Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

Physica A 365 (2006) 556–564

PHYSICA A

www.elsevier.com/locate/physa

Residual closeness in networks

Chavdar Dangalchev

5710 122nd Avenue SE, ap. 192, Bellevue, WA 98006, USA

Available online 18 January 2006

Abstract

A new characteristic (residual closeness) which can measure the network resistance is presented. It evaluates closeness after removal of vertices or links, hence two types are considered—vertices and links residual closeness. This characteristic is more sensitive than the well-known measures of vulnerability—it captures the result of actions even if they are small enough not to disconnect the graph. A definition for closeness is modified so it still can be used for unconnected graphs but the calculations are easier.

PACS: 89.75.Fb; 87.23.Ge; 89.65.−s

Keywords: Graph vulnerability; Closeness; Connectivity; Toughness; Vertex integrity

1. Introduction

Analyzing complex networks is a very important task in designing power grid systems, national security, sociology. There are two major aspects of this analysis—how the networks are created and how stable (robust) they are. Erdos and Renyi [1], with their random graph theory, started the study of the topology of the complex networks. Next follow the small-world model of Watts–Strogatz [2]. Barabasi and Albert [3] introduce the preference and growth as elements of creating scale-free networks. Dorogovtsev and Mendes [4] have proved that aging (preference growth based on the age of vertices) can also lead to scale-free networks. Dangalchev [5] gives examples that the scale-free topology can be created without the element of growth—in a network with a static number of vertices. The creation of the complex networks is far from being fully explained.

One of the most important characteristic of the complex networks is their stability, especially when designing a network—e.g., power grid systems. On description of the network resistance, finding critical vertices or links, are written many papers. In 1970s different measures of the graph vulnerability are introduced to study different aspects (not captured by the graph connectivity) of the graph behavior after removal of vertices or links. The graph toughness is introduced in Chvatal [6] and scattering number in Jung [7]. The connection between binding number and Hamiltonian circuits is proven in Woodall [8]. The vertex integrity is introduced and compared to the above characteristics in Barefoot et al. [9]. All these characteristics

E-mail address: Dangalchev@hotmail.com.

0378-4371/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.physa.2005.12.020

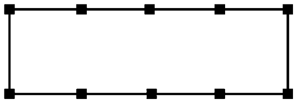

Fig. 1.

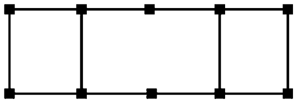

Fig. 2.

(except for the binding number) measure vulnerability by removing vertices or links so the resulting graph is disconnected. Their definitions can be found in Appendix A.

Measuring the ultimate robustness of the networks, knowing the number of vertices which have to be removed so the network will be totally destroyed (disconnected), is important. Not less important is to know how taking out one element will disturb the communications in social and biological systems, or how far from the optimal level the technological systems will perform. Our goal is to measure the vulnerability even when the actions (the removal of vertices or links) do not disconnect the graph. One possible measure could be how the closeness of the graph is changing after these actions. In this paper we introduce a new characteristic of the graph vulnerability—residual closeness. We have also calculated the residual closeness for some types of graphs to compare it with the other characteristics of graph vulnerability.

To explain better the need and advantages of the residual closeness as a measure for the graph vulnerability we will calculate the existing characteristics for the following graphs:

The graph from Fig. 2 seems to be more robust than the graph from Fig. 1, but both graphs have the same graph vulnerability characteristics: their connectivity is equal to 2, the toughness is 1, the vertex integrity is 6, the binding number is 1.5, and scattering number is 0. The residual closeness recognizes the difference between the above graphs— the graph from Fig. 1 has normalized link residual closeness equal to 0.84, while the graph from Fig. 2 has 0.88.

Centrality properties, useful tools in networks analysis, are reviewed in Freeman [10] and we will follow the definitions given there. A new definition for closeness from Latora and Marchiori [11] expands the use of this characteristic to the unconnected graphs. A definition of betweenness, based on random walk, is given by Newman in [12].

In this paper we consider simple undirected graphs, i.e., graphs without double links (links having the same pair of vertices) and without loops (a link with the same starting and ending vertex). We introduce here a modification of the definition of closeness, which as Latora and Marchiori's definition can be used for disconnected graphs. The main advantage of this definition is that it gives closed formulas for some graphs. Formulas (3) and (4), used for easier calculation of residual closeness of graphs, received by connecting two graphs, cannot be proven for Latora and Marchiori's closeness.

2. Closeness

The well-known definitions for closeness (see Ref. [10]) have one big flaw—they can only be used for connected graphs. In Latora and Marchiori [11] a new definition for point closeness is given

$$C(i) = \sum_{j \neq i} \frac{1}{d(i,j)}.$$

In the above formula $d(i,j)$ is the distance between vertices i and j. The interpretation of this definition is natural: $1/d(i,j)$ is the closeness between vertices i and j and the point closeness is the sum of closeness to all other vertices. This definition can be used for not connected graphs. For some items in the sum (when i and j are not connected) the closeness will be $1/d(i,j) = 1/\infty = 0$ but there will be items different than 0. The closeness of the graph is defined as

$$C = \sum_i C(i). \tag{1}$$

Let us see what we receive using Latora and Marchiori definition for path-graphs (n vertices connected with a line). For the closeness of the end-vertices of a path with 4 vertices we have: $C = \frac{1}{1} + \frac{1}{2} + \frac{1}{3} = \frac{11}{6}$. When $n = 8$ the end-vertex closeness will be $C = \frac{1}{1} + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{7} = \frac{1089}{420}$. It is difficult to find a formula to express the closeness in this case. If instead of closeness $1, 1/2, 1/3, \ldots,$ we use $1/2, 1/4, 1/8, \ldots,$ the calculations will be easier and again we can use this definition for not connected graphs.

Let us define closeness with

$$C(i) = \sum_{j \neq i} \frac{1}{2^{d(i,j)}}. \tag{2}$$

We use the standard definition $d(i,j)$ for distance between vertices i and j (with $d(i,i) = 0$). For the graph closeness we use formula (1).

Now the closeness of the end-vertices of a path with 4 vertices is $C = \frac{1}{2} + \frac{1}{4} + \frac{1}{8} = 1 - \frac{1}{8}$. For the path with n vertices the end-vertices have $C = 1 - (1/2^{n-1})$.

We will further illustrate how this definition makes calculations easier by calculating the graph formed from two given graphs. Let us have two graphs: G_1 and G_2. If we add the vertices of the two graphs without adding new links the graph $G = G_1 + G_2$ will have closeness equal to the sum of the two:

$$C(G_1 + G_2) = C(G_1) + C(G_2).$$

Let us connect graphs G_1 and G_2 with only one link: vertex "k" of graph G_1 with vertex "m" of G_2. For the closeness of graph G we receive (see Appendix B)

$$C(G_1 + G_2) = C(G_1) + C(G_2) + (1 + C(k))(1 + C(m)). \tag{3}$$

Let the same graphs G_1 and G_2 be connected using a new vertex "p" (vertex "p" is linked with vertex "k" of graph G_1 and vertex "m" of G_2). To calculate the closeness of this new graph we may follow the calculations in Appendix B. Here we will calculate it using formula (3) twice. Let first connect G_1 with vertex "p"

$$C(G_1 + p) = C(G_1) + (1 + C(k)).(1 + 0).$$

Now vertex p has closeness (within the graph $G_1 + p$) equal to $C(p) = 1/2 + C(k)/2$

$$C(G_1 + p + G_2) = C(G_1) + 1 + C(k) + C(G_2) + \left(1 + \frac{1 + C(k)}{2}\right)(1 + C(m)),$$

or finally

$$C(G_1 + p + G_2) = C(G_1) + C(G_2) + 2 + C(k) + C(m) + \frac{(1 + C(k))(1 + C(m))}{2}. \tag{4}$$

We will calculate the closeness for some graphs. An isolated graph (graph with n not connected vertices) has closeness 0—this is the reason in formula (2) to put the restriction $i \neq j$. The complete graph with n vertices (K_n) has closeness equal to $(n(n-1))/2$. The star graph (S_n) with n vertices (connected only to one, central vertex) has closeness

$$C(S_n) = \frac{n-1}{2} + (n-1)\left(\frac{1}{2} + \frac{n-2}{4}\right) = \frac{(n-1)(n+2)}{4}.$$

The closeness for the path (L_n) with n vertices is (see Appendix C)

$$C(L_n) = 2n - 4 + \frac{1}{2^{n-2}}.$$

For the cycle (C_n) with n vertices the closeness is

$$C(C_n) = 2n\left(1 - \frac{1}{2^{[n/2]}}\right),$$

where $[n/2]$ is the whole part of n after division by 2.

3. Definition of residual closeness

Let $d_k(i,j)$ be the distance between vertices i and j in the graph, received from the original graph where all links of vertex k are deleted. We can use formulae (1) and (2) to calculate the closeness after removing vertex k

$$C_k = \sum_i \sum_{j \neq i} \frac{1}{2^{d_k(i,j)}}.$$

The vertex residual closeness (VRC) of the graph can be defined as

$$R = \min_k \{C_k\}. \tag{5}$$

We can define analogously the graph link residual closeness (LRC)—the minimal of the closeness, calculated when only one link (k,p) is deleted

$$R = \min_{(k,p)} \{C_{(k,p)}\}. \tag{6}$$

To calculate it we use $d_{(k,p)}(i,j)$—the distance between vertices i and j in the graph, received from the original graph where only the link (k,p) is deleted.

To normalize the residual closeness we can divide it by the original closeness C

$$R' = R/C. \tag{7}$$

This normalization formula contains information about the original graph. After this point we will use the normalized residual closeness (7) instead of residual closeness (5) and (6). Different ways for normalization of the residual closeness can be found in the conclusion.

4. Residual closeness and closeness, degree, connectivity, betweenness

The example below shows that the vertex with the maximal closeness not always has the minimal residual closeness.

The vertex 1 has closeness equal to 2.75 (the biggest in the graph) and has normalized residual closeness 0.6452 (the residual closeness is 10) while the vertex 2 has the second closeness 2.5 and the smallest normalized residual closeness 0.5484 (residual closeness 8.5).

This graph can be an example, illustrating the same statement if we use Latora and Marchiori [11] definition of closeness. The values in this case for closeness, residual closeness and normalized residual closeness for vertex 1 are $5\frac{1}{2}$, $20\frac{1}{3}$ and 0.6489; 5, 17 and 0.5426 for vertex 2.

The graph degree also cannot be an indication for the residual closeness—vertex 1 of the graph from Fig. 3 has degree 5 (highest in the graph) while the vertex with the lowest residual closeness (vertex 2), has only 4.

Residual closeness is not as closely related to connectivity as it seems. The following example uses normalized link residual closeness.

If we delete the link $(7,8)$ from Fig. 4 we receive a residual closeness 0.8139 (the minimal residual closeness in the graph) and the graph is still connected. If we delete the link $(1,14)$ we receive a residual closeness 0.9183, the graph is disconnected (vertex 14 is isolated).

The corresponding numbers for the Latora and Marchiori definition are 0.8698 and 0.9006. An example for vertex residual closeness, similar to the graph from Fig. 4, can be constructed.

The residual closeness is not closely linked to betweenness either. The graph with 11 vertices in Fig. 5 illustrates this.

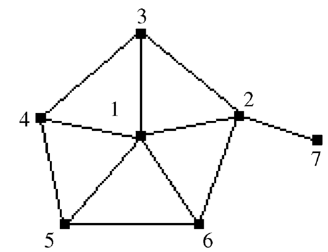

Fig. 3.

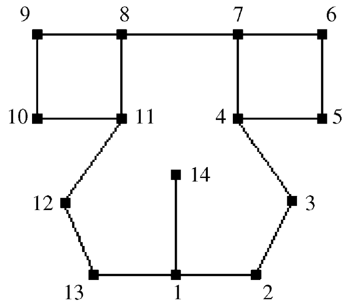

Fig. 4.

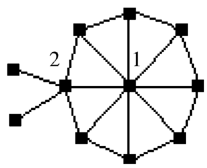

Fig. 5.

Deleting all links of vertex 2 will result in the smallest residual closeness 0.6029, while after deleting the links of vertex 1 the residual closeness will be 0.6066. Vertex 1 has the highest betweenness in the graph 0.480 and the highest random-walk (by Newman [12]) betweenness—0.552. Vertex 2 is second with 0.389 and 0.533.

5. Calculation of the residual closeness

The VRC of the star graph is 0—deleting the links of the hub will result in n isolated vertices.

The closeness of the complete graph with n vertices is $(n(n-1))/2$ and it is reduced to residual closeness equal to $((n-1)(n-2))/2$. The normalized vertex residual closeness (NVRC) is $(n-2)/n$. This number can become arbitrary close to 1. Hence NVRC belongs to the interval: [0,1).

The LRC of a graph with only one link and $n-2$ isolated vertices is 0—deleting the link will result in n isolated vertices. Deleting a link from a complete graph with n vertices will result in residual closeness equal to $(n(n-1)-1)/2$. The normalized link residual closeness (NLRC) is $1-1/(n(n-1))$, number close to 1. Hence NLRC belongs to the interval: [0,1).

The lower limit of NLRC (0) can be reached only when the resulting graph has isolated vertices. If the resulting graph is not completely disconnected the lowest link residual closeness is much higher. Let us connect two star graphs, each with number of vertices "k", with only one link. The resulting graph we will be denoted by $S_{k,k}$. The residual closeness of $S_{k,k}$ is

$$R(S_{k,k}) = 2\,\frac{(k-1)(k+2)}{4} = \frac{1}{2}\,(k^2 + k - 2).$$

The closeness of $S_{k,k}$ can be calculated using formula (3)

$$C(S_{k,k}) = \frac{1}{2}(k^2 + k - 2) + \left(1 + \frac{k-1}{2}\right)^2 = \frac{1}{4}\,(3k^2 + 4k - 3).$$

Finally for NLRC of $S_{k,k}$ we receive

$$\frac{R}{C}\,(S_{k,k}) = \frac{2k^2 + 2k - 4}{3k^2 + 4k - 3}.$$

For $S_{k,k+1}$ we can analogously receive

$$\frac{R}{C}\,(S_{k,k+1}) = \frac{2k^2 + 4k - 2}{3k^2 + 7k}.$$

Both NLRCs increase to $\frac{2}{3}$ with the increasing of k. The NLRC for connected graphs with $2k$ vertices belongs to the interval $[\frac{(2k^2+2k-4)}{(3k^2+4k-3)}, 1)$ or the interval $[\frac{(2k^2+4k-2)}{(3k^2+7k)}, 1)$ for $2k+1$. For not disconnected graphs with big number of vertices this interval is $(\frac{2}{3}, 1)$.

Some questions arise: is the lower limit of NLRC equal to $\frac{2}{3}$ when the resulting graph has two disconnected parts? What is the lower limit when the resulting graph is connected?

6. Comparison between different graph vulnerability characteristics

Let us connect two complete graphs with a number of vertices "k" and "m" with only one link. The resulting graph we will be denoted by $K_{k,m}$. We can calculate the residual closeness after deleting the connecting link

$$R(K_{k,m}) = R(K_k) + R(K_m) = \frac{k(k-1)}{2} + \frac{m(m-1)}{2}.$$

Using formula (3) we can calculate the closeness of the $K_{k,m}$ graph

$$C(K_{k,m}) = \frac{k(k-1)}{2} + \frac{m(m-1)}{2} + \left(1 + \frac{k-1}{2}\right)\left(1 + \frac{m-1}{2}\right)$$

For the normalized residual closeness of $K_{k,k}$ graph we receive

$$\frac{R}{C}\,(K_{k,k}) = k(k-1)\Big/\left(k(k-1) + \frac{1}{4}\,(k+1)^2\right) = \frac{4k^2 - 4k}{5k^2 - 2k + 1}. \tag{8}$$

When k increases then the normalized residual closeness of $K_{k,k}$ graph increases and has an upper bound 0.8. Calculating the normalized residual closeness of $K_{1,k}$ we receive

$$\frac{R}{C}\,(K_{1,k}) = k(k-1)/(k(k-1) + k + 1) = \frac{k^2 - k}{k^2 + 1} \tag{9}$$

which is close to 1.

Comparing the results of (8) and (9) we can see that deleting the connecting link from the graph $K_{k,n}$ when $k = 1$ is less damaging than deleting it when $k = n/2$. This shows that the residual closeness is very sensitive to the changes of k (like the vertex integrity while connectivity, toughness, and binding number do not respond properly to the changes of k). Even more—only the residual closeness satisfies the condition (iii) from Barefoot, Entringer and Swart [9]

(iii) If H is a proper subgraph of G then $R(H) < R(G)$. (10)

The sensitivity to the changes of k, satisfying condition (iii) and ability to capture the results of actions small enough not to disconnect the graph show that the residual closeness is a very good measure for graph vulnerability.

All other characteristics of the graph vulnerability are in the interval $[0, n]$. To compare better the residual closeness with them we can calculate $n\frac{R}{C}$

$$n\,\frac{R}{C}\,(K_{1,n-1}) = n\,\frac{(n-1)^2-(n-1)}{(n-1)^2+1} = n\,\frac{n^2-3n+2}{n^2-2n+2} = n-1-\frac{1}{n}\,\Delta,$$

$$n\,\frac{R}{C}\,(K_{\frac{n}{2},\frac{n}{2}}) = \frac{4}{5}\,n-\frac{24}{25}-\frac{1}{n}\,\Delta,$$

where $\Delta \ll n$.

The definition of $K_{k,m}$ is a little bit different from the definition of $G_{k,m+k}$, used in Barefoot, Entringer and Swart. In Ref. [9] the two complete graphs are connected with a new vertex and two links. A similar result (sensitivity to k for $G_{k,m+k}$) will be received using this definition and calculating VRC (instead of using formula (3) we will apply formula (4)).

7. Conclusions

In this paper we have introduced the residual closeness—a characteristic for graph vulnerability. The residual closeness is more sensitive than the other measures of vulnerability (like graph toughness, scattering number, vertex integrity). It captures the result of actions (deleting vertices or links) even if they are small enough not to disconnect the graph. The residual closeness is monotonous—it satisfies condition (iii) from Barefoot, Entringer, Swart [9], which is not satisfied by any of the graph vulnerability characteristics considered in Ref. [9]. The residual closeness is not so closely related to connectivity, degrees, closeness or betweenness as it seems—the given examples show that the vertex supplying the smallest residual closeness may not be the one with the highest degrees, closeness, betweenness or the one maintaining the same connectivity of the graph.

In this paper we modify the definition of closeness—it can be used for unconnected graphs and it is easier for calculations. With this definition of closeness and formulae (3) and (4) the residual closeness (vertex and link) of graphs $K_{k,n}$, $S_{k,n}$ and other examples from Barefoot, Entringer, and Swart can be easily calculated for comparison with the other characteristics for graph vulnerability.

We can consider not only VRC and LRC, but also relative residual closeness (the residual closeness divided by the number of vertices or links) or average residual closeness. We can define different ways to normalize the residual closeness (and use them to compare graphs with a different number of vertices or links, or compare them to different graph characteristics). A different (than the one used in this paper) way to normalize the residual closeness is to divide it by the maximal possible residual closeness of the graph with the same number of vertices. If after removing a vertex or a link the result is a complete graph with $n-1$ vertices then the residual graph closeness is $(n-1)(n-2)/2$. For the normalized residual closeness we can use

$$R' = 2R/(n-1)(n-2).$$

Another way to normalize the residual closeness is to use $(C-R)/C$ or $(C-R)/R$.

Acknowledgment

The author would like to thank the anonymous referee for the comments and suggestions.

Appendix A. Different measures of graph vulnerability

For a graph G with vertices V we denote the number of vertices by $n(G)$, the number of connected components of G by $c(G)$, and the maximum number of vertices in a component of G by $m(G)$.

$$\text{Separator S}: \; S \subset V \text{ and } c(G[V \backslash S]) > 1.$$

$$\text{Connectivity}: \; k(G) = \min |S|.$$

Toughness : $t(G) = \min\{|S|/c(G[V\backslash S])\}$.

Scattering number : $s(G) = \max\{c(G[V\backslash S]) - |S|\}$.

Vertex integrity : $i(G) = \min\{|S| + m(G[V\backslash S])\}$.

In the above definitions S is a separator of G.

Binding number : $b(G) = \min\{|N(V\backslash W)|/|W|\}$,

where all vertices of V/W are adjacent to at least one vertex of subset W of V.

Appendix B. Calculation of the closeness of two graphs

Let us connect graphs G_1 and G_2 with only one link: vertex "k" of graph G_1 with vertex "m" of G_2.

$$
\begin{aligned}
C(G) = \sum_i \sum_{j \neq i} \frac{1}{2^{d(i,j)}} &= \sum_{i \in G_1} \sum_{i \in G_1, j \neq i} \frac{1}{2^{d(i,j)}} + \sum_{i \in G_2} \sum_{i \in G_2, j \neq i} \frac{1}{2^{d(i,j)}} + 2 \sum_{i \in G_1} \sum_{j \in G_2} \frac{1}{2^{d(i,j)}} \\
&= C(G_1) + C(G_2) + 2 \sum_{i \in G_1} \sum_{j \in G_2} \frac{1}{2^{d(i,k)+1+d(m,j)}} \\
&= C(G_1) + C(G_2) + \sum_{i \in G_1} \frac{1}{2^{d(i,k)}} \sum_{j \in G_2} \frac{1}{2^{d(m,j)}} \\
&= C(G_1) + C(G_2) + \left(\sum_{i \neq k} \frac{1}{2^{d(i,k)}} + \frac{1}{2^{d(k,k)}} \right) \left(\sum_{j \neq m} \frac{1}{2^{d(m,j)}} + \frac{1}{2^{d(m,m)}} \right).
\end{aligned}
$$

Finally, for the closeness of graph G, formed by connecting with only one link graphs G_1 and G_2, we receive formula (3)

$$C(G_1 + G_2) = C(G_1) + C(G_2) + (1 + C(k))(1 + C(m)).$$

Similar calculations can not be done using the definition of Latora and Marchiori.

Appendix C. Calculation of the path–graph closeness

Let us calculate the closeness for the path (L_n)

$$
\begin{aligned}
C(L_n) = \sum_{i=1}^{n} \sum_{j=1, j \neq i}^{n} \frac{1}{2^{d(i,j)}} &= \sum_{i=1}^{n} \sum_{j=1, j \neq i}^{n} \frac{1}{2^{|i-j|}} \\
&= 2(n-1)\frac{1}{2^1} + 2(n-2)\frac{1}{2^2} + \cdots + 2.2\frac{1}{2^{n-2}} + 2.1\frac{1}{2^{n-1}}.
\end{aligned}
$$

To calculate the above sum we start from the equation

$$1 + X + X^2 + \cdots + X^{n-1} = \frac{X^n - 1}{X - 1}.$$

Differentiating both sides of the equation

$$1 + 2X + 3X^2 + \cdots + (n-1)X^{n-2} = \frac{nX^{n-1}}{X-1} - \frac{X^n - 1}{(X-1)^2},$$

substituting X with 2:

$$1 + 2.2 + 3.2^2 + \cdots + (n-1).2^{n-2} = n.2^{n-1} - 2^n + 1$$

and dividing both sides by 2^{n-2} we receive

$$C(L_n) = 2n - 4 + \frac{1}{2^{n-2}}.$$

References

[1] P. Erdos, A. Renyi, Publ. Math. Inst. Hung. Acad. Sci. 5 (1960) 17.
[2] D.J. Watts, S.H. Strogatz, Collective dynamics of 'small-world' networks, Nature 393 (1998) 440.
[3] A.-L. Barabasi, R. Albert, H. Jeong, Mean-field theory for scale-free random networks, Physica A 272 (1999) 173.
[4] S.N. Dorogovtsev, J.F.F. Mendes, Evolution of networks with aging of sites, Phys. Rev. E 62 (2000) 1842.
[5] Ch. Dangalchev, Generation models for scale-free networks, Physica A 338 (2004) 659.
[6] V. Chvatal, Tough graphs and Hamiltonian circuits, Discrete Math. 5 (1973) 215.
[7] H.A. Jung, On a class of posets and the corresponding comparability graphs, J. Combin. Theory B 24 (1978) 125.
[8] D.R. Woodall, The binding number of a graph and its Anderson number, J. Combin. Theory B 15 (1973) 225.
[9] C.A. Barefoot, R. Entringer, H. Swart, Vulnerability in graphs—a comparative survey, J. Combin. Math. Combin. Comput. 1 (1987) 13.
[10] L.C. Freeman, Centrality in social networks: conceptual clarification, Social Networks 1 (1979) 215.
[11] V. Latora, M. Marchiori, Efficient behavior of small-world networks, Phys. Rev. Lett. 87 (2001) 19.
[12] M.E.J. Newman, A measure of betweenness centrality based on random walks, arXiv:cond- mat/0309045 v1., 2004.