A Social Network Analysis approach on the Placement of Engineering graduates using 2-mode Bipartite graphs

Pretty Liju George
Faculty, Department of Information Technology
Believer's Church Caarmel Engineering College

Abstract— **In the last couple of decades, Engineering Education has gained much appreciation in India and has become the 'de-facto' graduate degree of every student. And as a result a time has reached where there is surplus Engineering graduates and they are experiencing a joblessness. When every other person became and Engineer, the demand for engineering graduates have dramatically diminished leading to unemployment.**

Campus placements, a green channel of fresh Indian engineering graduate, for getting a job, is noticed to be concentrated in specific parts of the country, and particularly in certain cities. Based on previous statistical studies along with data that is collected, certain hypothesis could be formulated regarding the placement trend of few companies. The Network analysis approach to this issue will help to get a technical and statistical analysis on this theoretical issue.

This paper tries to uncover the unpleasant statistics and the possible reasons that could have led to such a biased placement pattern across the country by using the concept of Social network analysis and graph theories.

Keywords— Placements, Engineering graduates, Engineering Colleges, Multinational Companies, Recruitment, Employability, SNA, Bipartite graph, Density, Degree

I. INTRODUCTION

The number of engineering colleges in India while as of year 1996, took a shoot from a count as small as 416 to a whooping figure of 4472 Institutes till date [1].
The number of engineering colleges in India while as of year 1996, took a shoot from a count as small as 416 to a whooping figure of 4472 Institutes till date [1].
Aspiring Minds, an employability solutions company, performs scientific assessments on statistics of recruitment that took place across the country. Their latest report showed some interesting facts about the employability of engineering graduates in India [2][3].

Few highlights of their report that is of our concern,

- Total of 38,52,014 students admitted for engineering programs this year, in a total of 10327 colleges in India
- Around 1.5 million graduates released to the job market every year
- States like Delhi, Bihar, Jharkhand shows highest employability having fewer number of colleges.

- High Correlation between number of colleges and Employability of that state. States with higher number of colleges like Tamil Nadu, Andhra Pradesh and Maharashtra shows the least Employability in their graduates.
- Government Colleges shows employability 35-50% higher than Self-financing colleges.
- 55% of the Employable pool lies in the lesser known colleges and are being missed by Corporates
- A student from Tier 3 College has 24% less chance of getting a job than an equally meritorious student of Tier 1 College.
- In 2016, Kerala rose to the first 25 percentile bin (top quartile) from the third quartile, which shows remarkable increase in the employability of their graduates.
- Institutions in the cities shows better employability, which could be due to a better exposure to communication and technical skills from an early age.
- IT sector carries out the highest number of recruitments from the pool of Engineers and ironically only 20% of them are employable for Software jobs

State	Institut-ions	Total Intake	Employabil-ity Percentile Bin
Kerala	208	65963	I
Karnataka	268	109493	II
Andhra Pradesh	456	194460	III
Telangana	447	180583	III
Tamil Nadu	574	288717	IV
Delhi	25	10080	I
Bihar	28	9080	I
Jharkhand	18	7545	I

Table 1: Employability Percentile of Various states in India

Table 1 shows the number of institutions, their total intake and which employability percentile bin few southern states and the top states belong to. Quarter I is the bin containing states with highest employability and Quarter IV the lowest.

II. MATTER OF CONCERN

It is interesting to note that the average intake of students into the colleges in Kerala, is too low, that the seats are not fully capacitated. This could be primarily because of the new technical university (KTU) undertaking of all engineering colleges in the state. The low placement history could also be an added factor in the low intake of students.

Placement in Kerala is very low that a large portion of students from this employable pool does not get an opportunity to even attend a placement drive. Meritorious students belonging to Tier 3 institutions doesn't even get noticed pass as average student of a Tier 1 institute. Corporates are stuck with their regular tie-up institutions majority of them belonging to the northern cities, and they end up missing the talented and skillful graduates from down south.

On the contrary, state like Tamil Nadu with the highest number of colleges and annual intakes produce least number of Employable graduates. Reason could be, the compromised quality of education and leniency in examination system.

The governing bodies should limit the number of institutions and get more focused in improving the quality of the already existing institutions. There should also be an unbiased and distributed system for placements across the country, giving meritorious students of even a Tier 3 college an equal opportunity into the job market as a student from the Tier 1 College.

III. BACKGROUND COMPUTATIONAL STRATEGY- SOCIAL NETWORK ANALYSIS

A network is a structure of nodes inter-connected using edges. Nodes can be anything varying from molecules, bus stops, computers, places to people. Edges connect two nodes representing a relationship between them. Network analysis is performed on various networks of computer or connecting equipment to measure the shortest path or even the traffic.

When the network referring to are dealing with people, it is called as Social network analysis. The edges could be the existing relationship between them. As defined by [4] Social network analysis is an approach and set of techniques used to study the exchange of resources among actors (i.e., individuals, groups, or organizations). It is a relational study of social structures which is applied to variety of substantial problems. Social Network Analysis is widely used in behavioral science, organizational science, statistics, criminology etc.[5][9].

Resources could be anything that flows between the two actors such as information, people, money etc. Regular patterns of resources exchange between the actors brings out the various characteristics of the network helping us to get to some conclusions. Social network analysis assesses information opportunities for individuals or groups of individuals in terms of exposure to and control of information. Link prediction analysis can be performed by mining patterns on the existing graph to recommend and predict better and futuristic edges.

To understand a network in its fullness, visual representations are necessary. Visualization generates a qualitative and quantitative interpretation of raw data. Network analysis tools with visualization capabilities such as Gephi, NetMiner, iGraph, Cytoscape, NetworkX are used to create meaningful images of high quality to bring out the best of information and interpretation of the network.

We require a set of concepts and analytical tools by statistical methods and graph theory to retrieve interesting information from a graph. SNA software generate these characteristics from raw network data transformed to an edge list, adjacency list, or adjacency matrix (also called Sociomatrix), along with nodal attributes. Though the majority of network analysis software uses a plain text ASCII data format, some software packages contain the capability to utilize relational databases to import and/or store network features.

Generally, a network is represented as G(V,E), where V is the vertex (node) and E is the edge (ties). Connections between the vertices can be represented by adjacency matrix A, where

$$\mathbf{A_{ij}!=0, (V_i, V_j)\ \varepsilon\ E.}$$

If Aij =0, (V_i,V_j) ε E, which implies there exist no ties between the vertex i and j [4].

The prime measures of centrality of a graph is Degree, Closeness, Betweenness and Eigen vector apart from other performance measures like robustness, effectiveness and diversity[7][11].

- Degree centrality of a node i is the number of edges incident on it and is represented as

$$d(i) = \sum_j m_{ij}$$

  where m_{ij} = 1 if there is a link between nodes i and j, is no such link.
- Number of vertex |V|=n and number of edges |E|=m.
- Shortest path between two nodes V_i and V_j is d(i,j).
- Neighbors of node V is N(V)

For Bipartite graphs, there exist two sets of vertices. It can also be considered as 2-mode data that is ties between two sets of entities, where mode is a class of entity [10][11].

Bipartite graph with two sets of nodes, U and V can be represented as G(U,V,E) , where E denoted the weight of the edge that connects a node of set U with a node of set V.

- Density of a bipartite graph is

  D= n1n2/ (n1+n2)(n1+n2-1)

  0 < D < 1, where 0 implies no links and 1 denoted all possible links present.

IV. RELATED WORKS (SNA IN EDUCATIONAL DOMAIN)

Many studies related to this topic has been carried out in the western countries. In the paper *"How Does the Image of Engineering Affect Student Recruitment and Retention?", H. Oner Yurtseven Of Indiana University* attempts to trace the image of an engineer throughout history and to describe how the modern day engineer is viewed in the USA. Some of the projects funded by various engineering societies and foundations in hopes of enhancing the image of engineering in the USA are also included in the paper.

"*The Correlation between Industrial Placements and Final Degree Results: A Study of Engineering Placement Students*" *Richard Mendez, University of Leicester* examined the efficacy of industrial work placements on improving academic performance amongst engineering students. Evidently, the findings illustrate a causal link between placements and improved academic performance. Ying Zhang et al. [17] studied the relationship between network centrality and academic performance among a group of PhD students of UNU-MERIT.

Sunil Mani & Arun M did a case study on the engineering education on Kerala in their work "Liberalization of technical education in Kerala: has a significant increase in enrolment translated into increase in supply of engineers?". They tried to point out that the actual outturn rates of engineering graduates have been steadily declining since 2004. They found out that many students who gain admission to engineering colleges do not have the basic capability, which can be built only by improving school education.

And further an extensive study by Suresh Kumar N et al. [16], tries to explore various parameters of teaching that plays a pivotal role in modeling the students at various stages of selection process. They came to a conclusion that it is certainly possible that one's personality and emotional temperament would influence one's academic abilities, and, regardless of the variations in language and classification, there is some evidence of an association between affective characteristics and academic performance.

Fabrice Coulon, Division of Innovation - LTH Lund University, Sweden "The use of Social Network Analysis in Innovation Research" I gained a clear outlook about the use of Network analysis in my problem. The review introduces network analysis then discusses why and how it has been used in innovation research so far. This paper argues that studies using social network analysis tend to focus too much on change in the relationships between interacting units or nodes of the network to the detriment of change within units/nodes. Therefore, a combination of case study and social network analysis can offer a solution to that problem by providing the best of both methodologies.

To represent a high school transcript data, a two-mode bipartite network data (Frank, K. A et al, 2008) was constructed using students as one set of data and courses available as another set. Edges were formed from a student to a course if that student enrolls for that course [6]. Analysis of two-mode network data, or bipartite graphs has set in a new trend in the social network analysis (see the May 2013 special issue of *Social Networks*, 35(2), pp. 145–278)

V. IMPLEMENTING SNA TO THIS PROBLEM

SNA is a study of relationship among a given set of entities. In my study, the entities are of two categories, firstly the different Engineering Institutions and second being the various Multinational Companies. Therefore the graph that would be created will be a bipartite 2-mode graph with connections only from nodes of one category to a node of the other category as shown in Figure 1

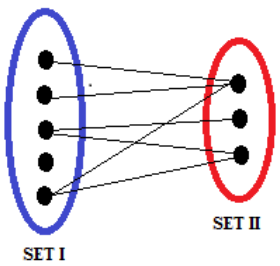

Figure 1. Incomplete 2-mode Bipartite Graph

Recruitment analysis of fresh graduates by multinational companies being the topic of our interest, this becomes the key relationship between the two sets. An edge will exist between a college and a company, if and only if a recruitment has taken place in that institution by that company. Each edge can be represented as a weighted link; the weight being the number of placements from that college. As several companies recruit students from a given college and at the same time students from several institutions are placed in a given company, we will be able to generate a 'many to many' network. The Edges can also be assigned weights based on the number of recruitment that happened between the company and the institution that

being connected by it. Yet, this network is an incomplete bipartite graph as not all companies recruit from every college. In fact, there are colleges which lacks any placement at all and companies that doesn't participate in campus placements. The graph can also be considered as a directed graph where flow happens from only one category to the other and not vice versa, that is from colleges to companies.

Shown in Fig.2 is a network generated from the raw placement statistics collected from few handpicked colleges in Kerala using a graph generating tool 'Gephi'. In order to visibly distinguish between the two categories, both are allotted different colors; red for colleges and blue for companies.

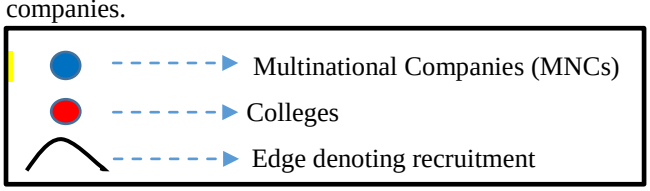

Fig 2. Network Entities

Several computation can be performed on this graphs to extract many facts and figures about the placement scenario of a certain district, sector, state or even a whole country as a whole.

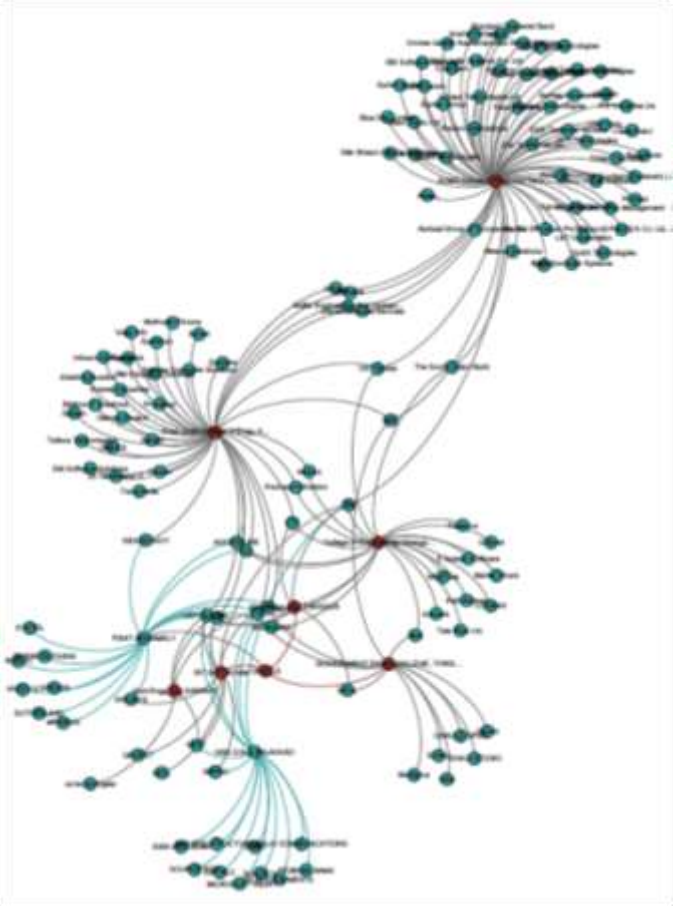

Fig 3. Sociogram of Placement on sample dataset. (Using Gephi)

VI. CONCLUSION

This paper is a review on Social network analysis concepts that have been implemented in the Education domain and highlighting some prominent graph theories that could be useful for this type of study. In addition to that a close look on the Engineering education of our country is made to bring into focus certain statistics that has become a matter of concern for us. It's been understood that increasing the number of Engineering colleges in states like Kerala, Tamil Nadu and Andhra Pradesh, can decrease the Employability of fresh graduates of that state. The placements are seen happening at particular colleges alone leaving aside a number of other colleges, depriving meritorious students of those college from being placed in good companies. Therefore an efficient n distributed measure has to be adopted for the recruitment process to provide an even and fair recruitment throughout the state and our country. This gives the privilege to all meritorious students to attend the recruitment process in spite of the institution or the place that he/she studies.

And moving onto the scientific side of the study, Social network analysis could bring out the graph pattern, the trendsetting colleges and companies and spot out the central area or college of the recruitment process. Using the various tools that are commercially available, several measures like degree centrality, density, cluster coefficient, link predictions etc., can be computed to derive certain conclusions on the Engineering graduate's placement in our country. So this paper should be considered as an attempt taken to kick start a new work that collaborates the Engineering Education domain, the Social network Analysis concepts and also basic Graph theory.

REFERENCES

[1] All India Council of Technical Education (2016), Report of the High Power Committee for Faculty Development in Technical Institutions, (Chairman: Prof.Anil D. Sahasrabudhe)
[2] Aspiring Minds (2014), Employability of Engineers State Wise, Excerpts from National Employability Study, Aspiring Minds Research Cell.
[3] Aspiring Minds (2011), Employability of Engineers State Wise, Excerpts from National Employability Study, Aspiring Minds Research Cell.
[4] Wasserman, S., & Faust, K. (1994). Social Network Analysis: Methods and Applications. Cambridge: Cambridge University Press
[5] M. Newman, A.-L. Barabasi, and D. J. Watts, editors. The Structure and Dynamics of Networks. 2006.
[6] Frank, K. A., Muller, C., Schiller, K., Riegle-Crumb, C., Strassman-Muller, A., Crosnoe, R., & Pearson J. (2008). The social dynamics of mathematics course taking in high school. American Journal of Sociology, 113(6), 1645–1696
[7] Fabrice Coulon, "The use of Social Network Analysis in Innovation Research: A literature review" Division of Innovation – LTH Lund University, Sweden. 2005
[8] Borgatti, S. P., & Everett, M. G. (2013). Analyzing Social Networks. New York, NY: Sage Publications
[9] Hanneman, R. A., & Riddle, M. (2005). "Introduction to social network methods", University of California Riverside.
[10] S. P. Borgatti and E. M. G., "Network analysis of 2-mode data," Social Networks, vol. 19, pp. 243-269, 1997.
[11] 2-Mode Concepts in Social Network Analysis, Stephen P. Borgatti
[12] G. Sabidussi, "The centrality index of a graph," Psychometrika, vol. 31, pp. 581-603, 1966.

[14] L. Freeman, Graphic techniques for exploring social network data. New York: Cambridge University Press, 2005, pages 248-269 in 'Models and methods in social network analysis' edited by P.J. Carrington, J. Scott, and S. Wasserman.

[15] Mendez, R. 2008. The Correlation between Industrial Placements and Final Degree Results: A Study of Engineering Placement Students. Paper presented at ASET Conference, September 2-4, in Plymouth.

[16] Suresh Kumar N , Prasanth MK, Ajith Sundaram , "Campus placements in Kerala-An empirical study at the selected Engineering Colleges in Kerala" , International Journal of Scientific and Research Publications, Volume 3, Issue 1, January 2013

[17] Ying Zhang, Iman Rajabzadeh, and Rodolfo Lauterbach, "Student Network Centrality and Academic Performance", Evidence from United Nations University, 2009