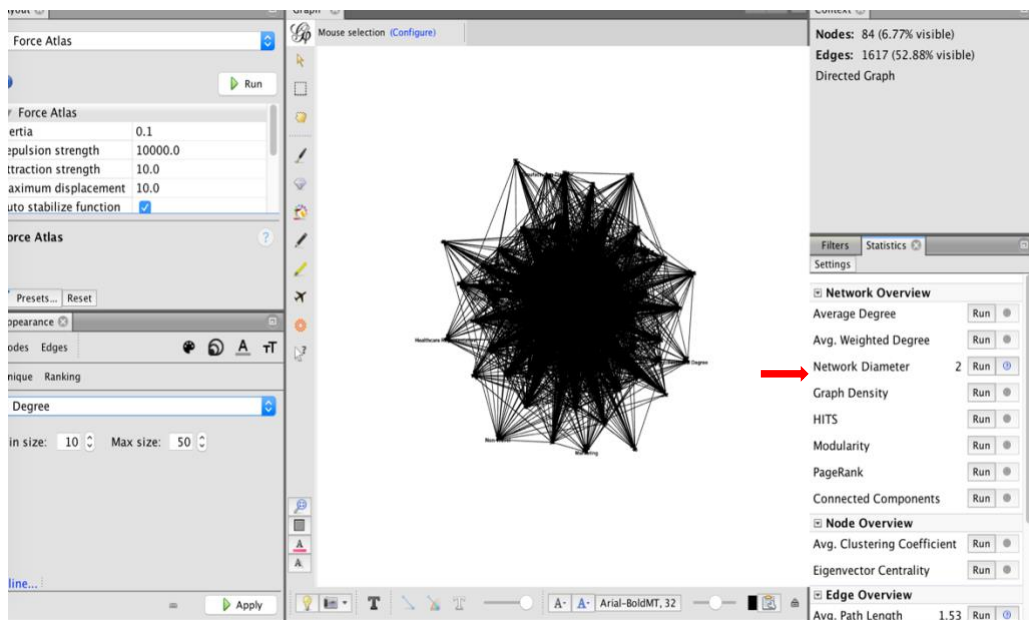


Figure 8 Network Diameter equal 2

6. Set size based on attribute (Betweenness centrality)

Big nodes in the picture appear the most central nodes in the network.

- Minimum value (10) is the smallest betweenness
- Maximum value (50) is the highest betweenness
- Medium value is somewhere between 10-50

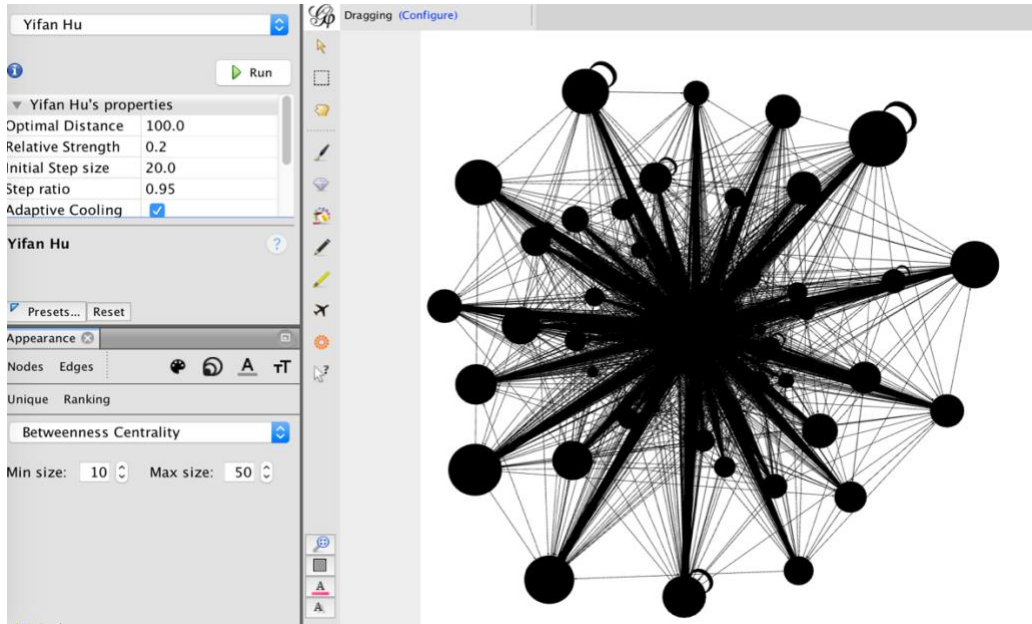


Figure 9 Betweenness Centrality

7. Set color based on attribute (closeness centrality)

- Red color: shows the lower closeness centrality
- Yellow color: shows the middle closeness centrality
- Green color: shows the higher closeness centrality

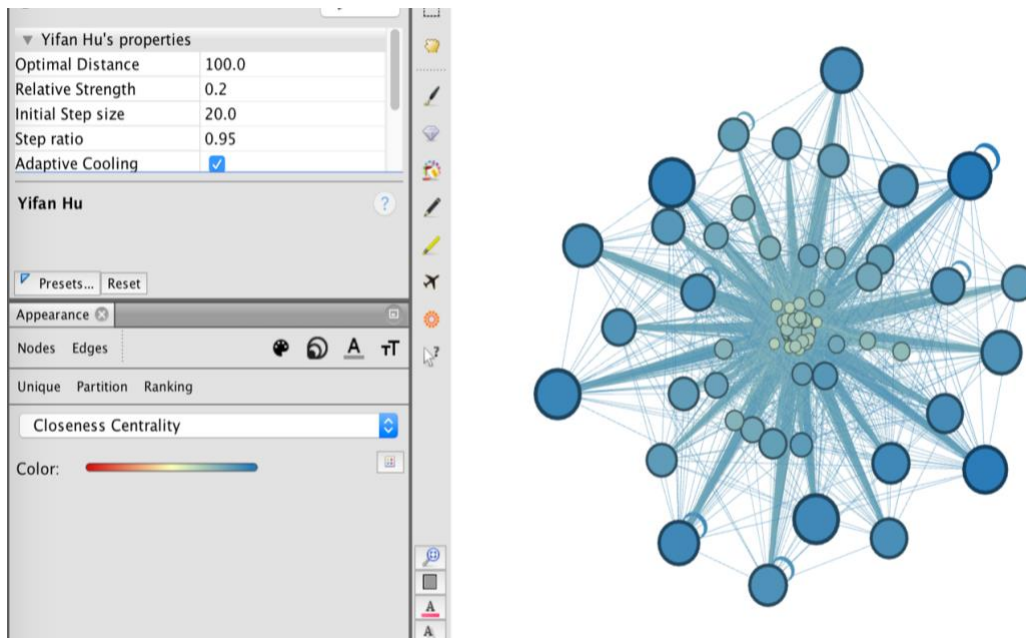


Figure 10 Closeness Centrality

8. Statistic (The average length between edges)

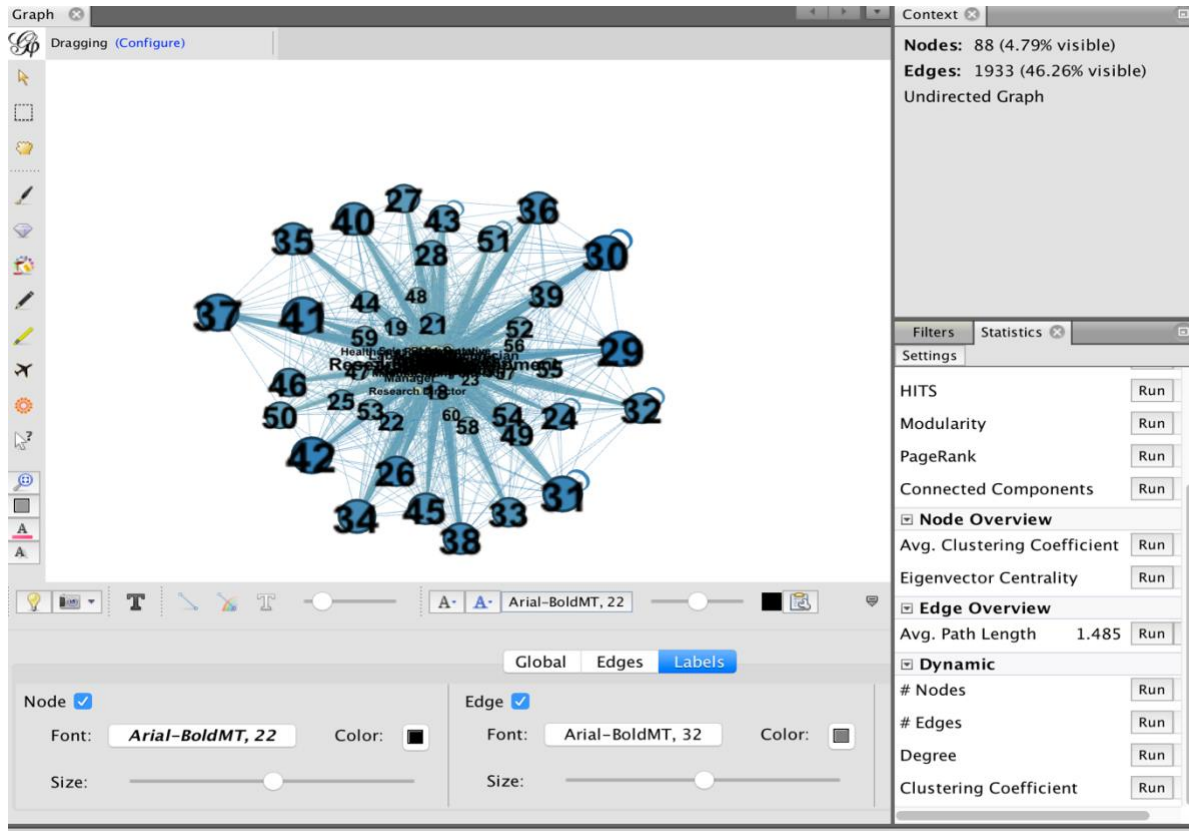


Figure 11 Edge overview

9. Set color based on attribute (Degree)

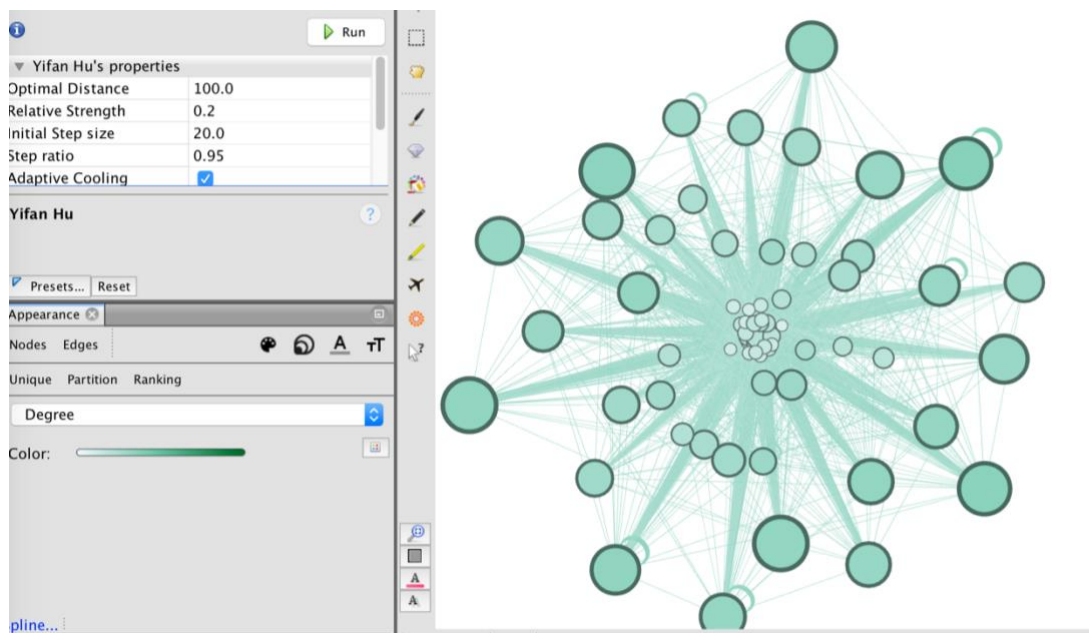


Figure 12 Degree

10. Show up the label on each node

Choose “Node size” it’s selected the size of the text to match the size of the nodes.

The big nodes represent the most hourly rates for employees which is 42, 29 and 37.

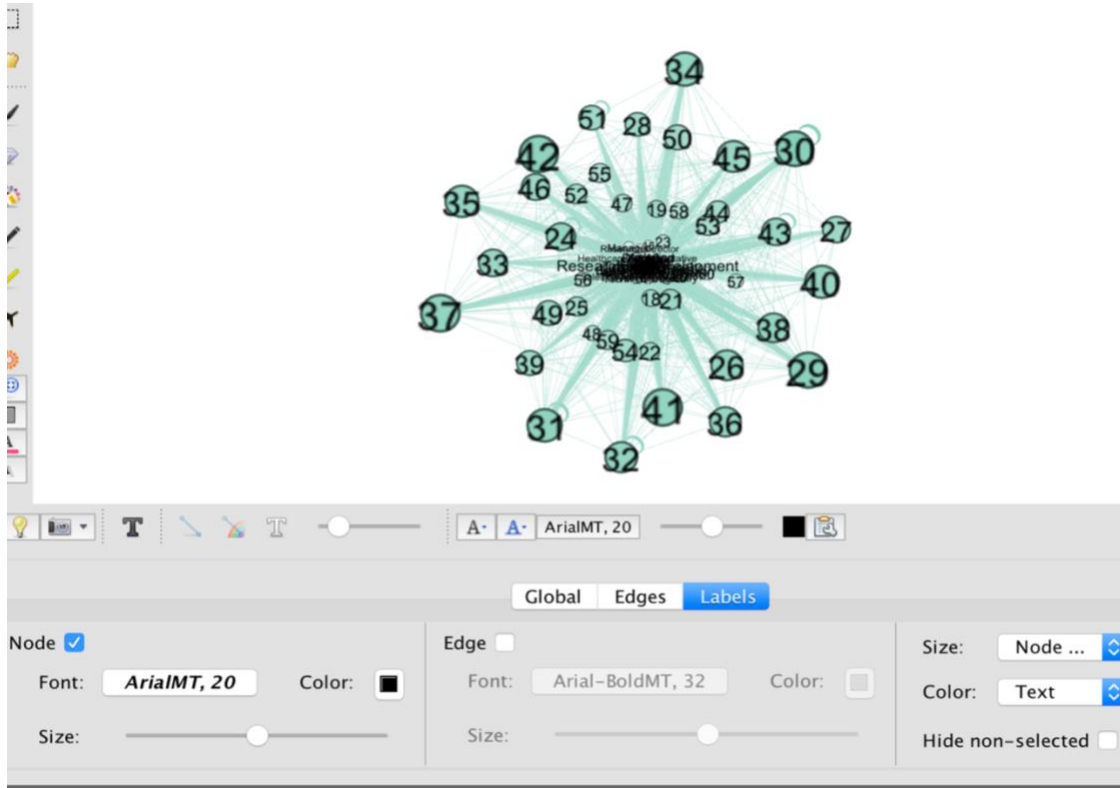


Figure 13 Label

As mentioned earlier the company wants to determine the factor that keep employees at the company. Some of these factors are

- Hourly rate
- Job involvement
- Monthly income
- Overtime
- Job satisfaction
- Years at company

Big nodes represent employees hourly rate. One of the main factors that lead employees to leave company is that employee who work overtime and not rewarded like increase the hourly rate.