Comprehending Nodes Essentiality through Centrality Measures in Biological Networks

Fatima Sughra Zaidi[†], Ubaida Fatima[††], Bilal Ahmed Usmani[†], Ali Raza Jafri[†]

[†]Department of Biomedical Engineering, NED University of Engineering and Technology, Karachi, Pakistan
[††]Department of Mathematics, NED University of Engineering and Technology, Karachi, Pakistan

Summary
The study of any complex system in the form of a network structure has always been an efficient approach, being the underlying aspect of graph theory. For the topological and structural network analysis, the concept of identifying and understanding the influential nodes is a beneficial method based on the connectivity of its structure. For this purpose, centrality measures are computed and the elements of the network are ranked through the obtained centrality scores. This method has been widely used for social networks, however, it gained emerging importance in biological networks and different areas of application. In this study, we have computed and compared degree centrality, closeness centrality, betweenness centrality, Katz centrality and PageRank algorithm on a biological network of Saccharomyces Cerevisae (eukaryotic organism) protein interaction. These measures predicted hundreds of important nodes interpreting the essential proteins. The biological significance of our result was sought through established literature. Out of top 30 proteins (i.e. 5 for each measure) we predicted, 29 were found to be highly significant which depicted the fact that absence of these proteins may result in lethality or destruction. Through these findings we concluded that for structural analysis of a complex biological network, centrality measures are proved to be helpful based on the strong prediction of relevant information regarding underlying biological mechanisms. The integration of centrality metrics with the biological knowledge developed an improved index for identification of network essentiality.

Key words:
Centrality measures, graph theory, biological networks, essential nodes

1. Introduction

Interlinked elements functioning together as a part of any mechanism results in forming a complex system. The representation of a system in the form of a network may ease the complexity to explore the actual system which helps to yield a better understanding of its structure. The initial step taken to explore interaction networks can be of identifying the important nodes, an approach under the perspective of graph theory. It is generally known that in most networks the few important components are placed in specific positions and are more influential than others, such as important genes and proteins in biological networks or people in social networks [1]. Determining the meaningful vertices in complex networks has been a crucial matter and many studies identified different methods for this purpose [2]. One of those methods for network analysis is computing centrality measures, which are at times useful for identifying the dominant nodes in a network through ranking. Centrality measures have greatly been a part of studies which deals in analyzing different types of networks involved in social [3][4], biological [5][6], traffic [7], biometric [8], epidemiological [9] and information systems [10]. The primary goal of static centrality measures is filtering of large or complex data available, and then retrieving the significant information related to network components [11]. Some centrality measures are widely used including degree centrality (DC), closeness centrality (CC) [12], betweenness centrality (BC), eigenvector centrality (EC) [13], and Katz centrality (KC). These measures have acquired great relevance regarding simple or complex network analysis by mainly evaluating the topological properties of networks [14]. In the network structure, the top ranked vertices are most relevant, and are supposed to be key players in the systematic mechanism under consideration. This results in highlighting important apexes in the network, usually from a very large network having thousands of nodes. However, it is to be noted that centrality measures may be influenced by the type of network chosen, the direction of links and the possible weights present on the links [15].

Biological networks have recently been evolved as an excellent measure to model biological processes. Among components at cellular level, biological functions are mostly dependent on their structural properties. Usually, most significant biological activities are not the result of a particular component but depends on the integrated effects of multiple components interacting with each other, generating networks. Thus, to study biology under the context of networks is very cardinal and fruitful [16][17]. Various biological interactions in the form of networks mainly include protein-protein interactions, gene regulatory and metabolic networks. To use biological networks and analyze them, the approach to evaluate the network's topological structure is said to be efficient. Topological analysis enlightens the possible network behavior in the regulation of biological processes and aids in discovering the basic mechanisms. In order to explore the network the topological framework may include:

Manuscript received September 5, 2019
Manuscript revised September 20, 2019

(1) Collective overall behavior: Global properties such as diameter, scale-free or small world characteristics of a network,
(2) Subnetwork behaviors: Sub-graphs, clustering, functional motifs identification.
(3) Individual behaviors: Ranking of important nodes through centrality indices of various network elements. [16] [18] (This perspective will stay our focus throughout this study.)

A biological network consists of a set of nodes or vertices representing the biological components and the edges or links that denote relationships between node pairs. In protein–protein interaction networks, vertices are proteins and links are their physical interactions [19]. Apart from the difficulties, these networks concurrently provide convenience to understand the cellular biology. For its structural analysis the method of ranking the nodes and predicting their influence is essential based on its structural connectivity [20]. Centrality metrics method is specifically helpful to classify the components playing a key role in biological mechanisms. More than one centrality metrics must be sought for biological network analysis as considering only one is not sufficient [21] [22].

There are broad applications in network-based approaches for biological sciences. Such as by considering the disease networks, we may get a good understanding of mechanisms behind the disease by identifying the elements and pathways that are causing the disease outbreak. It may offer goals for prevention, prediction or control through drug development or changes in social adaptations [23]. It is exhibited that in protein-protein interaction (PPI) networks, highly connected vertices are functionally significant and the removal of those vertices may link to lethality [24]. The topological study carried out in metabolic and transcriptional regulatory (TR) networks have been successfully helpful in identifying the essential components that have a dominant role in important functional biological processes for few microorganisms [25][26]. Comparison of different centrality measures was done on gene regulatory (GR) network of E.coli which concluded that motif-based centrality outstood in identifying 83% (i.e. 15 out of 18) global regulator genes of the network by combining the underlying biological knowledge [20].

In this paper, different centrality metrics proposed for network analysis are discussed and correlation between them has been calculated to evaluate their performance. The metrics were computed on a yeast (i.e. the most extensively used eukaryotic organism in scientific research [27]) PPI network data. The proteins having high centrality scores were then compared with their biological significance present in established literature.

2. Network and Centrality Metrics

A network is a graph structure that may be used to model simple or complex systems mathematically. Such a graph contains vertices ***'V'*** or nodes and edges ***'E'*** or links, i.e. ***G=(V,E)***. In graph theory, the importance of these vertices and their meaningful interpretation can be done using centrality measures [28]. These graphs exhibits interactions among elements and can be completely interpreted by its adjacency matrix ***'A'***. For an undirected network the adjacency matrix is of symmetric type in which the element is equals to 1 if nodes ***'i'*** and ***'j'*** share an edge or link and zero otherwise, whereas in the case of directed networks the adjacency matrix is asymmetric as if node $\boldsymbol{A_{ij} = 1}$, than it may be possible that node $\boldsymbol{A_{ij} \neq 1}$ [29] [30].

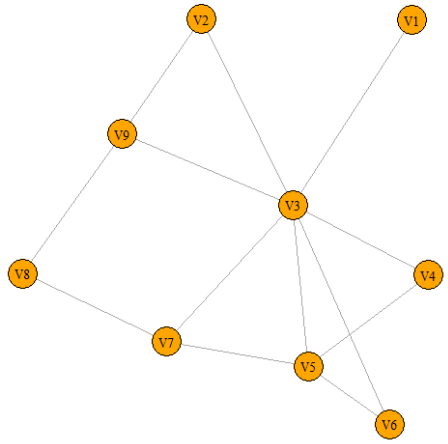

Fig. 1 A simple nine nodes (vertices) network 'G' [2]

Centrality can be referred to as a function ***'C'*** that assigns some value to all of the vertices of a graph ***'G'***. The vertex having a greater value is considered to be more important within a graph. Node centrality metrics are extensively used being essential in many network applications such as exploration and ranking of its elements for analysis. Many such metrics have been proposed up till now based on the fact that in what aspect the importance of a node is defined [31]. Centrality measures consider the network topology mainly through the network's adjacency matrix as discussed above. The metrics may be referred to as static because the temporal dimensions of any network are oversighted. These should not be applied to any system without focusing on its properties because it is known that this method of analysis may result in flaws if computed for dynamically emerging systems [11]. However, centrality measures are useful as a basic analytical tool of network analysis.

2.1 Degree

Degree centrality is uncomplicated yet the most well-known. It is a local measure that considers the direct neighborhood of a particular node by counting the number of direct links connected to a node. It is a kind of measure that can be applicable on all types of networks and can be efficiently computed for large networks as well [10]. There are two degree centralities for directed type of networks, in-degree and out-degree. In-degree centrality only recognizes incoming edges whereas out-degree is for the outgoing edges [20]. Degree shows the node's importance regardless of its position in the network. It may work more better in particular tasks such as determining disease spreaders more than other sophisticated metrics. The incoming edges in majority cases may be depicted as positive connection or signal from the neighbor vertex, such as in a network of research publications for academic search evaluation, a manuscript's incoming edge shows its earned citations, i.e. usually a positive endorsement. However, it may sometimes be misleading if pointing out the flaws of the paper [11]. The computational formula used for degree centrality is defined as follows:

$$Degree\ Standardized\ Score = \frac{D_i}{n-1} \quad (1)$$

Where, D= degree of node, i= target node, n= no. of nodes

2.2 Closeness

Closeness centrality interprets information about the shortest paths in a network, since it is a diameter related measure. Closeness is adequate in distinguishing nodes having low degree [10]. It is the inverse of the sum of shortest distances of a node to all other nodes in the network. Closeness can be efficiently computed for strongly connected networks as the path distance between nodes is specified for connected nodes in a pair [20]. A vertex is more central if it is near to a great number of other vertices i.e. carrying the highest values of closeness centrality. It aids a vertex for communication within the entire graph. Similarly, a vertex with lowest value is far from other vertices and will have difficulty in communicating or passing the information. [32] As being the reciprocal of farness, closeness centrality is mathematically defined as:

$$Standardized\ Closeness_i = \frac{n-1}{\Sigma\, no.of\ paths_i} \quad (2)$$

Where, i= target node, n= no. of nodes

2.3 Betweenness

Betweenness is the measure which shows that how much a node lies between the shortest path of other nodes. The node that is central accordingly can control the flow of information between vertices, as it acts as a connector among the nodes and therefore it is known as the most popular type of flow related measure [10] [20]. As there can be multiple shortest paths or geodesics between two nodes, the geodesics passing through the target node must be tracked [30]. Betweenness can be computed through the following formula, however as compared to degree and closeness it is more complex to calculate for large networks.

$$Betweenness = \sum \frac{Total\ no.of\ paths\ through\ node_i}{Total\ no.of\ paths} \quad (3)$$

Where, i=target node

2.4 Katz Centrality

Katz centrality metric is based on the eigenvector centrality, i.e. a node is said to be important if it is linked to other important or influential nodes [11]. Eigenvector is not efficiently applicable on directed networks because it works well only if the network is highly connected. An undirected network is strongly connected whereas in a directed network not all vertices have incoming links, which results in a null eigenvector score of those respective vertices. In such cases, Katz centrality overcomes this limitation and can be used for such networks [14]. Katz centrality solves this null or zero issue by initially assigning a minimum score to all nodes of the network, regardless of the node's position and influence. In this way none of the score remains null. Katz centrality covers a wider part of the network as compared to eigenvector centrality [12]. The network must be loop-free to calculate this metric [33][34]. Many variants of Katz has been proposed in this era getting more importance in other fields as well including biological and epidemiological systems. Examples may include predicting the neuronal activity [35] and an algorithm inspired from Katz centrality was used for the prioritization of disease genes by integrating genome scale protein interaction network [36]. In this following study [37], Katz centrality can be mathematically defined as:

$$KC_i = \alpha \sum_j A_{ij}\ V_j + \beta \quad (4)$$

Where,
α= penalty on the distant connections to a vertex centrality score i.e. $\alpha < 1/\lambda max$
β=preassigned constant centrality value,
A=adjacency matrix of graph G with eigenvalues λ,
V=highest eigenvector of A.

2.5 PageRank

PageRank is a famous ranking algorithm which was proposed with the goal of ranking web pages in a web network. This metric played a primary role in the phenomenal achievement of Google's internet search engine. [11] This algorithm evaluates the relative importance of a web page that is based on the web's graph or network and is proved to be relevant for directed graphs

[38]. According to this study [37], PageRank algorithm is presented as:

$$PR_i = \alpha \sum_j A_{ij} \frac{V_j}{K_j} + \beta \quad (5)$$

Where,
α=penalty on the distant connections to a vertex centrality score,
β=preassigned constant centrality value,
A=adjacency matrix of graph G,
V=highest eigenvector of A.

3. Data and Experimental Methodology

3.1 Data Description

A Saccharomyces Cerevisae (budding yeast) protein-protein interaction network consisting of 2361 nodes, 7182 edges and 536 loops is used for the computation of centrality measures to identify the important proteins. The nodes in this network denote the proteins present in yeast whereas the edges are directed and unweighted showing the physical interaction among them. The data is acquired from http://vlado.fmf.unilj.si/pub/networks/data/bio/yeast/yeast.htm.

3.1.1 Data Processing

To compute the centrality measures, adjacency matrix of the interaction data was required. An adjacency matrix can be said as a square matrix that is used to represent a graph which labels the vertices based on their connectivity. Therefore, according to the connections the adjacency matrix was formed on Microsoft Excel, i.e. value of ***1*** if two nodes ***(i,j)*** are connected and ***0*** otherwise. After complete formation, the adjacency matrix was used for further computations.

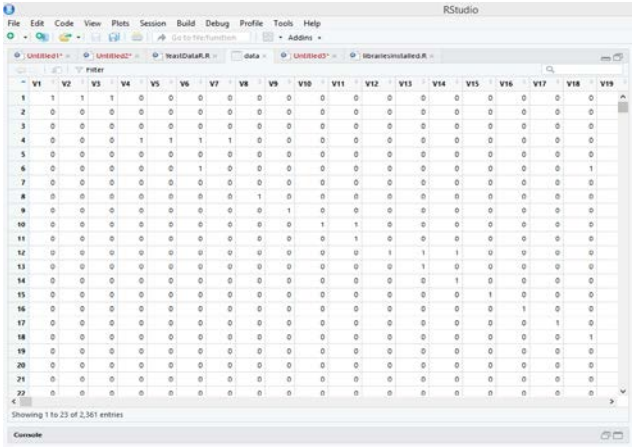

Fig. 2 A part of the data's adjacency matrix in Rstudio environment after processing

3.1.2 Data Visualization

The experimental data can be modelled as a network and shown as a node-link diagram. In this study, the yeast PPI data described above was used and by Gephi-0.9.2 software, a structural network arrangement of the connection data was generated.

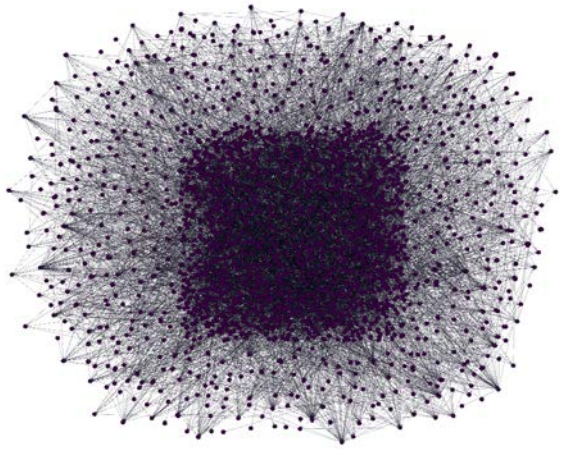

Fig. 3 Visualization of Yeast Protein-Protein Interaction Network

3.2 Computation of Centrality Measures

The discussed centrality measures including indegree, outdegree, betweenness, incloseness, outcloseness, Katz centrality and PageRank algorithm were computed for the protein interaction data to identify the significant yeast proteins among all. Analysis was performed through MATLAB R2017a software. Before computing Katz centrality, the self-interactions were removed from the network as the graph must be loop free as mentioned in the study of [22] and more. Based on the high centrality scores, 5 top proteins were separated for each metric, then the biological importance of those respective proteins were evaluated.

3.3 Correlation Coefficients Calculation

The comparison between the applied centrality metrics and their performance was done by calculating correlation coefficients represented by a value '*r*'. Two kinds namely Pearson's and Kendall's Tau coefficients were computed and their results were compared.

3.3.1 Pearson's Correlation Coefficient

Pearson correlation measures the extent of a linear relationship between any two variables that are showing a set of numerical data. It is the most commonly used correlation static. According to the study of Cohen (1988), the standard proposed can be availed to determine the

correlation and thus evaluate the strength of the possible relationship. Correlation values of 'r' between 0.10 and 0.29 are classified as a small association i.e. less correlation, coefficients between 0.30 and 0.49 show medium association, and coefficients of 0.50 and above till 1, interpret a large and strong relationship [39].

3.3.2 Kendall's Tau Correlation Coefficient

Kendall's Tau coefficient is particularly appropriate for centrality measures because the problems of distribution and normalization that may differ between the metrics do not alter it. It is calculated pairwise i.e. combination of two measures. It is a non-parametric measure which determines the strength of similarity between two groups of ranks that are given to the same objects. As a result it gives the value in the range [-1,1], where ideal correlation is interpreted at (+1), where (-1) shows no correlation and scores near zero show weak correlation [14].

4. Results and Discussion

By using the applauded method of network node analysis explained in section 2, the influential proteins in budding yeast were evaluated. As shown in Table 1 the following centrality values were extracted as the top 5 ranked proteins from highest to lowest for each centrality metric.

Table 1: The top five computed values of each centrality measure for yeast proteins (with the respective standards and systematic names)

Highly Ranked Proteins of the Network			
S.No.	**Node No.**	**Protein Name**	**Betweenness (BC)**
1	1443	YKU80 (YMR106C)	**36407.8574**
2	549	HHF1 Histone H4 (YBR009C)	29879.6511
3	302	CKA1 Casein Kinase II (YIL035C)	28526.6120
4	644	PWP2 Periodic Tryptophan Protein (YCR057C)	27060.5144
5	566	RPC40 DNA-directed RNA polymerase I, III (YPR110C)	26867.3302
			Outdegree (ODC)
1	209	SEC27 Coatomer Complex (YGL137W)	**60.0000**
2	147	SRP1 Karyopherin-alpha (YNL189W)	54.0000
3	120	BUD20 (YLR074C)	53.0000
4	61	UTP22 (YGR090w)	49.0000
5	566	RPC40 DNA-directed RNA polymerase I, III (YPR110C)	48.0000
			Indegree (IDC)
1	2022	SEN15 tRNA splicing endonuclease delta subunit (YMR059W)	**47.0000**
2	1817	APG12(PF04110-domain name)	37.0000
3	644	PWP2 Periodic Tryptophan Protein (YCR057C)	32.0000
4	1896	TEM1 GTP-binding protein (YML064C)	31.0000
5	1202	SMX2 snRNP G protein (YFL017W-A)	30.0000
			Outcloseness (OCC)
1	599	HUL5 ubiquitin-protein ligase (E3) (YGL141W)	$\mathbf{9.9747 \times 10^{-6}}$
2	736	SRB8 and SRB10 (YCR081W)	9.9619×10^{-6}
3	230	MYO5 myosin I (YMR109W)	9.9039×10^{-6}
4	853	SHE4 (YOR035C)	9.9039×10^{-6}
5	758	RSM7 (YJR113C)	9.8473×10^{-6}
			Incloseness (ICC)
1	877	AUT10/ATG18 (YFR021W)	$\mathbf{9.9753 \times 10^{-6}}$
2	1099	RLI1 (YDR091C)	9.9546×10^{-6}
3	1834	THS1 threonyl tRNA synthetase (YIL078W)	9.9441×10^{-7}
4	1835	Arginyl-tRNA synthetase (YDR341c)	9.9441×10^{-7}
5	1836	URA1 dihydroorotate dehydrogenase (YKL216W)	9.9441×10^{-7}
6	1837	MLP2 (YIL149C)	9.9441×10^{-7}
			Katz Centrality Index (KC)
1	302	CKA1 Casein Kinase II (YIL035C)	1.7739
2	566	RPC40 DNA-directed RNA polymerase I, III (YPR110C)	1.7108
3	1443	YKU80 (YMR106C)	1.7047
4	784	HRR25 casein kinase I (YPL204W)	1.7041
5	209	SEC27 Coatomer Complex (YGL137W)	1.7008
			PageRank (PR)
1	1392	HEK2 (YBL032w)	**0.0068**
2	901	ORC2 origin recognition complex (YBR060C)	0.0059
3	1058	RFX1 repressor (YLR176C)	0.0053
4	919	YTH1 protein of the 3' processing complex (YPR107C)	0.0042
5	1422	CYM1 (YDR430c)	0.0041

After attaining the results i.e. the ranking of proteins based on centrality scores, the influential proteins out of a large number of them were be able to predicted without any

difficult experiments in a laboratory. Generally, it is challenging to predict essential proteins as required experimental approaches are time consuming, expensive, complex and laborious [40][41]. It is critical to note that some proteins (nodes) were repetitively seen important i.e., having high centrality scores in majority of the measures (i.e., CKA1 Casein Kinase II, YKU80, SEC27 Coatomer Complex and PWP2 Periodic Tryptophan Protein). The highlighted proteins can be directly considered for further study. Therefore, the computed results were linked to their biological functions as shown in Table 2 below by searching the previous studies relevant for the extracted yeast proteins.

Table 2: Biological Significance of the Top Ranked Yeast Proteins

S.No	Saccha-romyces Cerevisae Important Proteins	Primary Litera-ture (Author & Year)	Biological Importance on Basis of Literature
Betweenness			
1	YKU80	Boulton et al.(1996) [42]	Plays a primary role in repair of restricted enzymes and DNA double strand break pathway, saving it from an error-prone pathway.
2	HHF1 Histone H4	Prado et al.(2005), Ge Z et al.(2013), Thakr et al.(2019), Yeom et al.(2018) [43][44] [45][46]	Its partial depletion from yeast resulted in the formation of incomparable structures of DNA, increasing instability in genes. Has great significance in progression of cell cycle, DNA damage response.
3	CKA1 Casein Kinase II	Kanki et al(2013), Qiu et al. (2016), Matsuzaki & Shinohara(2018) [47][48][49]	CKA1 phosphorylates Dgk1 diacyglycerol (DAG) kinase and results in formation of PA (phosphatidic acid) in yeast and is involved in mitophagy.
4	PWP2 Periodic Tryptophan Protein	Shafaatian et al (1996)[50]	Part of a complex that plays a key role in cell separation and cell growth.
Outdegree			
5	SEC27 Coatomer Complex	Duden et al. (1994), Gabriely et al. (2007) [51][52]	Contains 45% identity in sequence to the mammalian coatomer subunit beta. It encodes membrane proteins along with ARF1 for ER to Golgi transport and is involved in former steps of protein endosomal sorting in yeast.
6	SRP1 Karyopherin-alpha	Tabb M. et al (2000), Chen & Medura (2014) [53][54]	Acts as a nuclear localization signal receptor and a new function was suggested of protein degradation. The removal of Srp1 is said to be lethal.
7	BUD20	Ni & Snyder (2001), Bassler et al. (2012) [55][56]	Bud20 zinc-finger type protein interacts with the pre-60s subunit and performs the pre-ribosome nuclear export facilitating its extension of N-terminal.
8	UTP22	MB et al. (2012), Albert et al. (2016) [57][58]	Serves in the coordination of ribosomal protein gene transcription with rRNA. Utp22 works with Utp8p to collect the aminoacyl-tRNA, then combines with Utp8p-tRNA complex to transport the aminoacyl-tRNAs to protein Los1p.
Indegree			
9	SEN15 tRNA splicing endonuclease delta subunit	Tsuboi et al. (2015) [59]	Involved in the cleavage of Cbp1 (cytochrome b mRNA processing 1) that is cotranslationally confined to mitochondria
10	APG12	Kuma et al. (2002) [60]	It is a part of protein complex involved in autophagy.
11	PWP2 Periodic Tryptophan Protein	Shafaatian et al (1996)[50]	(Repeated protein) Described above in betweenness section.
12	TEM1 GTP-binding protein	Shiramaya et al. (1994), Scarfone et al. (2015) [61][62]	It encodes a novel GTP-binding protein and the absence was lethal as it is highly required for the termination from the M phase. The spindle position checkpoint relies on the GTPase Tem1 and few others for activating the mitotic exit which otherwise may result in abnormality.
Outcloseness			
13	HUL5 ubiquitin-protein ligase (E3)	Fang et al. (2011), Fang & Mayor (2012) [63][64]	Highly required to manage the cell fitness in response to heat shock i.e. essential for quality control and degradation of misfolded proteins.
14	SRB8 and SRB10	Borggrefe et al. (2002), Larchan & Winston (2005) [65][66]	Together important for Gal1 transcription and Gal4 activation in yeast.
15	MYO5 myosin I	GA & Li (2004), Giblin et al.(2011) [67][68]	Myosin I protein is involved in the scission vesicles for the event of endocytosis, i.e. these are recruited and contain biochemical mechanisms at the endocytic sites.
16	SHE4	Wesche et al. (2003), Toi et al. (2003) [69][70]	Essential for proper function of myosin, hence it is involved in myosin based two events; mRNA localization and endocytosis. She4 is a motor domain myosin binding protein and controls the myosin function regulator.
Incloseness			
17	AUT10 (similar to hypothetical protein)	Not found	-
18	RLI1	Dong J et al. (2004), Kispal G. et al. (2005), Alhebshi A. et al. (2012) [71] [72] [73]	Depletion of iron sulphur protein RLI1 causes translation inhibition. It has dual functions; ribosome biogenesis and translation initiation. The reactive oxygen species (ROS) is increased by its dysfunction, increasing the oxidative stress.

19	THS1 threonyl tRNA synthetase	Ruan et al (2015) [74]	The mutations in yeast THS1 were studied and revealed pathology related mutations in mammalian cells due to similar damaging effects.
20	Arginyl-tRNA synthetase	Delagoutte et al. (2000) [75]	Involved in the catalytic pathway and highlights the aminoacylation reaction dynamics of structure.
Katz Centrality			
21	CKA1 Casein Kinase II	Kanki et al(2013), Qiu et al. (2016), Matsuzaki & Shinohara(2018) [47] [48] [49]	(Repeated protein) Described above in betweenness section.
22	RPC40 DNA-directed RNA polymerase I, III	Mann C. et al. (1987), Shpakovski & Shematorova (1999) [76][77]	This is essential in yeast for cell viability. It is the part of the polymerase core and minimizes the genetic damage.
23	YKU80	Boulton et al.(1996) [42]	(Repeated protein) Described above in betweenness section.
24	HRR25 casein kinase I	Hoekstra et al. (1991) Kafadar et al. (2003), Schaefer et al. (2006) [78] [79] [80]	In yeast it is defined by mutation Hrr25-1. However, its deletion disturbs the meiotic and mitotic cell division. It plays an important part in calcineurin signaling that is a phosphatase protein required in yeast in order to respond to a number of environmental stresses. The 40s ribosomal subunit maturation is regulated by the Hrr25 kinase activity.
25	SEC27 Coatomer Complex	Duden et al. (1994), Gabriely et al. (2007) [51][52]	(Repeated protein) Described above in out-degree section.
PageRank			
26	HEK2	Hasegawa Y. et al. (2008), Mauchi N. et al. (2010) [81] [82]	During transport of mRNA it acts as a translational repressor. It is also required for the localization of an mRNA at the bud tip.
27	ORC2 origin recognition complex	Matsuda K. et al. (2007), Kan J. et al. (2008) [83] [84]	A novel function was revealed of ORC in mediation of histone methylation however already involved in DNA replication, centromere and telomere function and transcriptional control.
28	RFX1 repressor-regulatory factor X	Zhang and Reese (2005), Zaim J. et al. (2005) [85] [86]	It acts as an effector at the checkpoint pathway during any DNA damage. It may also impact on deoxyribo-nucleotide synthesis.
29	YTH1 protein of the 3' processing complex	Barabino et al. (2000), Casanal et al. (2017) [87] [88]	It combines factors required for efficient and particular polyadenylation and helps in the coordination of mRNA 3'-end processing. It also takes part in cleavage site recognition by binding with mRNA CYC1.
30	CYM1	Jonson L. et al. (2004) [89]	It is an essential endoprotease that enhances peptide secretion if its disruption is caused.

The biological significance for the total 30 considered proteins was found from the relevant literature available, except for one which was unavailable, i.e. a result of closeness centrality. These findings show that this method of graph theory is useful for significant protein identification. Therefore, now it can be assumed that the absence of these proteins from the system may result in improper cell functioning. Closeness centrality is not considered much appropriate for directed type of social networks [14] and it has not been resulted to be a good essentiality indicator in biological networks for few organisms as mentioned in the study of [90]. Our findings in Table 2 indicate that majority of the important proteins predicted through closeness came out be biologically important as well. Fundamentally, the structure of social and biological networks differ from each other, particularly with respect to modularity [91]. Hence, there is a possibility that methodologies may behave slightly different in various kinds of networks due to their inherent topology.

It is not necessary to expect an accurate match between biological functions and network topology [18] but such tools of computation can definitely provide the opportunity for accurate prediction and logical guessing. Current work advocates to a great support for study of complex and large biological networks.

4.1 Comparative Analysis between Applied Measures

The properties of any network prominently have an impact on the correlation values of centrality metrics. The Pearson's and Kendall's correlation coefficients between pairs of both metrics were calculated to analyze the similarities and differences of their performance in our aimed network. The degree and closeness centralities were taken as a single measure including both in and out-degree and closeness respectively in order to calculate the correlation with the single values of PageRank and Katz centrality. The following tables 3 and 4 show the correlation values of both types.

In a previous study, three centralities including degree, closeness and betweenness were compared for identification of important components in three eukaryotic organisms including yeast. It concluded that the average centrality values for influential proteins is remarkably higher than the scores of non-influential proteins [92]. We mainly considered correlation between Katz centrality and PageRank with the other used centralities as these two have not been correlated with the rest previously.

Table 3: Pearson's Correlation coefficients (r) between the centralities computed for PPI network analysis

Metrics	BC	DC	CC	PR	KC
KC	0.752	**0.997**	**0.177**	0.961	1
PR	0.759	**0.974**	**0.114**	1	0.961

According to Pearson's correlation, Katz centrality and PageRank both have strong relation with degree centrality, then second highest relation is between themselves, as depicted by the above table. Betweenness also shows a pretty good correlation unlike closeness centrality which shows a positive but weak correlation. These values support that closeness may have been a weak indicator for the particular network, as mentioned earlier. However, rest of the metrics agree in prediction of influential proteins and resulted to be useful.

Table 4: Kendall's Pairwise Correlation coefficients (r) between the centralities computed for PPI network analysis

Metrics	BC	DC	CC	PR	KC
KC	**0.657**	**0.921**	0.720	0.687	1
PR	0.637	**0.839**	**0.464**	1	0.687

Kendall's correlation coefficients are similar to the ones obtained from Pearson's, the highest correlation is with DC. The main difference is the relation with CC, since here it represents a good correlation which leads to the certainty that it may identify influential nodes, perhaps not accurately but to some extent which may be proficient, as indicated in table 2 above from our biological significance analysis. Strong correlation also exhibits a point that instead of applying all metrics, computation can be made comparatively simpler and just the more versatile centralities could be applied, which in this case may be Katz Centrality and PageRank algorithm.

5. Conclusion

In this research, the ranking methodology of graph theory was studied through centrality measures. Having various applications in many fields, we applied the discussed metrics to a biological yeast PPI network and explored the literature to check their biological significance. Centrality measures were initially proposed in social networks but its usage kept evolving in other fields such as biological, information and epidemiological systems. This study is beneficial to exemplify the application of these measures on biological systems, specifically PPI networks considering the important fact that study of yeast proteins is valuable for human research on cellular level due to the protein structure similarities. The method employed has been seen advantageous in this area because biological networks are complex and may contain difficulties in developing procedures for analysis, therefore even providing basic information and hints through valid prediction can be justified to be already useful. However, some experimental proves are still required as mathematical computations cannot directly prove that the predicted elements of the network are important. In future if exhaustive implementation of these computations is done, research on such proteins may also be carried out whose docking is comparatively difficult. The further study by applying different metrics and combining the related biological knowledge can also explore other interesting features of a network.

References

[1] M. Ghasemi, H. Seidkhani, F. Tamimi, M. Rahgozar, and A. Masoudi-Nejad, "Centrality Measures in Biological Networks," Curr. Bioinform., vol. 9, no. 4, pp. 426–441, 2014.

[2] S. Wang, Y. Du, and Y. Deng, "Commun Nonlinear Sci Numer Simulat A new measure of identifying influential nodes : Efficiency centrality," Commun. Nonlinear Sci. Numer. Simul., vol. 47, pp. 151–163, 2017.

[3] A. Landherr, B. Friedl, and J. Heidemann, "A Critical Review of Centrality Measures in Social Networks," Bus. Inf. Syst. Eng., vol. 2, no. 6, pp. 371–385, Dec. 2010.

[4] A. H. Dekker, "Centrality in Social Networks: Theoretical and Simulation Approaches."

[5] A. Özgür, T. Vu, G. Erkan, and D. R. Radev, "Identifying gene-disease associations using centrality on a literature mined gene-interaction network," Bioinformatics, vol. 24, no. 13, Jul. 2008.

[6] E. Zotenko, J. Mestre, D. P. O'Leary, and T. M. Przytycka, "Why do hubs in the yeast protein interaction network tend to be essential: Reexamining the connection between the network topology and essentiality," PLoS Comput. Biol., vol. 4, no. 8, Aug. 2008.

[7] P. HOLME, "CONGESTION AND CENTRALITY IN TRAFFIC FLOW ON COMPLEX NETWORKS," Adv. Complex Syst., vol. 06, no. 02, pp. 163–176, Jun. 2003.

[8] K. K. A. Ghany, H. A. Hefny, A. E. Hassanien, and N. I. Ghali, "A Hybrid approach for biometric template security," in Proceedings of the 2012 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, ASONAM 2012, 2012, pp. 941–942.

[9] C. S. Riolo, J. S. Koopman, and S. E. Chick, "Methods and measures for the description of epidemiologic contact networks," J. Urban Heal., vol. 78, no. 3, pp. 446–457, 2001.

[10] U. Kang, S. Papadimitriou, J. Sun, and H. Tong, "Centralities in Large Networks: Algorithms and Observations," pp. 119–130, 2011.

[11] H. Liao, M. S. Mariani, M. Medo, Y. C. Zhang, and M. Y. Zhou, "Ranking in evolving complex networks," Phys. Rep., vol. 689, pp. 1–54, 2017.

[12] M. Newman, Networks: An Introduction. 2010.

[13] P. Bonacich, "Some unique properties of eigenvector centrality," Soc. Networks, 2007.

[14] U. Federal, D. O. Rio, G. Do, and F. Grando, "On the Analysis of Centrality Measures for Complex and Social Networks," 2015.

[15] J. Hua, M. L. Huang, W. Huang, and C. Zhao, "Applying graph centrality metrics in visual analytics of scientific standard datasets," Symmetry (Basel)., vol. 11, no. 1, pp. 1–19, 2019.

[16] X. Ma and L. Gao, "Biological network analysis: Insights into structure and functions," Brief. Funct. Genomics, vol. 11, no. 6, pp. 434–442, 2012.
[17] A. L. Barabási, N. Gulbahce, and J. Loscalzo, "Network medicine: A network-based approach to human disease," Nature Reviews Genetics. 2011.
[18] M. Jalili et al., "Evolution of centrality measurements for the detection of essential proteins in biological networks," Front. Physiol., vol. 7, no. AUG, 2016.
[19] U. Stelzl et al., "A human protein-protein interaction network: A resource for annotating the proteome," Cell, 2005.
[20] D. Koschützki and F. Schreiber, "Centrality analysis methods for biological networks and their application to gene regulatory networks," Gene Regul. Syst. Bio., vol. 2008, no. 2, pp. 193–201, 2008.
[21] S. Wuchty and P. F. Stadler, "Centers of complex networks.," J. Theor. Biol., vol. 223, no. 1, pp. 45–53, Jul. 2003.
[22] D. Kosch and F. Schreiber, "Comparison of Centralities for Biological Networks," Proc. Ger. Conf. Bioinforma., pp. 199–206, 2004.
[23] E. E. Schadt, S. H. Friend, and D. A. Shaywitz, "A network view of disease and compound screening.," Nat. Rev. Drug Discov., vol. 8, no. 4, pp. 286–95, 2009.
[24] H. Jeong, S. P. Mason, A. L. Barabási, and Z. N. Oltvai, "Lethality and centrality in protein networks," Nature, 2001.
[25] O. Resendis-Antonio et al., "Modular analysis of the transcriptional regulatory network of E. coli.," Trends Genet., vol. 21, no. 1, pp. 16–20, Jan. 2005.
[26] O. Resendis-Antonio, M. Hernández, Y. Mora, and S. Encarnación, "Functional Modules, Structural Topology, and Optimal Activity in Metabolic Networks," PLoS Comput. Biol., 2012.
[27] A. A. Duina, M. E. Miller, and J. B. Keeney, "Budding yeast for budding geneticists: A primer on the Saccharomyces cerevisiae model system," Genetics, 2014.
[28] A. Noori, "On the relation between centrality measures and consensus algorithms," Proc. 2011 Int. Conf. High Perform. Comput. Simulation, HPCS 2011, pp. 225–232, 2011.
[29] C. Sciarra, G. Chiarotti, F. Laio, and L. Ridolfi, "A change of perspective in network centrality," Sci. Rep., vol. 8, no. 1, 2018.
[30] F. Bloch and M. O. Jackson, "Centrality Measures in Networks," SSRN Electron. J., 2016.
[31] S. P. Borgatti and M. G. Everett, "A Graph-theoretic perspective on centrality," Soc. Networks, vol. 28, pp. 466–484, 2006.
[32] U. Fatima and S. Hina, "Efficient Algorithm for Maximal Clique Size Evaluation Broad Learning of its Relation with Centrality Metrics for Large Dataset Networks," 2019.
[33] A. A. Hagberg hagberg, lanlgov -Los, D. A. Schult, and P. J. Swart swart, "Exploring Network Structure, Dynamics, and Function using NetworkX," 2008.
[34] M. Ogura and V. M. Preciado, "Katz centrality of Markovian temporal networks: Analysis and optimization," in Proceedings of the American Control Conference, 2017.
[35] J. M. Fletcher and T. Wennekers, "From Structure to Activity: Using Centrality Measures to Predict Neuronal Activity," Int. J. Neural Syst., vol. 28, no. 02, p. 1750013, 2018.
[36] J. Zhao, T. H. Yang, Y. Huang, and P. Holme, "Ranking candidate disease genes from gene expression and protein interaction: A katz-centrality based approach," PLoS One, 2011.
[37] S. Oldham, B. Fulcher, L. Parkes, A. Arnatkeviciute, C. Suo, and A. Fornito, "Consistency and differences between centrality metrics across distinct classes of networks," 2018.
[38] T. A.Jilani, U. Fatima, M. Mahmood Baig, and S. Mahmood, "A Survey and Comparative Study of Different PageRank Algorithms," Int. J. Comput. Appl., vol. 120, no. 24, pp. 24–30, Jun. 2015.
[39] University of Connecticut, "Pearson correlation table," no. April, pp. 1–2, 2015.
[40] J. Zhong, J. Wang, W. Peng, Z. Zhang, and Y. Pan, "Prediction of essential proteins based on gene expression programming.," BMC Genomics, vol. 14 Suppl 4, p. S7, 2013.
[41] M. Li, R. Zheng, H. Zhang, J. Wang, and Y. Pan, "Effective identification of essential proteins based on priori knowledge, network topology and gene expressions.," Methods, vol. 67, no. 3, pp. 325–33, Jun. 2014.
[42] S. Boulton, "Identification of a Saccharomyces cerevisiae Ku80 homologue: roles in DNA double strand break rejoining and in telomeric maintenance," Nucleic Acids Res., 1996.
[43] F. Prado and A. Aguilera, "Partial depletion of histone H4 increases homologous recombination-mediated genetic instability.," Mol. Cell. Biol., vol. 25, no. 4, pp. 1526–36, Feb. 2005.
[44] Z. Ge, D. Nair, X. Guan, N. Rastogi, M. A. Freitas, and M. R. Parthun, "Sites of acetylation on newly synthesized histone H4 are required for chromatin assembly and DNA damage response signaling.," Mol. Cell. Biol., vol. 33, no. 16, pp. 3286–98, Aug. 2013.
[45] P. K. Thakre, A. Sv, U. Golla, S. Chauhan, and R. S. Tomar, "Previously uncharacterized amino acid residues in histone H3 and H4 mutants with roles in DNA damage repair response and cellular aging.," FEBS J., vol. 286, no. 6, pp. 1154–1173, Mar. 2019.
[46] S. Yeom, J. Oh, E.-J. Lee, and J.-S. Lee, "Positive Charge of Arginine Residues on Histone H4 Tail Is Required for Maintenance of Mating Type in Saccharomyces cerevisiae.," J. Microbiol. Biotechnol., vol. 28, no. 9, pp. 1573–1579, Sep. 2018.
[47] T. Kanki et al., "Casein kinase 2 is essential for mitophagy.," EMBO Rep., vol. 14, no. 9, pp. 788–94, Sep. 2013.
[48] Y. Qiu, A. Hassaninasab, G. S. Han, and G. M. Carman, "Phosphorylation of Dgk1 diacylglycerol kinase by casein kinase II regulates phosphatidic acid production in saccharomyces cerevisiae," J. Biol. Chem., 2016.
[49] K. Matsuzaki and M. Shinohara, "Casein kinase II phosphorylates the C-terminal region of Lif1 to promote the Lif1-Xrs2 interaction needed for non-homologous end joining.," Biochem. Biophys. Res. Commun., vol. 501, no. 4, pp. 1080–1084, 2018.
[50] R. Shafaatian, M. A. Payton, and J. D. Reid, "PWP2, a member of the WD-repeat family of proteins, is an essential Saccharomyces cerevisiae gene involved in cell separation," Mol. Gen. Genet., 1996.
[51] R. Duden, M. Hosobuchi, S. Hamamoto, M. Winey, B. Byers, and R. Schekman, "Yeast beta- and beta'-coat proteins (COP). Two coatomer subunits essential for endoplasmic reticulum-to-Golgi protein traffic.," J. Biol. Chem., vol. 269, no. 39, pp. 24486–95, Sep. 1994.

[52] G. Gabriely, R. Kama, and J. E. Gerst, "Involvement of specific COPI subunits in protein sorting from the late endosome to the vacuole in yeast.," Mol. Cell. Biol., vol. 27, no. 2, pp. 526–40, Jan. 2007.
[53] M. M. Tabb, P. Tongaonkar, L. Vu, and M. Nomura, "Evidence for separable functions of Srp1p, the yeast homolog of importin alpha (Karyopherin alpha): role for Srp1p and Sts1p in protein degradation.," Mol. Cell. Biol., vol. 20, no. 16, pp. 6062–73, Aug. 2000.
[54] L. Chen and K. Madura, "Yeast importin-α (Srp1) performs distinct roles in the import of nuclear proteins and in targeting proteasomes to the nucleus.," J. Biol. Chem., vol. 289, no. 46, pp. 32339–52, Nov. 2014.
[55] M. Snyder and L. Ni, "A genomic study of the bipolar bud site selection pattern in Saccharomyces cerevisiae.," Mol. Biol. Cell, 2001.
[56] J. Bassler et al., "The conserved Bud20 zinc finger protein is a new component of the ribosomal 60S subunit export machinery.," Mol. Cell. Biol., vol. 32, no. 24, pp. 4898–912, Dec. 2012.
[57] M. B. K. Eswara, A. Clayton, and D. Mangroo, "Utp22p acts in concert with Utp8p to channel aminoacyl-tRNA from the nucleolus to the nuclear tRNA export receptor Los1p but not Msn5p.," Biochem. Cell Biol., vol. 90, no. 6, pp. 731–49, Dec. 2012.
[58] B. Albert et al., "A Molecular Titration System Coordinates Ribosomal Protein Gene Transcription with Ribosomal RNA Synthesis.," Mol. Cell, vol. 64, no. 4, pp. 720–733, 2016.
[59] T. Tsuboi et al., "The tRNA splicing endonuclease complex cleaves the mitochondria-localized Cbp1 mRNA," J. Biol. Chem., 2015.
[60] A. Kuma, N. Mizushima, N. Ishihara, and Y. Ohsumi, "Formation of the ~ 350-kDa Apg12-Apg5·Apg16 multimeric complex, mediated by Apg16 oligomerization, is essential for autophagy in yeast," J. Biol. Chem., 2002.
[61] M. Shirayama, Y. Matsui, and A. Toh-E, "The yeast TEM1 gene, which encodes a GTP-binding protein, is involved in termination of M phase.," Mol. Cell. Biol., vol. 14, no. 11, pp. 7476–82, Nov. 1994.
[62] I. Scarfone, M. Venturetti, M. Hotz, J. Lengefeld, Y. Barral, and S. Piatti, "Asymmetry of the budding yeast Tem1 GTPase at spindle poles is required for spindle positioning but not for mitotic exit.," PLoS Genet., vol. 11, no. 2, p. e1004938, Feb. 2015.
[63] N. N. Fang, A. H. M. Ng, V. Measday, and T. Mayor, "Hul5 HECT ubiquitin ligase plays a major role in the ubiquitylation and turnover of cytosolic misfolded proteins," Nat. Cell Biol., 2011.
[64] N. N. Fang and T. Mayor, "Hul5 ubiquitin ligase: Good riddance to bad proteins," Prion. 2012.
[65] T. Borggrefe, R. Davis, H. Erdjument-Bromage, P. Tempst, and R. D. Kornberg, "A complex of the Srb8, -9, -10, and -11 transcriptional regulatory proteins from yeast," J. Biol. Chem., 2002.
[66] E. Larschan and F. Winston, "The Saccharomyces cerevisiae Srb8-Srb11 complex functions with the SAGA complex during Gal4-activated transcription.," Mol. Cell. Biol., vol. 25, no. 1, pp. 114–23, Jan. 2005.
[67] G. A. Jonsdottir and R. Li, "Dynamics of yeast myosin I: Evidence for a possible role in scission of endocytic vesicles," Curr. Biol., 2004.
[68] J. Giblin, I. M. Fernández-Golbano, F.-Z. Idrissi, and M. I. Geli, "Function and regulation of Saccharomyces cerevisiae myosins-I in endocytic budding," Biochem. Soc. Trans., 2011.
[69] S. Wesche, M. Arnold, and R.-P. Jansen, "The UCS domain protein She4p binds to myosin motor domains and is essential for class I and class V myosin function.," Curr. Biol., vol. 13, no. 9, pp. 715–24, Apr. 2003.
[70] H. Toi, K. Fujimura-Kamada, K. Irie, Y. Takai, S. Todo, and K. Tanaka, "She4p/Dim1p interacts with the motor domain of unconventional myosins in the budding yeast, Saccharomyces cerevisiae.," Mol. Biol. Cell, vol. 14, no. 6, pp. 2237–49, Jun. 2003.
[71] J. Dong, R. Lai, K. Nielsen, C. A. Fekete, H. Qiu, and A. G. Hinnebusch, "The essential ATP-binding cassette protein RLI1 functions in translation by promoting preinitiation complex assembly.," J. Biol. Chem., vol. 279, no. 40, pp. 42157–68, Oct. 2004.
[72] G. Kispal et al., "Biogenesis of cytosolic ribosomes requires the essential iron-sulphur protein Rli1p and mitochondria.," EMBO J., vol. 24, no. 3, pp. 589–98, Feb. 2005.
[73] A. Alhebshi, T. C. Sideria, S. L. Holland, and S. V Avery, "The essential iron-sulfur protein Rli1 is a primary cellular target accounting for the toxicity of reactive oxygen species," Mol. Biol. Cell, 2012.
[74] Z.-R. Ruan et al., "Identification of lethal mutations in yeast threonyl-tRNA synthetase revealing critical residues in its human homolog.," J. Biol. Chem., vol. 290, no. 3, pp. 1664–78, Jan. 2015.
[75] B. Delagoutte, D. Moras, and J. Cavarelli, "tRNA aminoacylation by arginyl-tRNA synthetase: induced conformations during substrates binding.," EMBO J., vol. 19, no. 21, pp. 5599–610, Nov. 2000.
[76] C. Mann, J. M. Buhler, I. Treich, and A. Sentenac, "RPC40, a unique gene for a subunit shared between yeast RNA polymerases A and C," Cell, 1987.
[77] G. V Shpakovski and E. K. Shematorova, "Rpc19 and Rpc40, two alpha-like subunits shared by nuclear RNA polymerases I and III, are interchangeable between the fission and budding yeasts.," Curr. Genet., vol. 36, no. 4, pp. 208–14, Oct. 1999.
[78] M. F. Hoekstra, R. M. Liskay, A. C. Ou, A. J. DeMaggio, D. G. Burbee, and F. Heffron, "HRR25, a putative protein kinase from budding yeast: association with repair of damaged DNA.," Science, vol. 253, no. 5023, pp. 1031–4, Aug. 1991.
[79] K. A. Kafadar, H. Zhu, M. Snyder, and M. S. Cyert, "Kafadar, K. A., Zhu, H., Snyder, M., & Cyert, M. S. (2003). Negative regulation of calcineurin signaling by Hrr25p, a yeast homolog of casein kinase I. Genes and Development, 17(21), 2698–2708. https://doi.org/10.1101/gad.1140603Negative regulation of cal," Genes Dev., 2003.
[80] T. Schäfer et al., "Hrr25-dependent phosphorylation state regulates organization of the pre-40S subunit.," Nature, vol. 441, no. 7093, pp. 651–5, Jun. 2006.
[81] Y. Hasegawa, K. Irie, and A. P. Gerber, "Distinct roles for Khd1p in the localization and expression of bud-localized mRNAs in yeast," RNA, 2008.
[82] N. Mauchi, Y. Ohtake, and K. Irie, "Stability control of MTL1 mRNA by the RNA-binding protein Khd1p in yeast.," Cell Struct. Funct., vol. 35, no. 2, pp. 95–105, 2010.
[83] K. Matsuda, M. Makise, Y. Sueyasu, M. Takehara, T. Asano, and T. Mizushima, "Yeast two-hybrid analysis of the origin recognition complex of Saccharomyces cerevisiae:

interaction between subunits and identification of binding proteins.," FEMS Yeast Res., vol. 7, no. 8, pp. 1263–9, Dec. 2007.
[84] J. Kan, L. Zou, J. Zhang, R. Wu, Z. Wang, and C. Liang, "Origin recognition complex (ORC) mediates histone 3 lysine 4 methylation through cooperation with Spp1 in Saccharomyces cerevisiae.," J. Biol. Chem., vol. 283, no. 49, pp. 33803–7, Dec. 2008.
[85] Z. Zhang and J. C. Reese, "Molecular genetic analysis of the yeast repressor Rfx1/Crt1 reveals a novel two-step regulatory mechanism.," Mol. Cell. Biol., vol. 25, no. 17, pp. 7399–411, Sep. 2005.
[86] J. Zaim, E. Speina, and A. M. Kierzek, "Identification of new genes regulated by the Crt1 transcription factor, an effector of the DNA damage checkpoint pathway in Saccharomyces cerevisiae.," J. Biol. Chem., vol. 280, no. 1, pp. 28–37, Jan. 2005.
[87] S. M. Barabino, M. Ohnacker, and W. Keller, "Distinct roles of two Yth1p domains in 3'-end cleavage and polyadenylation of yeast pre-mRNAs.," EMBO J., vol. 19, no. 14, pp. 3778–87, Jul. 2000.
[88] A. Casañal et al., "Architecture of eukaryotic mRNA 3'-end processing machinery.," Science, vol. 358, no. 6366, pp. 1056–1059, 2017.
[89] L. Jønson, J. F. Rehfeld, and A. H. Johnsen, "Enhanced peptide secretion by gene disruption of CYM1, a novel protease in Saccharomyces cerevisiae.," Eur. J. Biochem., vol. 271, no. 23–24, pp. 4788–97, Dec. 2004.
[90] K. Raman, N. Damaraju, and G. K. Joshi, "The organisational structure of protein networks: Revisiting the centrality-lethality hypothesis," Syst. Synth. Biol., 2014.
[91] M. E. J. Newman and J. Park, "Why social networks are different from other types of networks," Phys. Rev. E - Stat. Physics, Plasmas, Fluids, Relat. Interdiscip. Top., 2003.
[92] M. W. Hahn and A. D. Kern, "Comparative genomics of centrality and essentiality in three eukaryotic protein-interaction networks.," Mol. Biol. Evol., vol. 22, no. 4, pp. 803–6, Apr. 2005.

Fatima Sughra Zaidi received the B.E. degree from NED University of Engg. and Technology, Karachi, Pakistan, in 2017 and currently completing M.E. degree from same place (2017-present) and recently working on thesis related to network analysis using graph theory. Her research interest includes data analysis, computational biology and coding in MATLAB.

Ubaida Fatima received the B.E. in Biomedical Engineering (2011) and MS in Applied Mathematics (2014) degrees from NED University of Engineering and Technology. She has been completing her PhD degree in Applied Mathematics (2015-present) at NED University of Engineering and Technology. Her research interest are Big data analysis, Social network analysis, Artificial intelligence, Data mining and Statistics.

Bilal Ahmed Usmani completed his PhD in Mathematical Epidemiology from University of Glasgow, Scotland, in 2016. He is currently an Assistant Professor in the Department of Biomedical Engineering at NED University, Karachi. His research interest includes developing mathematical and statistical models applicable to fields as diverse as rainfall forecasting, econometric analysis to infectious diseases. Current research on Hepatitis-C (HCV) shares common grounds with ongoing infectious disease modeling research happening at the Department of Population Health Sciences, Bristol Medical School and, in particular, Dr. Aaron Lim's modeling work on the HCV epidemic in low- and middle-income countries, including Pakistan.

Ali Raza Jafri received his B.E. and M.E. degrees from NED University of Engineering & Technology, Karachi, Pakistan in 1996 and 2002 respectively. He received Ph.D. degree from Beijing Institute of Technology, Beijing, China in 2007. He has been working at NED University of Engineering & Technology for last 22 years as lecturer (from 1997), Asst. Professor (from 2002), Assoc. Prof. (from 2007) and professor (from 2010). He is heading Department of Biomedical Engineering, NED University since 2011. His research interest includes Robotics, Mechatronics, Automation, Medical Device design, Materials and Biostatistics. He is a professional Engineer and life time member of Pakistan Engineering Council.