Link Analysis on Institutional Repository web Network of Indian Institute of Technologies Registered in open DOAR-uncovering Patterns and Trends Hidden in the Network

Kutty Kumar*

ARTICLE INFO

Article history:
Received 03 April 2018
Revised 25 April 2018
Accepted 12 May 2018

Keywords:
IIT,
Website,
Institutional Repository,
Link Analysis,
Network

ABSTRACT

Institutional repositories (IR) are promising to be extremely advantageous to scholars especially in developing countries. IR initiatives started in India during the late nineties and the popularity of this concept is growing rapidly in the higher educational and research institutions to disseminate newly emerging knowledge and expertise. The purpose of this paper is to critically analyze the network links of IR websites among four IITs that are registered in open DOAR (Directory of Open Access Repositories) web portal. The Institutional Repositories chosen for the study are IIT Delhi, IIT Hyderabad, IIT Bombay, and IIT Kanpur. The analysis of the study focused on standard graph and network cohesion metrics, such as density, diameter, eccentricity and distances, and clustering coefficient; for an even more detailed analysis advanced centrality measures and fast algorithms such as clique census are used.

1. Introduction

Evolution of institutional repositories had essentially changed the landscape of academic and scholarly publishing, impelling patrons in graduate programs to reexamine historic assumptions about thesis and dissertation management and distribution that were developed in the age of print and microfilm (Clement, & Rascoe, 2013). Over the decades, many Institutional Repositories have been launched by Indian Universities. In this framework, the Indian Institute of Technologies has designed their Institutional Repositories containing numerous digital libraries each with varying numbers of ETD collections and they all have registered either with the Open DOAR (Directory of Open access Repositories) or ROAR (Registry of Open Access Repositories). IITs is on track digitizing their back volumes of thesis and dissertation collections on project basis and make them available through open access or campus wide institutional repositories.

The increasing use of sophisticated visualizations is probably the most significant development in relational altmetrics and has led to the creation of a new field: knowledge domain visualization,

* Assistant Professor, Library and Information Science, College of Veterinary Science, Sri Venkateswara Veterinary University, India (kumarkkutty@gmail.com)
International Journal of Knowledge Content Development & Technology, 8(2): 23-36, 2018.
http://dx.doi.org/10.5865/IJKCT.2018.8.2.023

within the information visualization research area (Chen, 2006). Analyzing the links among the network is used to evaluate the relationships or connections between network nodes. These relationships can be between various types of objects (nodes), including people, organizations and even transactions. The contribution of Indian Institutes of Technologies has massively helped flourish the Indian socio-economic and cultural sectors. These institutions are active in producing scholarly literature in the form of research projects, research papers, in-house publications, etc. But such intellectual literature/outputs are not fully available in the public domain although they are public-funded establishments. Identifying this gap, institutional repositories (IRs) an online platform for archiving and sharing of institutional intellectual-content, initiated by the IITs were reconnoitered (Hulagabali, 2015). The purpose of this paper is to visualize and analyze the network links of IR among four Indian Institutes of Technologies (IIT) that are listed in DOAR (Directory of Open Access Repositories) viz. IIT-Bombay, IIT-Delhi, IIT-Hyderabad, and IIT-Kanpur.

2. Literature review

Link analysis is essentially a kind of knowledge discovery used to visualize data to allow better analysis; especially in the context of links, network analysis has become a widely applied method in research and business for inquiring into the web of relationships on the individual and organizational level (Thellwal, 2008). As stated by Börner (2010) network maps, comprised of nodes and links between these nodes, have come to dominate the landscape. Yang, Liu, and Meloche (2010) had explored the link relationship between websites in mainland China use principal component, multi-dimensional scale and social network analysis methods, such as a co-link, Betweenness and k-core. The study detailed on websites that are typically linked to their influence at a greater level, whereas subject/regional relations are discovered through sub-network analysis at the micro-level.

Ortega and Aguillo (2008) aimed to study the link relationships in the Nordic academic web space comprised of 23 Finnish, 11 Danish and 28 Swedish academic web domains with the European one. Through social network analysis the authors' attempted to detect sub-networks within the Nordic network, the position and role of the different university web domains and to understand the structural topology of this web space. Kraker et al. (2015) analyzed the adequacy and applicability of readership statistics recorded in social reference management systems for creating knowledge domain visualizations. The authors investigated the distribution of subject areas in user libraries of educational technology researchers on Mendeley. The study also used co-readership patterns to map the field of educational technology.

Based on a case study conducted among a sample of Swiss management scholars, Hoffmann, Lutz and Meckel (2016) analyzed how centrality measures derived from the participants' interactions with the academic SNS Research- Gate related to traditional, offline impact indicators. The authors found that platform engagement, seniority, and publication impact contributed to members' in degree and eigenvector centrality on the platform, but less so to closeness or Betweenness centrality. They concluded that a relational approach based on social network analyses of academic SNS, while subjected to platform-specific dynamics, could add richness and differentiation to scientific impact assessment. Backstrom et al. (2012) reported the results of the first world scale social network graph distance computations, using the entire Facebook network of active users. The investigators studied the distance distribution of Facebook and of some interesting geographic sub graphs, looking also at their evolution over time. Edward, Gelaw, & Reyes analyzed URL references in 4,335 ETDs at the University of North Texas ETD collection. Further the study provided a preliminary framework for technical methods appropriate for the approaching analysis of similar data that might be applicable to other sets of documents or subject areas. Sterman (2014) sought to give libraries a plan for inter institutional cooperation for institutional repositories that will benefit all involved: researchers, institutions, and, ultimately, global scholarship. This research used repository studies, interviews with existing repository managers, and the input of libraries considering a repository to inform the exploration of the opportunities for collaboration in IR development and maintenance.

2.1 Objectives

Gain an intangible understanding of

- ➢ Visualizing Network data of IIT’s repositories under study
- ➢ Characteristics of network link relationships

3. Limitation

The uncertainty in digital evidence is not being evaluated at present, thus making it difficult to assess the reliability of evidence stored on and transmitted using computer networks (Saint-Charles & Mongeau, 2009). Even though, five IIT institutional Repositories (IIT Delhi, IIT Bombay, IIT Kanpur, IIT Hyderabad and IIT Roorkey) are registered in DOAR, only for institutional Repositories had an active URL, and that of IIT Roorkey was found nonfunctional at the time of data collection, hence the later excluded from the study.

4. Materials and Methods

Graphical representation of Institutional Repositories network is a prevailing approach, to understand and analyze the behavior of both the individual website and the overall network. This study contributes to the state of knowledge on Institutional Repositories of four IIT by focusing on the nature of network analysis in order to explore connectivity including the strength/weakness of connections, the distances, and the robustness of the connections, among other properties. The sample lists were browsed from Directory of Open Access Repositories (http://www.opendoar.org/). Data collected using Social Network Visualizer (SocNetV) licensed under the GNU General Public License 3 (GPL3). SocNetV includes a simple web crawler, which consists of two parts: a spider and a parser. The spider visits a given initial URL (i.e. A website or a web page) and downloads its HTML code. The parser scans the code for 'href' links to other pages (internal or external) and adds them to a queue of URLs (called frontier). As URLs are added in the queue, the spider visits them and downloads their HTML which is scanned for more links by the parser, and so on... The end result is the 'network' of all visited webpages as nodes and their real links as edges. The nodes/vertices in the graph represent the related websites of the network and the relationship between them is represented as edges/links. By default the spider will crawl both internal and external links constituting a maximum of 600 nodes or pages. Hence, for this study 600 nodes are considered. Every node and each of their corresponding links carries certain characteristics. As described by Dincer (2018) each node represents an entity, while every link carries attributes that define the nature of the relationship.

5. Results and Discussion

A Visual graph is an abstract representation of the same set of information contained within the adjacency matrix. In a Visual graph:

- ➢ A vertex represents the web link of IIT - IR network.
- ➢ An edge represents a link between IIT - IR web networks.
- ➢ Link indicates the "strength" of the connections, the frequency of the interaction and the intensity of the relationship.

In this study, four centrality measures are compared: Degree centrality (based on degree), Closeness centrality (based on average distances) and Betweenness centrality (based on geodesics). Table 1 provides the sample list taken for analysis.

Table 1. Institutional Repositories under study

S.No	Intuitional Repository	Established Year	Software Used	URL
1	IIT - Delhi	2005	Eprints	http://eprint.iitd.ac.in/
2	IIT - Bombay	1999	DSpace	http://dspace.library.iitb.ac.in/jspui/
3	IIT - Hyderabad	2015	Eprints	http://raiith.iith.ac.in/
4	IIT - Kanpur	2005	Agropedia	http://agropedialabs.iitk.ac.in/openaccess/

5.1 IIT-Delhi (IITD)

Indian Institute of Technology, Delhi is a public research university located in Delhi, India. It was declared to be Institute of National Importance by Government of India under the Institutes of Technology Act. IIT Delhi is one of the two educational institutes in India, which have been listed in Quacquarelli Symonds' (QS) list of top 200 universities globally in 2015.The Institutional Repository eprints@IIT Delhi is registered as e-theses repository in DOAR. Link nodes of the IIT Delhi portal presented in Fig. 1. The repository also has the ability to capture, index, and store, disseminate and preserve digital materials created in any part of the Institute. Faculty and researchers can register themselves with the digital repository and submit their pre-prints (pre-refereed version of an article), post-prints (post-refereed final version) and publisher PDFs (if allowed by the publisher).

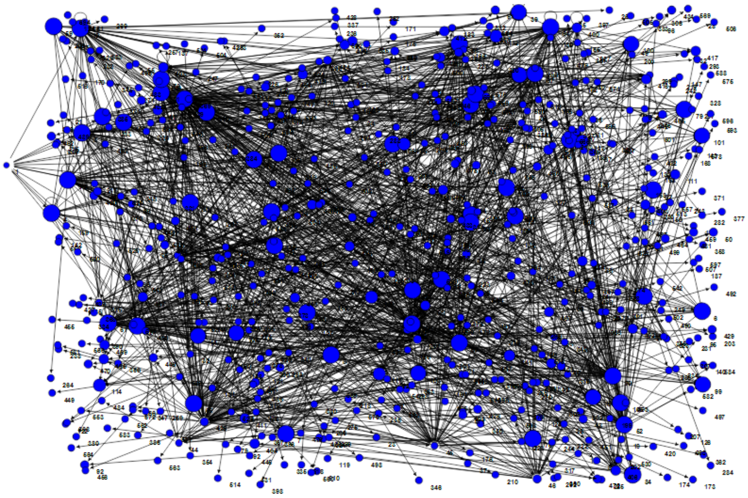

Fig. 1. Link Nodes of IIT Delhi IR Portal
(http://eprint.iitd.ac.in/)

5.2 IIT Bombay (IITB)

The Institutional repository of IIT Bombay is a research produced at IITB. The repository includes full-text of book chapters, conference/proceeding papers, technical reports, journal pre-prints & post-prints, working papers, Patents and others like annual reports. Link nodes of IITB -IR illustrated in Fig. 2.

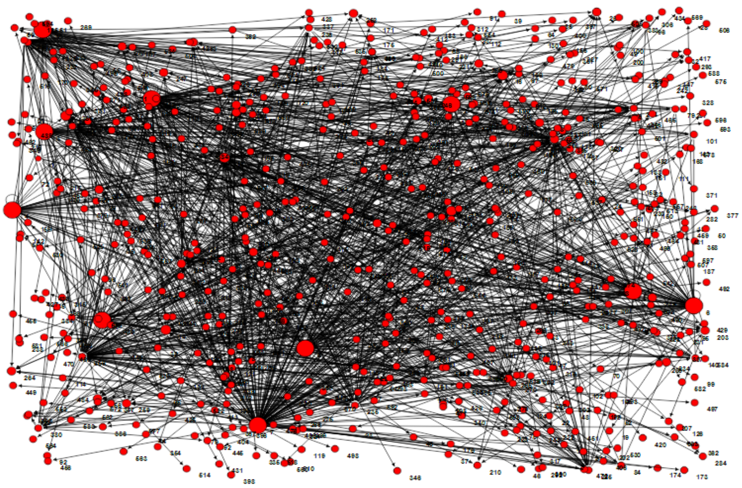

Fig. 2. Link Nodes of IIT Bombay IR Portal
(http://dspace.library.iitb.ac.in/jspui/)

5.3 IIT Hyderabad (IITH)

The Indian Institute of Technology, Hyderabad, (IITH) launched its Research Archives of Indian Institute of Technology, Hyderabad (RAIITH) (http://eprints.iith.ac.in) in the year 2015 (Kimidi, Mallikarjuna, & Asthana, 2017),This repository contains all the scholarly content authored by the academic community of the IITH and the Library team works with departments, individual faculty, and students to select, submit and preserve their 'intellectual output for long-term preservation and worldwide electronic accessibility'. Fig. 3 denotes a web portal link of RAIITH.

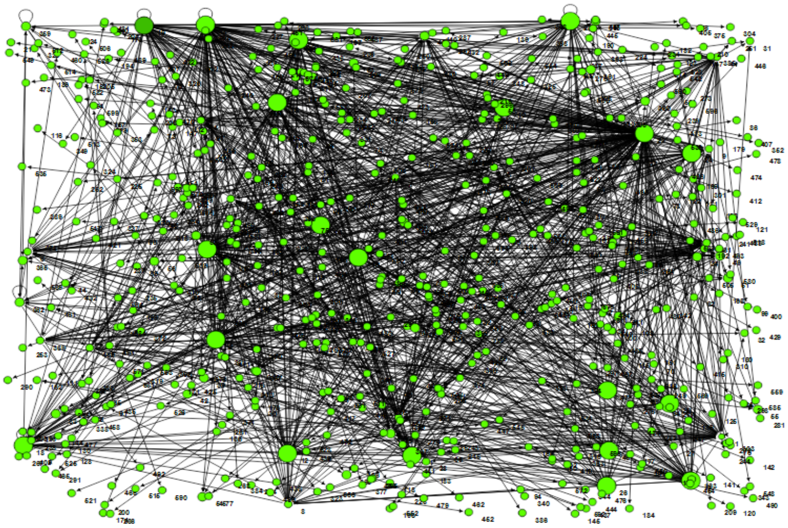

Fig. 3. Link Nodes of RAIITH Portal
(http://raiith.iith.ac.in/)

5.4 IIT Kanpur (IITK)

Since the Indian Institute of Technology, Kanpur had started its research programs from 1963 onwards, it had over 9000 M.Tech and pH. D theses in different areas of Science, Engineering, and Humanities & Social Sciences. Openagri (Open Access Agricultural Research Repository) sponsored ICAR; NAIP Agropedia is a digital knowledge repository with the open platform for learning and sharing information related to Indian agriculture. The contents are semantically catalogued and easy to find. This agricultural encyclopedia is being designed as a sub project of the knowledge management initiative of National Agricultural Innovation Project (www.naip.icar.org.in) in support of agricultural extension and outreach. Fig. 4 provide link nodes of the IIT Kanpur IR portal.

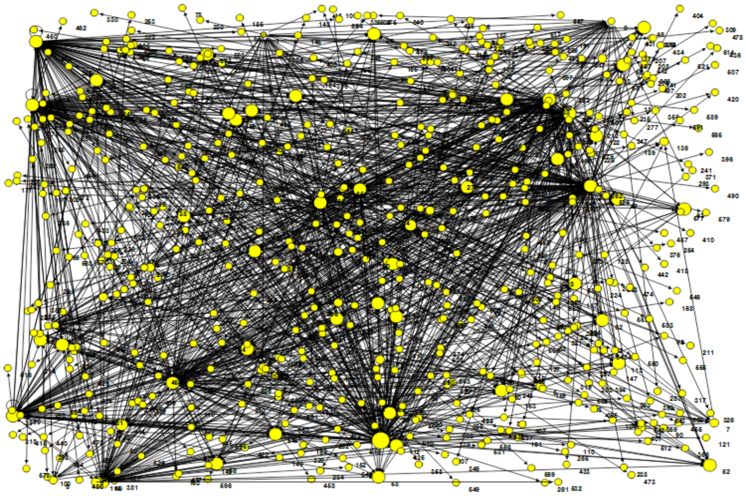

Fig. 4. Link Nodes of IIT Kanpur IR Portal
(http://agropedialabs.iitk.ac.in/openaccess/)

The initial point to begin in network analysis is at the global level, looking at properties that describe the entire network. For example, the counts associated with the graph provides some insight into its "density" the number of possible connections vs. the number of actually present connections. The larger the network is, the more connections are needed to effectively propagate information, which is why distance and its density are being explored. Table 2 provides data on link analysis, metrics of IIT repositories under study.

Table 2. Comparative data of Institutional Repositories in Link Analysis

S.No	Link Analysis Metrics	Institutions			
		IIT –Delhi	IIT- Bombay	IIT-Hyderabad	IIT-Kanpur
1	Average Distance	0.0919672	0.188297	0.140626	0.0829188
2	Diameter	6	9	8	4
3	Network Density	0.00386	0.00284	0.0027	0.00332
4	Total Edges (Arcs)	1386	1019	971	1193
5	Node Out Degree	15	10	47	30

5.5 Average Distance

The Average distance is the mean length of the shortest path between two vertices in a network. Nodes that are connected at short lengths or distances may have stronger connections. (Yan, Ding, & Zhu, 2010). Shortest paths in a network are the most important; the shorter the path from one network to the other, the quicker and more efficient the flow of information, advice, and knowledge. The more central a node is, the lower its total distance to all other nodes. In the study sample, it is found that IR websites of IIT Kanpur (0.08), and IIT Delhi (0.09) had least average distance between linked websites, while IIT Bombay had a longer path between nodes; however the network with shorter path are preferred concerning to speed of information exchange.

5.6 Diameter

The longest shortest path (distance) between any two nodes in a network is called the network's diameter it is a useful measure of the reach of the network (Giorgos & Cheliotis, 2017). It also indicates how long it will take at most, to reach any node in the network (sparser networks will generally have greater diameters), It is vivid from the table that IR of IIT Delhi (6) and IIT Kanpur (4) have a denser network among the study group; however the diameter of a graph is the length of the longest chain from one vertex to another in that graph.

5.7 Network Density

Network Density is the ratio of links to the number of possible links. Network density is a measure of the connectedness in a network. It is a number that varies between 0 and 1.0. When density is close to 1.0, the network is said to be dense, otherwise it is sparse. When dealing with direct ties, the maximum possible number of pairs is used instead. From the report presented in the Table 2, it is found that almost all the IIT - IR web network is less densely connected since their network density are not close to 1.

5.8 Degree

The degree of a node describes its level of connectedness in the network. The degree of a vertex is the number of edges incident to it. So, the degree of a node is the number of direct connections the node has in a network. The nodes with higher out degree is more central (choices made, In general, a vertex with higher degree is more connected compared to a vertex with lower degree, and thus the graph can in general resist more link/vertex failures. Edges can represent interactions, flows of information in the network. From the data, it could be noted that IIT, Hyderabad has more node out degree (47) with 971 edges, followed by IIT Kanpur (node, out degree: 30 with 1193 edges) indicating these IR websites play a predominant role in connecting the network among the IR under study. It is significant to note that, IIT Delhi has more number of edges (1386) but node out degree is merely 15, meaning the repository website is less central.

5.9 Aggregate Counts of Cliques

The clique algorithm produces an output that can be quickly interpreted by a human who is versed in the field. A (maximal) clique is a maximal set of mutually adjacent vertices (Makino & Uno, 2004). Cliques are important for their role as cohesive subgroups, but show up in many other contexts as well (Wasserman & Faust, 1994). Aggregate count of cliques under study in

Table 3. It is observed from the data that IIT Kanpur and IIT Delhi had maximum number of cliques implying that links are closely clustered.

Table 3. Aggregate Counts of Cliques

S.No	Institutions	2-Vertex cliques (max: 179700)	3-Vertex cliques (max: 35820200)	4-Vertex cliques (max: -22544270)
1	IIT Delhi	84	170	265
2	IIT Bombay	47	45	45
3	IIT Hyderabad	26	30	20
4	IIT Kanpur	74	220	495

5.10 Centrality measures

Centralization provides a measure of the extent to which a whole network has a centralized structure. Whereas centralization describes the extent to which this connectedness is organized around particular focal nodes; density describes the general level of connectedness in a network. Centralization and density, therefore, are important complementary pair measures. While a centralized network will have many of its links dispersed around one or a few nodes, the decentralized network is one in which there is little variation between the number of links each node possesses. Table 4 provides information on various centrality measures for the repositories under study.

Table 4. Centrality Measures

S.No	Metrics	IIT Bombay-Node		IIT-Delhi-Node		IIT-Hyderabad-Node		IIT-Kanpur-Node	
		Max	Min	Max	Min	Max	Min	Max	Min
1	Eccentricity	9	0	6	0	8	0	4	0
2	Betweenness Centrality	0.017466	0	0.018448	0	0.017641	0	0.017034	0
3	Degrees Centrality	0.1803	0	0.1419	0	0.37229	0	0.20701	0
4	Degree Prestige	0.048414	0	0.040067	0	0.023372	0	0.035058	0
6	Average local Cluster Coefficient	1	0	1	0	0.83333	0	0.91667	0
7	Influence Closeness Centrality (IRCC)	0.39122	0	0.2915	0	0.37229	0	0.51377	0

5.11 Eccentricity (e)

Conceptually, the number of nodes traversed to establish the connection between two websites. Distances can be scaled based on the size of the network. Geodesic distance is used to characterize the most efficient or optimal connection between two nodes. The largest geodesic distance for each node is called its “eccentricity”. The eccentricity “e” of a node is the maximum geodesic distance from that node to all other nodes in the network. Dense networks have mostly short geodesic distances. Therefore, “e” reflects farness: how far, at most, is each node from every other node. A node has maximum e when it has distance 1 to all other nodes (star node). In this study, node 9 (http://dspace.library.iitb.ac.in/jspui/subscribe) of IIT Bombay (e = 9), node 1 (http://eprint.iitd.ac.in//) of IIT Delhi (Max e = 6), node 3 (http://raiith.iith.ac.in//) of IIT Hyderabad (Max e = 8) and node 1 (http://agropedialabs.iitk.ac.in/openaccess//) of IIT Kanpur (Max e = 4) have maximum eccentricity indicating these web links located far from the network. On the other hand, node 17 (http://dspace.library.iitb.ac.in/jspui/browse) of IIT Bombay, node 2 (http://eprint.iitd.ac.in/) of IIT Delhi, node 2 (http://raiith.iith.ac.in/) of IIT Hyderabad and node 11 (http://agropedialabs.iitk.ac.in/openaccess//) of IIT Kanpur had a minimum eccentricity of 0, which indicate these IR websites possess efficient connections between links.

5.12 Betweenness centrality

It is defined as the extent to which a node lies between other nodes in the network. In other words, Betweenness centrality is a measure of the degree to which a given node lies on the shortest paths (geodesics) between the other nodes in the network Collins, Bradley, and Yassine (2010). If all paths have to go through the node the number is 1, if there is always an alternative path the number is 0. Here, the connectivity of the IIT IR websites is taken into account in order to provide a higher value for nodes which bridge clusters. This metric reflects the number of links that are connected indirectly through direct links. (Lecture 4). The generated values of study indicated that node 13 (0.017) of IIT Bombay (http://www.dspace.org/), node 7 (0.018) of IIT Delhi (http://eprint.iitd.ac.in/help/index.html), Node 69 (0.018) of IIT Hyderabad (http://raiith.iith.ac.in/cgi/search/advanced)and node 3 (0.017) of IIT Kanpur (http://agropedialabs.iitk.ac.in/openaccess//) had highest betweenness, meaning they have higher inuence on the connectivity of their respective IR network .node 1 of IIT Bombay (http://dspace.library.iitb.ac.in/jspui//), IIT Delhi (http://eprint.iitd.ac.in//), IIT Hyderabad (http://raiith.iith.ac.in//), and IIT Kanpur (http://agropedialabs.iitk.ac.in/openaccess//) had least Betweenness and they represent alternative path for communication.

5.13 Degree Centrality (DC)

Nodes with the higher degrees are more central. The degree is simply the number of nodes at distance one. Though simple, degree is often a highly effective measure of the influence or importance of a node (Cecilia, 2010). With regard to degree centralization, the analysis revealed that node 7 (0.1803) of IIT Bombay (http://dspace.library.iitb.ac.in/jspui/browse?type=title), node 5 (0.1419) of IIT Delhi (http://eprint.iitd.ac.in/browse?type=title), node 186 (0.37229) of IIT-Hyderabad (http://www.eprints.org/uk/index.php/eprints-software/) and node 37 (0.20701) of IIT Kanpur (http://agropedialabs.iitk.ac.in/openaccess//) showed maximum degree centrality hence tend to have more power and more visible in their respective IR network. The data also present nodes with minimum degree centrality (DC,’=0) viz. Node 17 (http://dspace.library.iitb.ac.in/jspui/browse) of

IIT Bombay, node 13 (http://www.dspace.org/) of IIT Delhi, node 2 (http://raiith.iith.ac.in/) of IIT, Hyderabad and node 11 (http://agropedialabs.iitk.ac.in/openaccess//) of IIT Kanpur respectively.

5.14 Degree Prestige (DP) or In-Degree Centrality

Prestige measures are usually computed for direct networks only, since for this measure the direction is important property of the relation. Prestige is usually tied to the number of "choices" a node has which is related to the in-degree (as opposed to just the degree) of the node. A website that is linked to often has high prestige. The statistical data showed that node 2 (DP' =0.048) of IIT Bombay (http://dspace.library.iitb.ac.in/jspui/), node 2 (DP' = 0.040) of IIT Delhi (http://eprint.iitd.ac.in/), node 2 (DP' = 0.023) of IIT Hyderabad (http://raiith.iith.ac.in/) and node 3 (DP' = 0.0350) of IIT Kanpur (http://agropedialabs.iitk.ac.in/openaccess//) had maximum in-degree centrality.

5.15 Average Local Clustering Coefficient

The clustering coefficient (Watts & Strogatz, 1998), when applied to a single node, is a measure of how complete the neighborhood of a node is. When applied to an entire network, it is the average clustering coefficient over all of the nodes in the network. The clustering coefficient is a real number between zero and one that is zero when there is no clustering, and one for maximal clustering, which happens when the network consists of disjoint cliques (Newman, Strogatz, & Watts, 2001). The statistical report illustrated that the average local clustering coefficient of IIT-IR networks under study are almost close to 1, meaning the networks are completely clustered with adjacent nodes.

5.16 Influence Closeness Centrality (IRCC)

This refers to the degree with which an individual is nearer to all others in a network either directly or indirectly. Further, it reflects the ability to access information through the cluster of network members. In this way, closeness is considered to be the inverse of the sum of the shortest distance between each website and all others available on the network (Borgatti, 2005). The IRCC index is the ratio of the fraction of nodes reachable by each node to the average distance of these nodes from it. This index is optimized for graphs and directed graphs which are not strongly connected. The more central a node is, the lower its total distance to all other nodes. Closeness can be regarded as a measure of how long it will take to spread information from center point to all other nodes sequentially. The study found that largest possible closeness is present in node 7(0.391) of IIT Bombay (http://dspace.library.iitb.ac.in/jspui/browse?type=title), node 5 (0.291) of IIT Delhi (http://eprint.iitd.ac.in/browse?type=title), node 186 (0.372) of IIT Hyderabad (http://www.eprints.org/uk/index.php/eprints-software/) and node 3 (0.513) of IIT Kanpur (http://agropedialabs.iitk.ac.in/openaccess//); Smallest possible closeness (IRCC = 0) are seen in node 17 (http://dspace.library.iitb.ac.in/jspui/browse) of IIT Bombay, node 13 (http://www.dspace.org/) of IIT Delhi, node 2 (http://raiith.iith.ac.in/) of IIT Hyderabad and node 11 (http://agropedialabs.iitk.ac.in/openaccess//) of IIT Kanpur IR network respectively.

6. Conclusion

Visualization of an Institutional repository network is a popular way to understand and analyze the behavior of both the individual website and the overall network. Visualizing networks is of immense help for researchers in understanding new ways to present and manage data and to effectively convert the data into meaningful information (Tantipathananandh, Berger-Wolf, & Kempe, 2007). The study evidently visualized the IR network sites of Indian Institute of Technologies under study. Even though enormous studies are made on IR (Kamila, 2009; Thakuria, Das & Karmakar, 2010; Krishnamurthy & Kemparaju, 2011), the present study is one of its kind, since it had made an attempt to explore how well these networks are connected and identify the pattern of information flow from one node to another. The findings of the study revealed IIT institutional repository web networks are thinly connected; IIT, Hyderabad partakes more node out degree (47) with 971 edges, followed by IIT Kanpur (node, out degree: 30) with 1193 edges indicating these IR play predominant role in connecting the network among the Institutional Repositories under study. It is significant to note that, IIT Delhi has more number of edges (1386) but node out degree is merely 15, meaning the repository website is less central. IIT Kanpur and IIT Delhi had least average distance between linked websites, while IIT Bombay had a longer path between nodes; however the network with shorter path are preferred concerning to speed of information exchange. It is observed that the IIT Kanpur IR network had maximum number of cliques indicating all possible ties present among themselves. Even though, network analysis has become an important tool for investigators, all the necessary information is often distributed over a number of Web servers. To date, network centrality measures of academic SNS have not been considered in the perspective of impact assessment. The study results, while based on a small, exploratory attempt, suggest that such measures do relate to established impact metrics and therefore might be helpful, at least in supplementing existing forms of academic emergence.

References

Backstrom, L., Boldi, P., Rosa, M., Ugander, J., & Vigna, S. (2012, June). Four degrees of separation. In Proceedings of the 4th Annual ACM Web Science Conference (pp. 33-42). ACM.

Borgatti, S. P. (2005). Centrality and network flow. *Social networks, 27*(1), 55-71.

Börner, K. (2010). Atlas of Science: Visualizing What We Know. Cambridge, MA: The MIT

Cheliotis, G. (2017). Social Network Analysis (SNA) including a tutorial on concepts and method Social Media. Communications and New Media, National University of Singapore.

Chen, C. (2006). *Information visualization: Beyond the horizon*. Springer Science & Business Media.

Clement, G. P., & Rascoe, F. (2013). ETD management and publishing in the ProQuest system and the university repository: a comparative analysis. *Journal of Librarianship and Scholarly Communication, 1*(4), eP1074. http://dx.doi.org/10.7710/ 2162-3309.1074JJour.

Collins, S. T., Bradley, J. A., & Yassine, A. A. (2010). Analyzing product development task networks to examine organizational change. *IEEE Transactions on Engineering Management, 57*(3), 513-525.

Dincer, B. (2018). Introduction to Social Network and Link Analysis. [online] Academia.edu. Available at:

Edward Phillips, M., Gelaw Alemneh, D., & Reyes Ayala, B. (2014). Analysis of URL references in ETDs: a case study at the University of North Texas. *Library Management, 35*(4/5), 293-307.

Hoffmann, C. P., Lutz, C., & Meckel, M. (2016). A relational altmetric? Network centrality on ResearchGate as an indicator of scientific impact. *Journal of the Association for Information Science and Technology, 67*(4), 765-775.

Hulagabali, S. C. (2015). Institutional Repositories Initiated by Indian Institutes of Technology and Indian Institutes of Management: A Case Study. *DESIDOC Journal of Library & Information Technology, 35*(4).

IIT Bombay available at: http://dspace.library.iitb.ac.in/jspui/

IIT Delhi Details available at: http://library.iitd.ac.in/pdf/LibraryGuide.pdf

Kamila, K. (2009). Institutional repository projects in India.

Kimidi, S. S., Mallikarjuna, C., & Asthana, S. (2017). RAIITH: Institutional Repository at IIT Hyderabad and its impact in scientific community.

Kraker, P., Schlögl, C., Jack, K., & Lindstaedt, S. (2015). Visualization of co-readership patterns from an online reference management system. *Journal of Informetrics, 9*(1), 169-182.

Krishnamurthy, M., & Kemparaju, T. D. (2011). Institutional repositories in Indian universities and research institutes: A study. *Program, 45*(2), 185-198.

Link analysis details available at: https://www.techopedia.com/definition/30336/link-analysis

Luenberger, D. G. (1979). *Introduction to dynamic systems: theory, models, and applications* (Vol. 1). New York: Wiley.

Makino, K., & Uno, T. (2004, July). New algorithms for enumerating all maximal cliques. *In Scandinavian Workshop on Algorithm Theory* (pp. 260-272). Springer, Berlin, Heidelberg.

Newman, M. E., Strogatz, S. H., & Watts, D. J. (2001). Random graphs with arbitrary degree distributions and their applications. *Physical review E, 64*(2), 026118.

Ortega, J. L., & Aguillo, I. F. (2008). Visualization of the Nordic academic web: Link analysis using social network tools. *Information Processing & Management, 44*(4), 1624-1633.

Registry of Open Access Repositories available at: http://roar.eprints.org/

Saint-Charles, J., & Mongeau, P. (2009). Different relationships for coping with ambiguity and uncertainty in organizations. *Social Networks, 31*(1), 33-39.

Sourceforge.net. (2018). Social Network Visualizer/Git/ [6fff1f] /manual/generate.txt. [online] Available at: https://sourceforge.net/p/socnetv/git/ci/6fff1ff93c8b37026924549d66a910013b7cbf4d/tree/manual/generate.txt#l119

Sterman, L. (2014). Institutional repositories: An analysis of trends and a proposed collaborative future. *College & Undergraduate Libraries, 21*(3-4), 360-376.

Tantipathananandh, C., Berger-Wolf, T., & Kempe, D. (2007, August). A framework for community identification in dynamic social networks. In *Proceedings of the 13th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 717-726). ACM.

Thakuria, J., Das, R., & Karmakar, G. S. (2010). Institutional Repositories in India: A Comparative Study of Dspace and Eprints.

The Directory of Open Access Repositories available at: http://www.opendoar.org/

Thelwall, M. (2008). Social networks, gender, and friending: An analysis of MySpace member profiles. *Journal of the Association for Information Science and Technology, 59*(8), 1321-1330.

Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications* (Vol. 8). Cambridge university press.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of ‘small-world’ networks. *nature, 393*(6684), 440.

Yan, E., Ding, Y., & Zhu, Q. (2010). Mapping library and information science in China: A coauthorship network analysis. *Scientometrics, 83*(1), 115-131.

Yang, B., Liu, Z., & Meloche, J. A. (2010). Visualization of the Chinese academic web based on social network analysis. *Journal of Information Science, 36*(2), 131-143.

[About the author]

Kutty Kumar, Ph.D is presently working as Assistant Professor, Department of Library and Information Science in Sri Venkateswara Veterinary University, College of Veterinary Science, Proddatur. His research focuses on Digital Library Initiatives in Engineering Educational Institutions in Rayalaseema Region of Andhra Pradesh. He has 15 years’ experience in librarianship in both Engineering and Medical Educational Institutions. Around 43 articles were published in peer reviewed journals, 26 conference proceedings (both National and International) and had attended about 12 workshops and seminars. His subjects of interest include Digital Library, Web Technology, Cloud Computing, Data Mining and Computer Networks.