Node Status Detection and Information Diffusion in Router Network using Scale-Free Network

A.W.Mahesar, A.A.Shah and M.R.Wahiddin

Department of Computer Science, Institute of Information and Communication Technology, International Islamic University Malaysia
e-mail: Abdul.waheed@live.iium.edu.my;
{messikh|asadullah|mridza}@iium.edu.my

Abstract

In the field of computer networks various routing and internetworking algorithms and protocols have been introduced according to many performance metrics like network topology, scalability, speed, and congestion control requirements. In this paper we have used the concept of scale-free network theory to design a more robust data dissemination approach which can be used in one dynamical Autonomous System (AS) to know the appearance and disappearance of nodes, and speedily propagate the information to all nodes in the routers network. By taking advantage of the features of scale-free network behavior as found in inhomogeneous structure, short path lengths, highly clustered and epidemiological spreading an enhanced algorithm has been introduced which effectively finds the node status in the network and speedily broadcasts the information of status to all nodes in the network.

Keywords

Scale-free networks, degree distribution, preferential attachment, graph theory

1. Introduction

The concept of scale-free nature of many artificial (manmade) or natural complex networks in the world has been extensively studied during the last decade (Strogatz, 2001) (Barabasi, 2001) (Wang, 2002). In the category of technological networks the Internet (Faloutsos et al. 1998), World Wide Web (Reka Albert, Hawoong Jeong, 1999) and electrical power grid (Amaral et al. 2000), are scale-free networks. Similarly, the transportation networks of airways (Barrat et al. 2004), movie-actor collaboration network (Newman et al. 2001), scientific collaboration network (Dame, 2002) and web of human sexual contacts (Liljeros et al. 2001) have been proved to be scale-free networks. The complex structure of these networks has introduced great thrust and interest among the researchers to investigate the internal and hidden organizing principles or rules behind the emergence of these complex networked systems, and their resilience towards breakdown.

Complex networks formation has been seen and observed in many fields of life due to availability of vast amount of data gathered and analyzed with the help of high processing and storage capabilities of modern processing systems. Unfortunately, in spite of the availability of a huge amount of data, there are still many microscopic phenomena needed to fully understand the complexity level and dynamic behavior of these networks. Therefore, there is a severe need to understand the dynamics inside these networks, once understood then it will be easy to fully analyze the behavior with topology of these complex networks.

The most convenient way to represent any network is the graph theory. In graph theory, vertices represent the nodes and edges represent the links in the network. For example, the complex network of internet is a network of domain or routers. The Erdos and Renyi (ER) (A.Renyi and P.Erdos,1959) introduced the concept of random graph theory in the classical mathematical graph theory. According to them, complex networks topology can be best described by random graph. For example, if we have a large number of nodes in the network, and if we connect pair of nodes with the equal probability *p* then ultimate outcome will be physical example of ER random graph. The (ER) model remained very popular from late fifties to late nineties and, the modeling of almost all complex networks was based on random graph theory. According to this model the nodes degree distribution follow the uniform distribution. In this way, the network will be homogeneously connected with long length paths and low clustering coefficient. These features show that the connectivity distribution is homogeneous in networks as shown in Figure 1. In (ER) model the node degree distribution shows very different behavior as compared to real world complex networks. The Barabasi and Albert analyzed the network of World Wide Web (Albert-Laszlo Barabasi and Albert Reka, 1999) and found that the node degree of WWW does not follow random graph connections rather, it is scale-free graph and its degree distributions follow power law form as given in equation (1).

$$p(k) \sim k^{-\gamma} \tag{1}$$

Where p(k) is the probability of node degree distribution and γ is scaling exponent which is a numerical parameter called connectivity distribution exponent. In fact, γ is a scale-free parameter in the sense that it does not depend on a characteristic scale of the network. Further, exponent gamma γ has been measured as well as confirmed in a number of research studies to be approximately 2.1 (Goh et al. 2005). It means, the node degree k and the number of links a node can have, follows the power-law distribution relation. Thus power-law implies that few nodes in the network can have large number of links whereas majority of nodes have very small number of connections. Also, (Faloutsos et al.1998) has shown that, from the autonomous systems perspective internet is also scale-free network.

Therefore, for explaining the power law distribution in complex networks Barabasi and Albert (Albert-Laszlo Barabasi and Albert Reka, 1999) proposed the model that is known as Barabasi-Albert (BA) model in short. According to them there are two main features of scale free behavior of any growing complex network. These features are continuous growth and preferential node attachment. Later they noticed this behavior in many real world networks. Due to these two features, the node with many connections tends to have more chances to acquire links in future like also known as rich get richer phenomenon. Therefore, this is the reason for creation of giant nodes in the network with high node degree distribution. Both these factors influence the creation of inhomogeneous or heterogeneous structure of network topology with hubs and make networks more robust under random node failures.

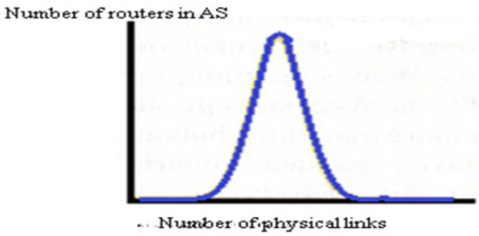


Figure 1: Bell shape curve distribution of routers linkages

Also, the inhomogeneous topological structure allows rapid information diffusion in networks due to short path lengths. Figure 1 shows the nodes degree distribution in ER random graphs, where it follows the Poisson distribution with bell shape curve. Figure 2 shows power law distribution with continuously decreasing curve as fat tailed curve behavior in (a) and decreasing slope on log-log scale of scale-free model in (b).

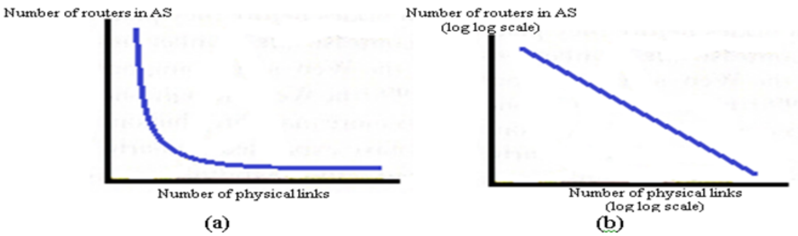


Figure 2: Power law distributions of routers linkages

According to the BA network model, the preferential node attachment plays crucial role in degree distribution of nodes in the network, with assumption that highly connected nodes have more chances to get more and more links. Figure 3 shows the preferential node attachment rule. It can be clearly seen that all the nodes have different importance and this scale-free behavior is the cause of inhomogeneous structure in the growing complex networks.

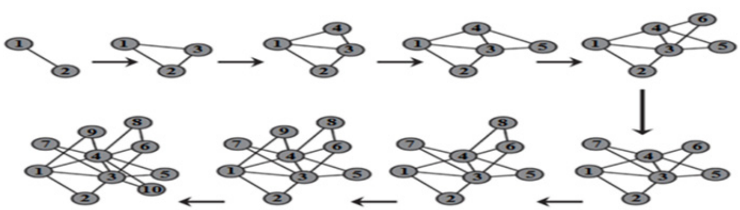


Figure 3: The formation of a scale-free network as an effect of the preferential node attachment (Oliver Hein, 2006)

Therefore, the formation of routers network in autonomous system based on preferential attachment can lead the network towards the scale-free topology. This formation allows the nodes in the network to form inhomogeneous structure and create short path lengths between nodes. In this way by taking advantage of these features and by using them properly, optimum results can be achieved in routers network. The figure 3 shows the network formation as a result of preferential attachment. The rest of the paper is organized as follows. In Section 2 we present scale-free model of routers network. In Section 3 we present the mathematical modeling to know the appearance and disappearance of nodes, ratio based technique and SFN dynamic node status detection algorithm. In Section 4 we discuss simulation results and finally, Section 5 concludes the paper with future work.

2. Scale-free model for routers network

As scale free networks are based on two main features namely growth and preferential attachment, and this behavior has been shown in Internet from the autonomous systems perspective as well as routers perspective as given in (Romualdo Pastor-Satorras 2007). If we analyze this pattern of increasing nodes in the network of routers we can have network of routers as given in figure 4.

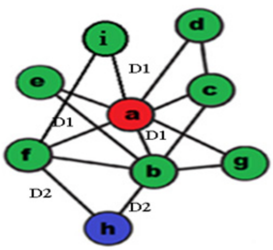


Figure 4: Schematic representation of scale-free network of routers connected in the Internet with D1 (Green) directly attached with a and D2 (h) (Blue) indirectly attached nodes with *a* as main node.

The scale-free network of the routers in the Internet can be represented by graph G (V, E). Where V is the set of vertices or nodes represented by circles, and E is the set of links between nodes represented by lines. Further, D1 are directly connected routers with main router a (Red) and D2 is indirect neighbor or second degree router from a. Also, we can represent the total number of nodes N (t) and the total number of links L (t) as a function of time t. As we are concerned with the dynamic networks therefore the behavior of above network is assumed to be dynamic, as addition or deletion (appearance / disappearance) of nodes makes it highly dynamic, evolving or growing network. Moreover, we can observe the behavior of a network in case of appearance/disappearance of routers in the network by both the dynamical rules governing the nodes and flow occurring along the links in terms of information diffusion. For example, in such network we assume the preferential attachment rule for acquiring the links and information updates of (appearance/disappearance) between nodes as a flow of routing updates.

According to BA model of preferential attachment, there are two main steps as given below: (Barabasi, 2001)

- Growth: Starting with a small number (m_0) of nodes, at every time step, we add a new node with $m \leq (m_0)$ edges that link the new node to m different nodes already present in the system as shown in (Barabasi, 2001).

- Preferential attachment: When choosing the nodes to which the new node connects, we assume that the probability that a new node will be connected to node *i* depends on the degree *Ki* of node *i*, such that

$$\prod i = k_i \div \sum_j k_j \qquad (2)$$

where k_i is the degree of node i and k_j is the sum of the degrees of all nodes in the network.

After *t* time steps, there are

$$N(t) = \lim_{t\to\infty}(t + m_0) \quad \text{nodes} \qquad (3)$$

and

$$L(t) = \lim_{t\to\infty}(mt) \quad \text{edges} \qquad (4)$$

Now, we represent the above network in Figure 4 with the help of the adjacency matrix X= {xij}. This will be the NxN matrix defined such that

$$X_{ij} = \begin{cases} 1 & \text{if (i, j), is directly connected} \\ 0.5 & \text{if (i, j), is connected through second degree node connection} \\ 0 & \text{if (i, j), is not connected} \end{cases}$$

where, i and j are any nodes in the network.

	a	b	c	d	e	f	g	h	i
a	0	1	1	1	1	1	1	0.5	1
b	1	0	1	0.5	1	1	1	1	0.5
c	1	1	0	1	0.5	0.5	0.5	0.5	0.5
d	1	0.5	1	0	0.5	0.5	0.5	0.25	0.5
e	1	1	0.5	0.5	0	0.5	0.5	0.5	0.5
f	1	1	0.5	0.5	0.5	0	0.5	1	1
g	1	1	0.5	0.5	0.5	0.5	0	0.5	0.5
h	0.5	1	0.5	0.25	0.5	1	0.5	0	0.5
i	1	0.5	0.5	0.5	0.5	1	0.5	0.5	0

Figure 5: Mathematical representation of network in Figure 4 as undirected graph, where 0.5 represents the second degree node from main node a.

3. Mathematical modeling of the concept “node detection or disappearance and appearance of nodes in the network”.

To know the appearance and disappearance of nodes in the network, the main node will have to send the periodic update or keep alive messages to all nodes in the network. Let us assume that the sending period to keep alive messages is 10ms (milli seconds (it can be varied depending on the sensitivity or rapid response of nodes in the network). It means after each 10ms update message will be sent to all nodes in the network. In case of second degree or indirect nodes the message will be sent through nearest neighbor node from main node in the network.

Further, if we have two nearest neighbors then less congested or more reliable route can be selected as in many traditional routing protocols. In above network of figure 4, there are two routes to node h (Blue) from node a, first from node a (Red) via node b (green), and second from node f (green). So, in traditional routing the preferred route will be fastest for example, which has more bandwidth or less congestion or it can be any other network metric depending on the type of networks like social network, covert (terrorist) network, airport- air flight network etc. But, here we find the shortest route based on the probability of degree distribution of neighboring nodes, for example from the probability of (a to b) or (a to f) in above scenario of network given in figure 4. If probability from a → b with respect to links of b is high as compared to a → f, then the route from a→ b →h will be selected as shortest route for h. mathematically,

$$\prod_{a\to b} = \frac{Kb}{\sum_j Kj} \tag{5}$$

Where k_b is the degree of node b and k_j is the sum of the degrees of all nodes in the network.

Or

$$\prod_{a\to f} = \frac{Kf}{\sum_j Kj} \tag{6}$$

Where, k_f is the degree of node f and k_j is the sum of the degrees of all nodes in the network.

3.1. The ratio based technique for node detection:

To know the presence and absence of nodes, we find the value of the ratio **r** between sending and receiving messages among nodes from the main node to all. The result of ratio will decide the appearance and disappearance of nodes. Also we set threshold value of ratio ***r*** to further know the stability of nodes in the network. Based on the values of ratio, we can have three cases.

First case:

If the main node sends three messages after each 10ms and if it gets three replies from the receiving nodes in 30ms interval then it will be assumed that the nodes are alive from the ratio of sending and receiving messages. The ratio of messages like three messages and three replies, so ratio is 1, and this ratio will trigger the appearance of node in the network.

Second case:

If main node sends three messages and receives two replies then the ratio will be 3:2 and it will be assumed that the node now may be present in the network. Similarly, if main and receiving node has 3:1 ratios, it will be assumed that the node is not in good condition or it is trying to reduce the chances to get connected or the node is not stable in the network.

Third case:

If main node doesn't receive any reply and the ratio is 3:0, then main node will be able to know that the particular node has disconnected or disappeared and main node will broad cast the message accordingly to all nodes that particular node is no more part of the network now. This can be represented as

Let R represent the ratio, if R= (N received messages / N sent messages) > 0.5 it implies that the node is connected.

If the ratio R= (N receive / N sent) > 0.0 and R < 0.5, then we can assume that node is in unstable condition.

But, if ratio R= (N receive / N sent) is 0.0 it means that node has disappeared from the network.

This effect can be shown mathematically in the adjacency matrix with (0 as disappeared and 1 as present) entries in the rows and columns of the matrix for connectivity. And the value 0.5 represents the presence of indirect nodes in the network. When main node observes from the result of ratio that the particular node is approaching a low threshold, it will be assumed that node is going to reduce the participation in the network in advance.

Further, if we observe from the epidemic modeling perspective, as scale-free nature allow the presence of hubs and giant structure due to power-law distribution, therefore the information in main node can propagate to all nodes like infection propagates. Therefore, by taking advantage of this behavior and disseminating information in an epidemic way in the network can be more effective and threshold value of ratio decides the triggering effect for the main node to decide when to propagate. Also, we can assume that the main router ***a*** in the figure 4 or the router with second largest connections of links ***b*** in figure 4 has advanced configuration like high bandwidth, high processing speed and large buffer size for handling incoming and outgoing information. Moreover, we can say that there are two scenarios, first when the main node try to receive the message from the nodes as pull based. It is something like when node tries to extract new information from its neighbors in above case to know the status (Ganesan et al. 2002). In second scenario when a node sends new information to a selected or all neighbors as push based. Finally, as we have to deal with dynamic or evolving networks, therefore the proposed algorithm for above case must rapidly propagate updates, because some tasks or information have to be activated as soon as possible and newly assigned tasks make the older one useless or obsolete.

3.2. SFN node status detection dynamic routing algorithm.

Algorithm 1: The algorithm for detecting the disappearance and appearance of nodes in scale-free based topology of routers network.

VARIABLE Mn =Main node, Dn =Direct neighbor, In = Indirect neighbor

Step1. Every node in the network will send or exchange their degree distribution to

their neighbor's at each time interval of addition of links.

Step2. Node with higher degree Max: f (Ki) will be selected as the main node (Mn)

of network from degree distribution probability.

Step3. Mn (Main) node will send its IP and node degree to all nodes (direct Dn/

Indirect In) in the network.

Step4. If node is directly connected neighbor THEN

Update will be send directly.

ELSE

In case of indirect neighbor, probability of neighbor's degree distribution

will decide the path.

IF the degree of node b is higher THEN

b will be selected as next route.

ELSE

'f' will be the next route.

Step5. Mn will send periodic message k three times to all connections after each 10 seconds.

Step6. After receiving reply messages from all nodes ratio will be calculated.

IF ratio R=1 THEN

node is connected

ELSE

IF ratio is zero THEN

node will be considered as disappeared

ELSE

IF ratio is in between threshold 0.5< and >0.3 THEN

the node is not in stable condition.

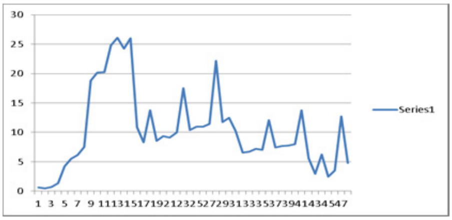
Step7. The unstable node information will be used to select best path in advance in network.

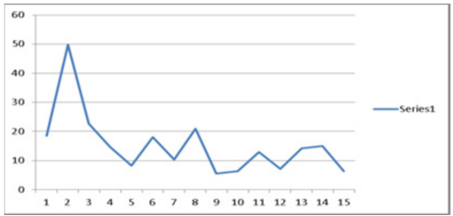
Step8. The disappearance will be disseminated as broad cast message to all nodes in the network epidemically.

---

4. Results and findings

We have implemented and tested this algorithm in Java2SE as research tool. The following graphs show the relation between time in milliseconds and number of nodes in the network as it grows based on preferential node attachment concept. The graph in Figure 7(a) shows the time taken by main node after addition and dissemination of this information to all nodes as network grows. The result shows as network grows based on preferential link attachment and due to inhomogeneous topology caused by few hubs in the network, the time is decreasing to diffuse the information in network as it is growing. In graph 7(a) vertical axis represents time in milliseconds and horizontal axis shows the number of new nodes attached as it grows.





(a) **(b)**

Figure 7: (a) Node appearance and information diffusion time. (b) Node disappearance and information diffusion time

The graph in Figure 7(b) shows node disappearance from network and its information diffusion time is decreasing as network grows based on preferential link attachment. In graph 7(b) vertical axis represents time in milliseconds and horizontal axis shows number of nodes deleted from the network.

5. Conclusion and future work

Based on scale-free features like preferential node attachment and growth, the routers network has been analyzed. By taking advantage of heterogeneous topology formation from preferential node attachment, an efficient way of data transmission to many nodes at once has been achieved. Further, this paper has added the link detection algorithm in this type of topology, which will greatly help to know the status of their detached as well as newly added nodes in the network. Also, the stability of nodes can be determined from the results of ratio and better route can be decided in advance for future configuration in the network. The nodes sudden appearance and disappearance also can be analyzed in many real as well as artificial complex networks and better strategies can be formulated to properly know the behavior and effects in these networks. Also, epidemic spreading and other features in scale free networks can be used in various networks like sensor, ad hoc and wireless networks to better devise powerful and robust algorithms and this may constitute a future work.

6. Acknowledgement:

This research is funded by the Malaysian Grant No. FRGS 11-042-0191

7. References

A.Renyi and P.Erdos, 1959. On Random Graphs. *Publ.Math.Debrecen* 6, p.290-297.

Amaral, L. A. N., A. Scala, M. Barthélémy, and H. E. Stanley, 2000. Classes of behavior of small-world networks. *Proc. Natl. Acad. Sci.* U.S.A. 97, 11 149.

Barabasi, A.L and Albert, R, 1999. Emergence of Scaling in Random Networks. *Science* 286 (5439): 509-12

Barabasi, A.L. and Albert.-R., 2001. Statistical Mechanics of Complex Networks. *Review of Modern Physics*, 77, pp.41–71.

Barrat, A., Barth, M. & Vespignani, A., 2004. Weighted evolving networks: coupling topology and weights dynamics.,(1), pp.1-4.

Dame, N., 2002. Evolution of the social network of scientific collaborations. *Physica A 311*, 3(4), pp.590–614.

Ganesan, D. et al., 2002. An Empirical Study of Epidemic Algorithms in Large Scale Multihop Wireless Networks. *Intel Research, Berkeley Technical Report, IRB-TR-02-003*.

Goh, K., Kahng, B. & Kim, D., 2005. Evolution of the Internet Topology and Traffic Dynamics of Data Packets. , pp.235–250.

Liljeros, F., Edling, C.R. & Stanley, H.E., 2001. Distributions of number of sexual partnerships have power law decaying tails and finite variance. *Nature*, 1(6), pp.1–6.

Michalis Faloutsos, Petros Faloutsos, C.F., 1998. On Power-Law Relationships of the Internet Topology. *ACM SIGCOMM Computer Communication Review*, 29(4), pp.251–262.

Newman, M.E.J., Strogatz, S.H. & Watts, D.J., 2001. Random graphs with arbitrary degree distributions and their applications.

Oliver Hein, M.S. and W.K., 2006. Scale-Free Networks The Impact of Fat Tailed Degree Distribution on Diffusion and Communication Process. , 48, pp.267–275.

Reka Albert, Hawoong Jeong, A.-L.B., 1999. Diameter of the World-Wide Web Growth dynamics of the World-Wide Web Global methylation in eutherian hybrids. *Nature*, 401(September), pp.398–399.

Romualdo Pastor-Satorras, A.V., 2007. *Evolution and Structure of the Internet*, Cambridge University Press.

Strogatz, S.H., Watts, D.J, 1998. Collective dynamics of small-world networks.*Nature* 393 (6684) 409-10.

Wang, X.F., 2002. Complex Networks: Topology, Dynamics and Synchronization. *International Journal of Bifurcation and Chaos*, vol. 12, no. 5, pp. 885–916.