2006 IEEE 24th Convention of Electrical and Electronics Engineers in Israel

Dynamic Measure of Network Robustness

Yaron Singer
School of Computer Science
Tel Aviv University
yaronsin@post.tau.ac.il

Abstract— **In this paper the Dynamic Network Robustness (DYNER) method is introduced for the benefit of network robustness measurement. Its uniqueness is derived from consideration of backup efficiency in the network in instances of node deletion. By initiating artificial node deletions in the graph and calculating their effect on functionality, DYNER enables analysis of network robustness while considering backup in the network. Aside from discussion of its properties, results of application of DYNER on communication networks and synthetically generated networks with power-law degree distribution topology are presented in this work.**

I. Introduction

Study of network topology receives great attention in various research fields such as Physical Sciences [1], [2] Biological Sciences [3], [4], Computer Science [5], [6], [7] and Mathematics [8], [9], [10]. Extensive exploration of the network's topological properties contributes to comprehension of its functionality [3] and other inherent traits specific to the network studied [4]. Finding an appropriate metric for determining resiliency in power-law networks remains an open problem [11], and the study of network robustness receives great attention in all fields of network research.

Study of the Internet's resiliency under deliberate attacks and failures has experienced a growing interest in recent years. Such study has split into two branches; one focuses on the stability of Internet protocols in case of error and failures [12], [13] and the other, which receives attention outside the computer networking community, deals with the Internet's resiliency to attacks and failures on strategic locations [14], [15], [16]. Recently, Dolev *et al.* [17] reexamined the Internet resiliency to attacks taking into account the BGP valley free routing rules and have provided measures of stability unique to the Internet AS graph.

In Systems Biology, the difficulties of modeling complex cellular networks leaves the relationship between the network's topology, and function as an open question [18], [19], [20]. Specifically, the unavailability of mechanistic detail and kinetic parameters imposes difficulties on dynamic mathematical modeling of complex cellular networks. In contrast, the network topology is known in many cases and thus enables structural analysis which sheds light on key aspects of network properties such as functionality and robustness from its structure alone [4].

An important aspect of networks' analysis is their study under changing conditions. As networks are subject to changes which effect their function and performance, understanding of the effects due to such changes is a vital component of their analysis. Luscombe *et al.* have studied the Saccharomyces cerevisiae genomic regulatory network under changing biological conditions [21], to draw conclusions regarding the static network itself. The method introduced here, uses changing conditions of *node deletions* to obtain topological insight about the network. Though the likelihood of such an event varies from one network to another, in instances where network functionality is determined largely by path distances and connectivity of nodes, node deletion is a crucial event.

The Dynamic Network Robustness (DYNER) method suggested here, measures topological robustness of networks modeled by either directed or undirected, weighted or unweighed graphs. Its uniqueness derives from its consideration of backup efficiency in the network in instances of node deletion. By initiating artificial node deletions in the graph and calculating their effect on functionality, DYNER enables analysis of network robustness.

The rest of this paper is organized as follows. The next section discusses the topological measurement of functionality in networks and the concept of backup. Section III introduces the DYNER method, its motivation and some of its properties. Application on communication networks is presented in section IV. Finally, conclusions and future work are discussed in section V.

II. Functionality and Backup in Networks

This section discusses fundamental elements of the DYNER method. The principle idea behind DYNER is measurement of the effect of node failures on the network's functionality while considering backup. For this purpose, the key elements of topological centrality and backup vertices are presented in this section.

A. Topological Measure of Network Functionality

Topological properties are frequently used to measure networks' functionality, which in turn depends on path length and connectivity. For example, according to findings of Jeong *et al.* [22] the most highly connected proteins in the Saccharomyces cerevisiae are the most important for its survival. In the interest of applying an appropriate measure which enables classification of nodes by these criteria the centrality function below is used to determine a node's topological centrality. For a network $G = (V, E)$, the centrality function is simply an inverse of the average distances from a node $u \in V$:

$$\psi(u) = \frac{|V|}{\sum_{w \in V} \delta(u, w)} \tag{1}$$

Where $\delta(u, w)$ denotes the shortest distance between u and w with the following conventions:

1) $\delta(u, u) = 0 \quad \forall u \in V$;
2) $\delta(u, w) = \infty \iff w$ cannot be reached from u;
3) $\frac{c}{\infty} = 0$ for any constant c.

With appropriate calculation of δ, it is possible to apply ψ on vertices in both directed and undirected graphs, either weighted or unweighed.

Similar centrality functions have been used in the past to capture functionality in the network. In [14] the average distance between nodes in the network has been used to indicate functionality of the Internet under attacks and failures. For a node $u \in V$, note that $\psi(u)$ depends on the number of nodes connected to u and their distances from it; ψ monotonically increases with respect to rise in closeness of nodes to which it is connected. If there is some node $w \in V$ which cannot be reached by u, then $\psi(u) = 0$. Thus, this centrality measure penalizes instances of disconnection in the network.

B. Backup Vertices

Topological centrality is often used as an indication of the node's contribution to the network's functionality. To emphasize the significance of backup in context of network functionality the graph illustrated in figure 1 is presented. In this discussion v_k is represented by the vertex labeled k in the figure $\forall k \in \{1, \ldots, 8\}$. While v_1 has the highest centrality value in the graph ($\psi(v_1) = 0.8$), its extraction from the graph does not alter connectivity and shortest paths in the graph, and therefore does not effect centrality. In contrast, extraction of v_6 which has lower centrality value in the graph ($\psi(v_6) = 0.72$) would have a significant effect on centrality of other vertices since it is a disconnecting vertex in G. This seemingly counter intuitive fact can be reasoned by examining backup in the network; v_1 is fully backup up by v_5, whereas this is not the case for v_6 - all paths to v_7 and v_8 must pass through v_6, and thus it is not fully backed up.

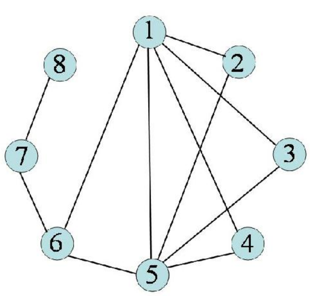

Fig. 1. Illustration of a graph in which v_1 has higher centrality than v_6 though the graph is less sensitive to its extraction.

The example above sheds light on the significance of backup in the study of network robustness. Denote C_v as v's set of direct children: $C_v = \{w | (v, w) \in E\}$. The following definition formalizes the intuitive concept of backup.

DEFINITION 1 (BACKUP VERTEX OF ORDER K): Let $G = (V, E)$. For $v \in V$ with a set of neighbors $C_v \neq \phi$, say that v has a *backup vertex of order k*, $k \leq |C_v|$, if there is a vertex $b \in V \backslash \{v\}$, such that b is connected to k of v's neighbors.

The vertex v_1 in figure 1 has a backup vertex of order $|C_{v_1}| = 5$, and v_6 has a backup vertex of order $2 < |C_{v_6}| = 3$.

In measuring the backup order of a vertex, one must also consider efficiency of reaching the backup vertex by other vertices in the graph. The focus of this work is on undirected graphs and thus reachability of the backup vertex by other vertices is equivalent to its order. In case of directed graphs it is possible to allow another parameter for the definition of backup which considers the efficiency of reaching the backup vertex. Note that it is also possible to extend the definition of backup vertex of order k to backup set of vertices, which is simply a set of backup vertices. It is easy to show that a backup set of vertices can be translated to an equivalent backup vertex by adding another vertex $b \notin V$ to the graph and re-directing edges from vertices of the backup set to b.

The connection between a vertex and its backup in the graph is a prominent one. The relation R_k, defined $\mathrm{R}_k = \{\langle u, w \rangle | \text{u is a backup of order of k of w}\}$ is symmetric. That is, u is a backup vertex of order k of $w \iff w$ is a backup vertex of order k of u.

III. DYNAMIC NETWORK ROBUSTNESS MEASURE

Having discussed topological centrality and backup which are its foundations, the DYNER measure is introduced in this section. To clarify the rational behind it, the following construction stages are presented.

For a given vertex $v \in V$, the effect of v's extraction on centrality of some vertex $u \in V$ in the network, is measured by:

$$\frac{|V|}{\sum_{w \in V} \delta(u, w)} - \frac{|V|}{\sum_{w \in V} \delta_v(u, w)} \tag{2}$$

Where $\delta_v(u, v)$ denotes the shortest distances in G after v is extracted from the graph. Such extraction simulates failure of v in the network. Note that the value of (2) monotonically increases with respect to increase in number of shortest paths in the graph from u to all vertices which pass through v. For brevity denote V_v to be $V \backslash \{v\}$. Effect of v's extraction on all vertices in V is measured by:

$$\sum_{u \in V_v} \frac{|V|}{\sum_{w \in V_v} \delta(u, w)} - \frac{|V|}{\sum_{w \in V_v} \delta_v(u, w)} \tag{3}$$

To establish a metric which measures the effect of extractions, for $P(v)$ denoting the probability of v's failure, the following measure is obtained:

$$\sum_{v \in V} P(v) \cdot \left[\sum_{u \in V_v} \frac{|V|}{\sum_{w \in V_v} \delta(u, w)} - \frac{|V|}{\sum_{w \in V_v} \delta_v(u, w)} \right] \tag{4}$$

Without prior knowledge of the probability for v's failure, the metric given here assumes that all nodes in the network are likely to have the same probability of failure, and therefore $P(v) = \frac{1}{|V|} \quad \forall v \in V$. Finally, the inverse is used in order to receive a metric which increases with respect to rise in robustness of the network. Thus, the DYNER measure for a graph $G = (V, E)$, denoted $\Gamma(G)$ is:

$$\Gamma(G) = \left[\sum_{v \in V} \sum_{u \in V_v} \frac{1}{\sum_{w \in V_v} \delta(u, w)} - \frac{1}{\sum_{w \in V_v} \delta_v(u, w)} \right]^{-1}$$

In instances where probability of failure is known, Γ may be adjusted in a manner which reflects the network's true robustness. Note that it is possible for Γ to return ∞. This indicates that the backup in the network is such that no single node failures would effect its functionality. For more complicated analysis it is possible to initiate multiple node extractions. Note however, that exhaustive examination of set extractions increases computational complexity of implementation. The algorithm below formally describes a simple implementation of the DYNER measure on unweighed graphs. In this description $\bar{\delta}$ represents the vector returned by the BFS algorithm which holds shortest distances from a source vertex to all other vertices in the graph.

Algorithm 1 Dynamic Network Robustness measure algorithm for unweighed graphs.

```
DYNER (G):
Γ' ← 0
for all v ∈ V do
    φ ← 0
    φ_v ← 0
  for all u ∈ V\{v} do
    δ̄ ← BFS(u, G)
    ψ' ← 0
    for all w ∈ V\{v, u} do
      ψ' ← ψ' + δ(u, w)
    φ ← φ + 1/ψ'
  G_v ← EXTRACTNODE(v, G)
  for all u ∈ V\{v} do
    δ̄_v ← BFS(u, G_v)
    for all w ∈ V\{v, u} do
      ψ'_v ← ψ'_v + δ(u, w)
    φ_v ← φ_v + 1/ψ'_v
  Γ' ← Γ' + φ − φ_v
Γ ← 1/Γ'
return Γ
```

The running time for the algorithm described is $O(|V|^2 \cdot |E|)$. For weighted graphs, one can substitute the BFS algorithm for a single-source shortest path algorithm for non-negative weighted graphs, such as Dijkstra's algorithm [23] and obtain $O(|V| \cdot (|V| \cdot \log |V| + |E|))$ running time. In work done by Demetrescu *et al.* [24] a deterministic oracle is suggested which enables shortest paths queries in instances of node-link failures in $O(1)$ time. Such an oracles can be constructed in $O(|E| \cdot |V|^2 + |V| \cdot \log |V|)$ time using $O(|V|^2 \cdot \log |V|)$ space. In their work, an oracle which in constructed in $O(|E| \cdot |V|^{1.5} + |V|^{2.5} \cdot \log |V|)$ and $O(|V|^{2.5})$ space is also presented. Such an oracle can be used to compute DYNER on weighted directed graphs at $O(|V| \cdot (|E| \cdot |V|^{1.5} + |V|^{2.5} \cdot \log |V|))$ time, as all distance queries are answered in $O(1)$ time by the oracle.

By simulating extractions in the graph and measuring reaction to such extractions through centrality measures, DYNER enables determining the efficiency of the network in instances of node failures. To provide complete analysis of the DYNER measure the following theorem provides formal proof of its qualities regarding backup.

Theorem 1: Let $G = (V, E)$ be a connected undirected graph. Then, $\Gamma(G)$ monotonically increases with respect to rise in backup in G.

Proof: Let $v \in V$ be some vertex which has a backup vertex b of order $0 < k < |C_v|$. Thus, there is some vertex $w \in C_v$ for which $\delta(v, w) = 1 < \infty$ and some vertex $u \in V$ for which $\delta(u, w) = \delta(u, v) + 1$ (specifically this is true for all $u \in C_v$), and no other path leads u to w. Since w cannot be reached by any other vertices in V, it follows that $\delta_v(u, w) = \infty \quad \forall u \in V \backslash \{v\}$.

Let $e \notin E$ be some edge which connects b to w. Thus, in $G' = (V, E \bigcup \{e\})$ v has a backup vertex of order $k + 1$. To prove the above it is enough to show that $\Gamma(G) \leq \Gamma(G')$.

Denote $\delta'(u, w)$ and $\delta'_v(u, w)$ to be the shortest distances between u and w in G' and shortest paths which bypass v in G' respectively. As w can be reached through b in G', clearly $\delta'_v(u, w) \leq \delta_v(u, w) = \infty \quad \forall u \in V \backslash \{v\}$. As value of (2) in G' is thus smaller than that in G, it follows that $\Gamma(G) \leq \Gamma(G')$. □

As discussed in section II backup sets can be used equivalently. For a backup set B, the proof above can be generalized with e connecting any vertex $x \in V$, and define the backup set as $B \bigcup \{x\}$.

As it monotonically increases with respect to increase in backup order, DYNER enables measuring network robustness while considering backup. The following section shows DYNER 's application of various network topologies and discusses this point further.

IV. Application on Communication Networks

This section displays application of the DYNER method on communication networks of various sizes and topologies. The networks shown here are IP networks taken from real measurements and networks synthetically generated using a network generator which generates networks with power-law degree distribution topologies.

The first data set used has been obtained from the DIMES [25] project from week 29 of 2006. The data used is of

Internet Autonomous Systems (ASes) subnetworks. The Internet consists of approximately 20,000 subnetworks, each with unique administrative management which are the ASes. The DYNER method has been applied on various AS subnetworks of different sizes on the IP level. Every unique IP address is represented by a node and links between IP addresses are represented by edges. The IP networks inside an AS are modeled by an undirected graph, unlike the AS graph where routing conforms to BGP policies [26]. In the data obtained not all vertices were connected. Therefore, in instances of disconnection in the network represented in the data, DYNER was computed on each connected component in the network and the sum over the entire connected component is considered the Γ value of the network.

Figure 2 shows application different IP networks. It is clear that the DYNER measure increases with respect to increase in size of the network.

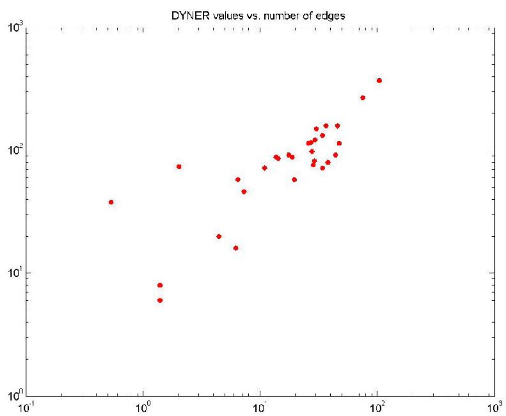

Fig. 2. DYNER values against number of edges applied on IP networks on a log-log scale.

More interesting analysis is the DYNER measure compared against node-edge ratios in the networks in figure 3. This comparison allows normalization of DYNER , as larger networks with more edges are normalized against the number of nodes. It is interesting to see that the incline is steep, and while there is a natural incline with relation to node-edge ratio in the network, there are networks where there are fewer edges per nodes, though more sustainable to node failures. It can be safely assumed that these instances are manifestations of the network topology and backup.

The DYNER measure has also been applied on synthetically generated networks with power-law degree distribution topologies. The Tel Aviv Network Generator (TANG) [27] generates networks according to models which are extensions to the Barabási-Albert model [28], and models power-law degree distribution topologies. Measurements have been conducted on 600 TANG networks, where number of nodes in the network has been chosen at random.

The result is shown in figure 4 where DYNER values are shown against network size. The number of edges in each

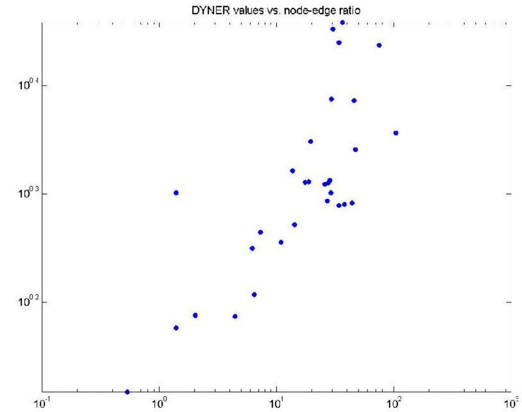

Fig. 3. DYNER values against node-edge ratio applied on IP networks on a log-log scale.

network has been chosen according to the default used in TANG, which uses a 1:4 node edge ratio. Clearly, as TANG generates networks according to strict topological criteria, the backup in the networks remains proportional to their size.

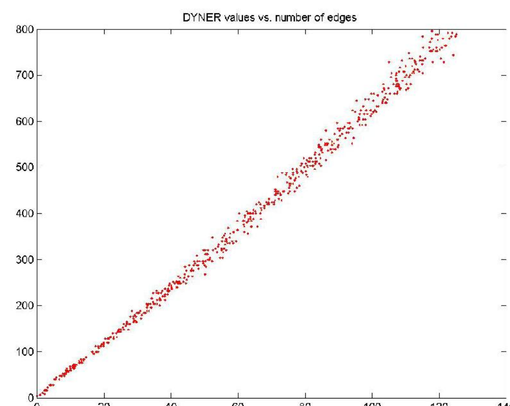

Fig. 4. DYNER values against number of edges on TANG networks where the average degree is set at 4.

Further study has been conducted on TANG networks where the average degree has also been chosen at random from discrete values of 1-4. Figure 5 displays results on the networks. This example shows DYNER measures following four separate slopes. Each one of these originates from a different node-edge ratio in the networks generated. It is clear that as the average degree increases, so does the resiliency of the network, and therefore it is well shown in this example that the DYNER metric captures robustness. The common topological patterns which are generated by TANG account for the steady monotonic rise in robustness in this example as well.

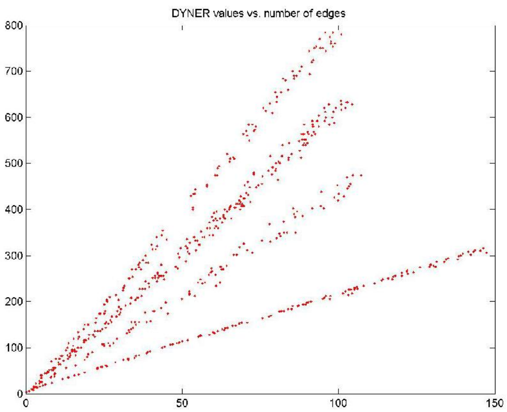

Fig. 5. DYNER values against number of edges on TANG networks where the average degree is randomly chosen between possible values of 1-4.

V. Conclusions and Future Work

This paper introduced DYNER - a new method to measure network robustness which examines networks' resiliency to node failures while considering backup. The elements of centrality and backup which are key elements in DYNER were introduced. Formal discussion regarding the DYNER properties has shown that it indeed monotonically increases with the respect to increase in the backup order in the network.

Regarding its application on IP networks and communication networks, it has been shown that DYNER is greatly affected by the network's topology. Experimentation has also been conducted on TANG networks which model power-law degree distribution networks of different sizes. It has been shown that the generation of similar topological patterns of networks of different sizes indeed produces networks which are resilient to node failures as a function of both their size in nodes and average degree.

For the future, further study on networks of different topologies is to be performed. Another interesting aspect would be to apply the DYNER measure with various probabilities of node deletion which could model different instances of node failure in various networks, rather than random failures as have been studied here. Other interesting implementations are ones where shortest distances are calculated according to minimum delay and minimum hop using valley-free paths.

Acknowledgment

The author wishes to express sincere gratitude to Dima Feldman, Anat Halpern, Ela Malinowski, Ohad Serfaty and Eran Shir of the DIMES lab in Tel Aviv University for all their great help and support. The author also wishes to thank Yuval Shavitt for stimulating discussions and encouragement.

References

[1] Doye, J. P. K. "Network Topology of a Potential Energy Landscape: A Static Scale-Free Network". Phys. Rev. Lett. 88, 238701 (2002).

[2] Scala, A., Amaral, L. A. N., and Barthelemy, M. "Small-World Networks and the Conformation Space of a Short Lattice Polymer Chain". Europhys. Lett. 55, 594600 (2001)

[3] Amitai G., Shemesh A., Sitbon E., Shklar M., Netanely D., Venger I., Pietrokovski S. "Network Analysis of Protein Structures Identifies Functional Residues". Journal of Molecular Biology, vol. 344, no. 4, pp. 113546, 2004.

[4] Stelling J., Klamt S., Bettenbrock K., Schuster S., Gilles E.D. "Metabolic Network Structure Determines Key Aspects of Functionality and Regulation" Nature. Nov. 2002. 190-193.

[5] Awerbuch B., Shavitt Y. "Topology Aggregation for Directed Graphs". IEEE/ACM Transactions on Networking, 9(1):82–90, Feb 2001. D.

[6] Dolev D., Mokryn O., Shavitt Y. "On Multicast Trees: Structure and Size Estimation". IEEE Infocom 2003.

[7] Tauro L., Palmer C., Siganos G., Faloutsos M. "A Simple Conceptual Model for the Internet Topology". Global Internet, Nov. 2001.

[8] Bollobàs, B. "Random Graphs". Academic Press, New York, 2nd ed.(2001).

[9] Karoński, M. "A Review of Random Graphs". Journal of Graph Theory 6, 349389 (1982).

[10] Erd´ós, P. and Rényi, A. "On the Strength of Connectedness of a Random Graph". Acta Mathematica Scientia Hungary 12, 261267 (1961).

[11] Faloutsos M., Faloutsos P., Faloutsos C. "On Power-Law Relationships of the Internet Topology". ACM SIGCOMM Aug./Sept. 1999.

[12] T. Gri.n, G.T. Wilfong. "An Analysis of BGP Convergence Properties". ACM SIGCOMM 1999. 277 - 288.

[13] C. Labovitz, G.R.Melan, F. Jahanian. "Internet Routing Instability" ACM SIGCOMM 1997, 1997.

[14] R. Albert, H. Jeong, A.-L. Barabasi. "Error and Attack Tolerance of Complex Networks". Nature 406 (2000). 378.

[15] R. Cohen, K. Erez, D. ben Avraham, S. Havlin. "Breakdown of the Internet under Intentional Attack". Physical Review Letters 86 (2001). 3682.

[16] R. Cohen, K. Erez, D. ben Avraham, S. Havlin. "Resilience of the Internet to Random Breakdowns". Physical Review Letters 4626 (2000). 85 - 89.

[17] Dolev D., Jamin S., Mokryn O., Shavitt Y. "Internet Resiliency to Attacks and Failures Under BGP Policy Routing" Computer Networks.

[18] Dandekar, T., Schuster, S., Snel, B., Huynen, M.,Bork, P. "Pathway Alignment: Application to the Comparative Analysis of Glycolytic Enzymes". Biochem. J. 343 (1999). 115 124.

[19] Edwards, J. S.,Palsson, B. O. "The Escherichia coli MG1655 in Silico Metabolic Penotype: Its Definition, Characteristics and Capabilities". Proc. Natl Acad. Sci. USA 97.(2000). 5528 5533.

[20] Cornish-Bowden, A., Cardenas, M. L. "Complex Networks of Interactions Connect Genes to Phenotypes". Trends Biochem. Sci. 26 (2001). 463 465.

[21] Nicholas M. Luscombe, M. Madan Babu, Haiyuan Yu, Michael Snyder, Sarah A. Teichmann, Mark Gerstein. "Genomic Analysis of Regulatory Network Dynamics Reveals Large Topological Changes". Nature, VOL 431, SEPTEMBER 2004. 308 - 312.

[22] H. Jeong, S.P. Mason, A.-L. Barabási, Z.N. Oltvai, "Lethality and Centrality in Protein Networks" Nature 411, (2001).41 - 42.

[23] Cormen T.H., Leiserson C.E., Rivest R.L., Stein C. "Introduction to Algorithms". 2001 MIT press.

[24] Demetrescu C., Thorup M., Chowdhury R. A., Ramachandran V. "Oracles for Distances Avoiding a Node-Link Failure". SODA 2002.

[25] Shavitt Y., Shir E. "DIMES: Let the Internet Measure Itself". ACM SIGCOMM Computer Communication Review, vol. 35, no. 5, Oct. 2005.

[26] Gao L. "On Inferring Automonous System Relationships in the Internet". IEEE Global Internet, Nov. 2000.

[27] Bar S., Gonen M., Wool A. "An Incremental Super-Linear Preferential Internet Topology Model". PAM 2004, Antibes Juan-les- Pins, France, Apr. 2004.

[28] Barabási A., Albert S. "Emergence of Scaling in Random Networks". Science, October 1999. 286. 509 - 512.