Faculty of Mathematics and Computer Science

Emanuel Bîscă

Social Network Analysis

Evolving Networks: An Investigation

2021

Department of Computer Science

Abstract

The vast majority of real-world systems, which are fundamental for network science, come with different complexities and are a favorable environment for the investigation of the processes that alter the network topology. These processes revolve around the contrast of potential which binds certain nodes to achieve links. My goal was to build an easy-to-understand system of ideas which must offer the theoretical background essential in gaining rewarding insights on how networks evolve in time, and most importantly what forces are acting upon them and trigger their evolution. This mathematical landscape is full of concepts as *initial attractiveness*, *removal of nodes and/or links*, *ageing of nodes* or even the *accelerated growth*. Moreover, extensive explanations are provided to confide safe passage for further individual explorations.

Contents

1 Conceptual Background **1**
- 1.1 Complex Networks 1
- 1.2 The Bianconi-Barabási Model 2
- 1.3 Degree Variations 3

2 Evolving Networks **3**
- 2.1 Initial Attractiveness 4
- 2.2 Internal Links 5
- 2.3 Node Removal 5
- 2.4 Accelerated Growth 6
- 2.5 Aging 6

3 Final Remarks **7**

References **8**

1 Conceptual Background

Various systems are characterized by various complexities, depending on the relationship between their components and the corresponding interactions. Let's consider a biological system, the cell working relies on an elaborated metabolic network, where the enzymes are considered to be the nodes, while the chemical interactions correspond to what's linking them (Jeong et al. 2000). Another nice example concerns the enormous social network that models the society, in which different people or companies represent the nodes, the links being shaped by their social interdependence (Wasserman and Faust 1994), or in the situation in which companies are the nodes, their links are represented by their trading behaviour. Of course, it is possible for us to describe in this manner also the world wide web, where links binding one domain page to another are considered, with html documents as nodes (Barabási and Albert 1999) (Albert, Jeong, et al. 1999).

1.1 Complex Networks

We cannot fully understand the meaning of evolving networks, without taking a look on how complex network are defined in the literature.

If there is to be found one ability that holds for all these complex systems, it would be about their capacity to continuously grow while the time passes. The direct implication of this, it's the fact that the resulting networks are constantly changing by the instrumentality of increasing or decreasing number of nodes and edges; so they're not static networks. The first level to be completed in the path of understanding the effects which are determining the large-scale topology of our systems, is to take a closer look at the active forces acting upon particular nodes.

An important phase to acquire this, was made together with the theory of scale-free models (Barabási and Albert 1999), which states that the evolution of networks is triggered by at least two characteristics that must coexist:

- the *growth*, which implies the constant expanding of networks, increasing the number of new nodes;
- the *preferential attachment*, assuming a greater likelihood of newborn nodes to be linked to nodes having a bigger number of nodes, by that time.

Bearing in mind these two concepts, it's trivial to observe that the rise of *power law connectivity distribution*, extensively presented in the literature, has been predicted by the scale-free model. Moreover, the scale-free model was extended to include instances about the *rewiring* (Albert and Barabási 2000) and the *aging* (Dorogovtsev and Mendes 2000) (Amaral et al. 2000), which were capable to take into consideration more practical sides of how networks are evolving, i.e., the connectivity distribution displays the presence of different cutoffs (shortcuts) as well as that of some so-called scaling exponents.

1.2 The Bianconi-Barabási Model

However, this cannot be the case for all systems, as multiple examples in the literature are proving that the evolution process of systems coming from reality and their connectivity are not dependant only on the age factor. This is a safe affirmation, because some individuals are simply fully equipped to turn some random meetings into strong, durable links, when we talk about the social systems; some companies have an interesting approach that makes their clients to be become devoted partners.

In the situation of world wide web, certain html documents are securing a great network of nodes related to them, surpassing outdated websites, as a result of the commitment to provide to users the best content on the web. If we're thinking to the emerging network of research, we can say that special papers are constructing larger numbers of citations in a short period of time. The feature that all these nodes are sharing is an inherent quality that successfully drives them further in the competition of acquiring more nodes. This ability is called by specialists, a node's *fitness*.

Barabási (2016) proposed in the Barabási-Albert model, the idea that the degree alone determines the growing process of a node. So that the purpose of the *fitness* will be incorporated, let's consider the preferential attachment to be triggered by the product between two important values, the fitness of that node, denoted by ξ and k, the corresponding degree. The result, is a model known in the literature as the *fitness model* or the Bianconi-Barabási model (Bianconi and Barabási 2001). To obtain it, the next two phases are assessed:

- *evolution*

Any new time state adds, to the current network, a newborn node y, with h denoting the number of links and ξ_y its fitness. Two important remarks need addressing here: ξ_y is randomly picked from a fitness distribution $\rho(\xi)$ and it's allocation happens only once for each node.

- *preferential attachment*

It declares the proportionality of the likelihood which assigns a link connecting a newborn node with the node x and the previously mentioned product between the degree of the node x, which is k_x and the matching fitness, $\xi(x)$.

$$\Pi_x = \frac{\xi_x k_x}{\sum\limits_y \xi_y k_y} \tag{1}$$

Now, starting from that we must be able to conclude some ideas regarding certain dependence relations. Firstly, the dependence existent between Π_x and k_x, makes the nodes with greater degrees to be easily observable, thus they are more appropriate to gain many connections. Also, because Π_x depends on ξ_x it's safe to assume that in the case of nodes characterized by equal degrees, the one having the better *fitness* will be a more likely option. Therefore, Equation 1 offers the assurance that even a recent node, which is connected only to few of the other nodes in its initial state, is capable of growing a stronger network around him, in a quite short period of time, if it possesses a more generous fitness, than most of the remaining nodes.

1.3 Degree Variations

The *continuum theory* (Bianconi and Barabási 2001) can be utilized to emit predictions related to the evolution through time, of each node. The next result is a consequence of Equation 1:

$$\frac{\partial k_x}{\partial t} = h \frac{\xi_x k_x}{\sum_y \xi_y k_y} \tag{2}$$

Next, we consider that a power law distribution characterizes the evolution in time of k_x, where $\beta(\xi_x)$ is a term that depends on fitness.

$$k(t, t_x, \xi_x) = h \left(\frac{t}{t_x} \right)^{\beta(\xi_x)} \tag{3}$$

Combining these last two equations, we come to an interesting result, which reveals that the exponent β meets the requirements of:

$$\beta(\xi) = \frac{\xi}{G} \tag{4}$$

$$G = \int \rho(\xi) \frac{\xi}{1 - \beta(\xi)} d\xi \tag{5}$$

When we talk about the Barabási-Albert model, the corresponding degree for any node is getting increased as a square root of time since its existence, as the exponent $\beta = 0.5$. But in the case of Bianconi-Barabási model, considering the facts of Equation 4, we are aware that β varies proportionally with the fitness of each node, ξ, concluding that any given node would hold a particular dynamic exponent β. Therefore, the degree growth of nodes with bigger values of fitness would be accelerated.

2 Evolving Networks

The purpose why the (B-A) Barabási-Albert model (Barabási and Albert 1999) appeared in the literature, was to meet the needs of a model capable to expose the processes in charge for the materialization of the scale-free property, in various networks. Since then, studies shows the existence of some limitations:

- it estimates the degree exponent γ to be equal to 3, but the extensive experiments being completed through time, show that γ has fluctuating values in $[2, 5]$;
- it's predicting a power-law degree distribution, although real-world networks are displaying orderly deflections from genuine power-law functions;
- it disregards a multitude of vital mechanisms which exist in most of the real-world systems, such as the deletion of links or further addition of links between existing nodes.

All these facts have determined scientists to consider further research, to enlighten even more the mechanisms which stay behind the concept of *network topology.* I reserved this section to grasp more meaningful information on how an extension of the initial B-A model, can offer us details about the substantial spectrum of phenomena that influence real-world network topology.

2.1 Initial Attractiveness

Even though it's practically impossible for an unconnected node to achieve even a small number of links in the B-A model, due to the preferential attachment which dictates that if a node has $k = 0$, then the probability of a newly-born node to become linked to it is equal to zero, in networks coming from reality such nodes characterized by solitude can become linked to new nodes at a certain time. In order to make this linking possible, the preferential attachment equation is altered to look like that:

$$\Pi_k \sim A + k. \tag{6}$$

So, a new constant, A, was added, known in the literature as initial attractiveness. In the particular situation, when $\Pi(0) \sim A$ is true, then the likelihood of acquiring the first link in the immediate time stage, corresponding to a specific node is pro rata with A.

Two main effects of the presence of initial attractiveness in real networks, have been spotted:

1. The degree exponent is increased and it satisfy the equality: $\gamma = 3 + \dfrac{A}{h}$;

The meaning being that A boosts the homogeneity of the network, thus it consistently reduces the hubs dimensions.

2. The corresponding degree distribution of such a network doesn't satisfy the requirements of a genuine power-law one, instead it gains the shape of: $p_k = G(k + A)^{-\gamma}$;

We can conclude saying that A intensifies the likelihood of nodes with smaller degrees to become linked to a newly-born node, which determines the degree of such nodes to grow bigger, faster.

2.2 Internal Links

In the vast majority of networks, linking doesn't happen just between newly-appeared nodes and already existing ones, as old nodes can develop links among themselves. The measures performed on the collaboration networks prove that a double preferential attachment holds when it comes to internal links; basically we can formalize the internal connection likelihood between any two nodes having the degrees k and k', via:

$$\Pi(k, k') \sim (A + Tk)(A + Tk)' \tag{7}$$

The influence of internal links on the topology of real networks, is better understood after we take a look at the two frontier cases of Equation 7:

1. When A is equal to zero, we're in the double preferential attachment case;

Let's assume that in the B-A model, any time step comes with a newly added node characterized by h links, then s internal links, with respect to the particular likelihood of this individual case. In this situation, we are aware that the degree exponent satisfies the equality: $\gamma = 2 + \frac{h}{h+2s}$, which indicates that the degree exponent is somewhere in $[2, 3]$. See Albert and Barabási (2000) Ghoshal et al. (2013). The correct conclusion states that the degree exponent is lowered by this approach, therefore the network is becoming more heterogeneous.

2. When T is equal to zero, the situation of random attachment emerges;

This situation assumes that the nodes degree do not play any role when establishing a connection with an internal link between the considered nodes. In other words, the addition of internal links, happens between random couples of nodes. We assume once more that we start from the B-A model, in which when a newly added node comes in the network, also a number of s links are introduced between random nodes. It's proven that the next equation displays the degree exponent of such a network: $\gamma = 3 + \frac{2s}{h}$ (Ghoshal et al. 2013). Therefore, it's obvious now that any combination of h and s will provide us $\gamma \geq 3$, suggesting that a network constructed using this philosophy would be characterized by homogeneity, more than one that neglects the internal links.

2.3 Node Removal

A high number of systems from reality are displaying a behaviour that leads to the disappearing of nodes and links. Indeed, if we consider a social network modeling the relationships from a corporation, there is the possibility that some people would trade their job for a better one, in another work environment.

The B-A model is the ground for the case study around deletion of nodes (and links as a consequence). Each stage through time is considered to be defined by two parameters; a newly added node connects to h other nodes, while other node is removed with respect to a rate r. The studies (Saavedra et al. 2008) (Bauke et al. 2011) around r have come to the conclusion that three distinct scaling directions happen:

1. The scale-free stage

This stage guarantees that the network is continuously growing, as more nodes are added to the network in each time step, than are deleted; so the removal rate is $r < 1$ At this time, the degree exponent takes the form dictated by the equality: $\gamma = 3 + \dfrac{2}{1-r}$.

2. The exponential stage

In the special case of $r = 1$, the number of nodes which exist in the network in any time step is the same, because the deleted nodes are equal to the nodes added. The scale-free property no longer holds.

3. The declining networks

As the name suggest, we can already know that in this phase, the deletion rate is greater than 1, concluding that the newly-added nodes are no longer exceeding the number of deleted nodes, not even equals it, creating a declining trend in the network. A great example which gives consistent insights on this matter, are the extensive studies performed in ecology's exploration of the sequential loss of natural territories (Srinivasan et al. 2007).

2.4 Accelerated Growth

The most defining characteristic of a network experiencing accelerated growth, is that the rate in which newly added nodes come to existence is lower than the speed of creating new links. This phenomena is nothing unusual for many real-world systems.

The investigation around this accelerated growth, must be done by assuming that the next relation defines how the multitude of links arrive with any node:

$$h(t) = h_0 t^{\lambda} \tag{8}$$

The $\lambda = 0$ guarantees that links are equally distributed for any newly added node. On the other hand, if the equality $\lambda > 0$ is true, then an accelerated growth influences that network. Keeping in mind the result of Equation 8, we can safely affirm that accelerated growth in the B-A model would make the degree exponent to follow the relation:
$\gamma = 3 + \dfrac{2\lambda}{1-\lambda}$.
So, the homogeneity of such a network is triggered by the accelerated growth, as it drives the degree exponent γ to accept values above 3.

2.5 Aging

In reality, the dynamic systems have particular properties that impose enforce the nodes to become unable to acquire new links, to connect to other nodes. It's easy to understand this concept, if we think of the actors which have time-limited careers; so, in a point in time they stop acting and their process of growing their network interrupts.

To gain more insights on the influence of aging, let's consider that a new node become connected to node x following the likelihood of: $\Pi(k_x, t - t_x)$, in which t_x defines the time-step when the node x appeared. A formal way of modeling the aging phenomena:

$$\Pi(k_x, t - t_x) \sim k(t - t_x)^{-\psi} \tag{9}$$

where ψ is an adjustable feature that permits to establish a relation of dependence between the likelihood of preferential attachment and how long have the node been in the network. Again, three scaling regimes have been discovered to exist; each depends on the eventual values of ψ:

1. Negative ψ;

In this situation, older nodes act as high interest of connection for new nodes. Somehow, we can say that $\psi < 0$ gives a lot of importance to the preferential attachment; talking extremely, if $\psi \to -\infty$, all the new nodes will desire to link the most aged node. The scale-free nature of such a network is conserved, and even though the γ is under 3, the network display heterogeneity.

2. Positive ψ;

As it can be foreseen already, this situation where $\psi > 0$ determines the most recent nodes to be of interest for the newly-added nodes; talking extremely, if $\psi \to \infty$, any node will be forced to link to its most appropriate precursor. There is no demand for high values of ψ to observe the influence on the aging phenomena; the γ diverges while ψ approaches 1. Therefore, successive ageing makes the network more homogeneous, by diminishing the power of the bigger hubs with old nodes.

3. The case of $\psi > 1$;

This special case, leads to damages on the scale-free nature because the contribution of the preferential attachment was conquered by the ageing phenomena.

3 Final Remarks

In the end, I'd like to say a few words about the results I've been discussing in my research report. My hope is that I successfully explained each process and the corresponding impact on the dynamism of evolving networks. My main goal was to offer valuable insights on the authentic capabilities of growing networks model, building the conceptual framework which give us the mathematical means necessary to study the forces that act upon the evolution of networks, remodeling their topology.

References

Albert, R. and Barabási, A.-L. (2000). “Topology of Evolving Networks: Local Events and Universality”. In: *Physical Review Letters* 85.24, pp. 5234–5237.

Albert, R., Jeong, H., and Barabási, A.-L. (1999). “Diameter of the World-Wide Web”. In: *Nature* 401, pp. 130–131.

Amaral, L. A. N., Scala, A., Barthelemy, M., and Stanley, H. E. (2000). “Classes of small-world networks”. In: *Proceedings of the National Academy of Sciences* 97.21, pp. 11149–11152.

Barabási, A.-L. (2016). *Network Science*. first edition. Cambridge, UK: Cambridge University Press.

Barabási, A.-L. and Albert, R. (1999). “Emergence of Scaling in Random Networks”. In: *Science* 286.5439, pp. 509–512.

Bauke, H, Moore, C., J.B., Rouquier, and Sherrington, D. (2011). “Topological phase transition in a network model with preferential attachment and node removal”. In: *European Physical Journal B* 83.4, pp. 519–524.

Bianconi, G. and Barabási, A.-L. (2001). “Competition and multiscaling in evolving networks”. In: *Europhysics Letters* 54.4, pp. 436–442.

Dorogovtsev, S. N. and Mendes, J. F. F. (2000). “Evolution of networks with aging of sites”. In: *Physical Review Letters* 62.2, pp. 1842–1845.

Ghoshal, G., Chi, L., and Barabási, A.L. (2013). “Uncovering the role of elementary processes in network evolution”. In: *Scientific Reports* 3.1, pp. 1–8.

Jeong, H., Tombor, B., Albert, R., Oltvai, Z. N., and Barabási, A. -L. (2000). “The large-scale organization of metabolic networks”. In: *Nature* 407, pp. 651–654.

Saavedra, S., Reed-Tsochas, F., and Uzzi, B. (2008). “Asymmetric disassembly and robustness in declining networks”. In: *Proceedings of the National Academy of Sciences of the United States of America* 105.43, pp. 16466–16471.

Srinivasan, U.T., Dunne, J.A., Harte, J., and Martinez, N.D. (2007). “Response of complex food webs to realistic extinction sequences”. In: *Ecology* 88.3, pp. 671–682.

Wasserman, S. and Faust, K. (1994). *The large-scale organization of metabolic networks.* Cambridge, UK: Cambridge University Press.