

# An Implementation of Preferential Deletion in Dynamic Models of Web-Like Networks

Carlos Santiago Bañón

College of Engineering and Computer Science - University of Central Florida - Orlando, FL

---

## Abstract

This paper attempts to replicate the preferential deletion model for generating random graphs defined by Narsingh Deo and Aurel Cami. In this study, the data brought forth by this new implementation agrees seamlessly with the original model. The data presented by Deo and Cami is validated and corroborated, as similar results are found. Further, this paper serves to once again show that the expected fraction of nodes with degree  $k$  in the graph generated by this process does indeed decrease asymptotically as  $k^{-1-(\frac{2p}{2p-1})}$ .

**Keywords:** Web-Like Networks, Interconnection Networks, Dynamic Random Graph Modeling, Preferential Node Deletion, Scale-Free, Power-Law Degree Distribution

---

## 1. Introduction

The purpose of this paper is to replicate the results and conclusions obtained by Narsingh Deo and Aurel Cami in their 2005 paper *Preferential Deletion in Dynamic Models of Web-Like Networks*. In their paper, they carefully study the properties of “web-like” networks, which they define as “real-life networks such as the network of phone calls, power-distribution networks, citation network, science-collaboration network, movie-actor collaboration network, the network of sexual contacts, neural networks, and various infrastructure networks.”<sup>[1]</sup>

They begin their paper by introducing this concept, following with a discussion on how to properly simulate such networks. Quickly stating that the Erdős-Rényi model does not provide an accurate picture for these networks, they propose a new model: the *Preferential Deletion Model*.

This model focuses on mirroring the properties of a dynamic random graph model, which they formally define as “a discrete-time process which starts out with a small fixed graph and in each subsequent time step a new node is added to the graph or an existing node is deleted from the graph.”<sup>[1]</sup> For their model, this process of adding and deleting nodes is controlled by some stochastic rules, which are described in the formal definition of the model below.

## 2. Notation

As this paper is an implementation of the paper by Deo and Cami, the same notation is used, specified in Figure 1.

## 3. Definition of the Preferential Deletion Model

Let  $G_1$  be a graph that begins with a single node and a self-loop. For each discrete time step  $t + 1$ ,  $t > 0$ , either one of these processes can occur:

### 3.1. Birth Process

Deo and Cami define the birth process as follows: “with probability  $p$ , a new node is added, together with a new edge incident on it. The other end-node  $u$  of the new edge is chosen preferentially from among all the existing nodes based on a *linear preferential attachment rule*.”<sup>[1]</sup>

Further, this linear preferential attachment can be defined as:

$$\mathbb{P}_{t+1}[u] = \frac{d_t(u)}{\sum_{w \in V_t} d_t(w)} = \frac{d_t(u)}{2m_t}. \quad (1)$$

Symbol	Meaning
$G(V, E)$	A graph with set of nodes $V$ and set of edges $E$
$n(G)$ or $n$	Number of nodes in $G$
$m(G)$ or $m$	Number of edges in $G$
$d(u)$	Degree of node $u$
$\bar{d}(G)$	Average degree of graph $G$
$G_t$	Graph obtained after the $t$ th step of a random graph generating procedure
$n_t$	Number of nodes in $G_t$
$m_t$	Number of edges in $G_t$
$d_t(u)$	Degree of node $u$ in $G_t$
$\bar{d}_t$	Average degree of $G_t$
$N_{k,t}$	Number of nodes of degree $k$ in $G_t$
$N_{k,t}^{(1)}$	Number of neighbors of degree $k$ of a node chosen for deletion from $G_t$
$N_{k,t}^{(1)}(u)$	Number of neighbors of degree $k$ of node $u$ in $G_t$
$\mathbb{P}[k]$	Fraction of nodes of degree $k$
$\mathbb{P}_t[u]$	Probability of vertex $u$ being selected as an end-node of a new edge or for deletion during the $t$ th step of the random graph process
$\mathbb{P}[A]$	Probability of event $A$
$\mathbb{E}[X]$	Expectation of random variable $X$

Figure 1: Notation used in this paper.

### 3.2. Death Process

If a new node is not created, then a node will be deleted from the list. This process is defined as follows: “with probability  $q = 1 - p$ , a node  $u$  is chosen for deletion along with all the edges incident on it in  $G_t$ . To make small-degree nodes more likely candidates for deletion than the higher-degree ones, node  $u$  is chosen according to the probability distribution defined below.”<sup>[1]</sup>

$$\mathbb{P}_{t+1}[u] = \frac{n_t - d_t(u)}{n_t^2 - 2m_t} \quad (2)$$

Finally, they state that, if during any step  $t < 0$ , the graph becomes empty, then it starts anew with a single node and a self-loop.

### 4. Number of Nodes

The original paper begins its discussion regarding the expected number of nodes in the resulting random graph by stating that, based on their model, the graph will seldom vanish. That is, there is a very low probability, shown mathematically as  $\mathbb{P}[n_t = 0]$ , that the graph will disappear after some step  $t > 0$ . Further, after some derivation, they formally established the expected number of nodes after running their model to be:

$$\mathbb{E}[n_t] = (p - q)t + 2q. \quad (3)$$

This strongly suggests that the order of growth of  $n_t$  is:

$$\mathbb{E}[n_t] = \Theta[(p - q)t]. \quad (4)$$

With this in mind, the new implementation of their model was run with similar parameters in order to test this definition. With the average data collected from the simulations, very similar results were obtained, thereby confirming this mathematical expectation. The expected number of nodes obtained from this new implementation of the model can be seen in Figure 2.

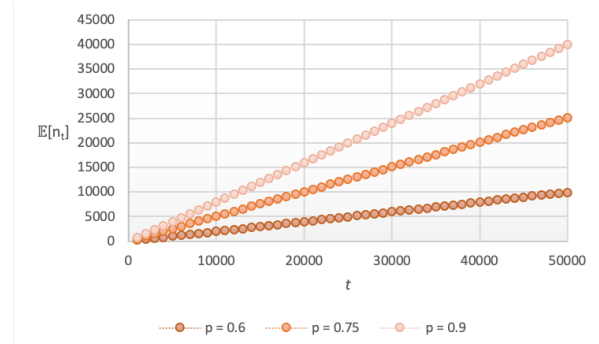


Figure 2: Expected number of nodes.

### 5. Number of Edges

Deo and Cami then continue by discussing the expected number of edges in a random graph created by their model. After deriving from various probabilities, they arrive at the conclusion that the expected number of edges of graph  $G_t$  can be modeled by the order of:

$$\mathbb{E}[m_t] = \Theta[p(p - q)t]. \quad (5)$$

As before, the new implementation of the model was also used to corroborate these findings. The new results, seen in Figure 3, clearly show an agreement between the original paper and this new implementation. The expected number of edges of graph  $G_t$  can be seen in Figure 3.

### 6. Degree Distribution in the First Neighborhood of the Deleted Node

Further, the discussion then shifts to an in-depth analysis of the degree distribution of a random graph  $G$  generated with their model. However, Deo and Cami first focus on the degree distribution in the first neighborhood of a deleted node in  $G_t$ . They formally portray

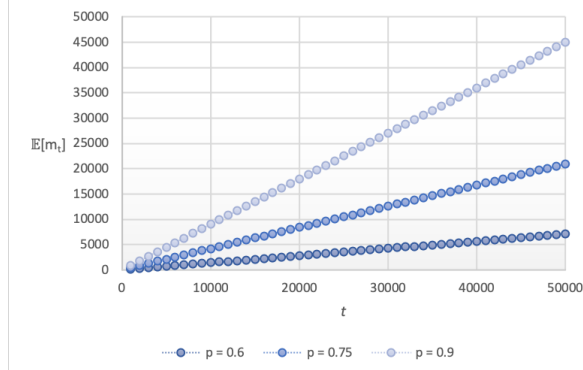


Figure 3: Expected number of edges.

this idea as the expected “number of neighbors of degree  $k$  of a node chosen for deletion from  $G_t$ .”<sup>[1]</sup> This can be formally defined as:

$$\mathbb{E}[N_{k,t}^{(1)}] \approx k \mathbb{E}[N_{k,t}] \mathbb{E}\left[\frac{1 - 2\bar{d}_t/n_t}{n_t - \bar{d}_t}\right]. \quad (6)$$

This distribution was also gathered from the new implementation of the model, and this new data can be seen in Figure 4. As evident, the new implementation agrees seamlessly with the original paper.

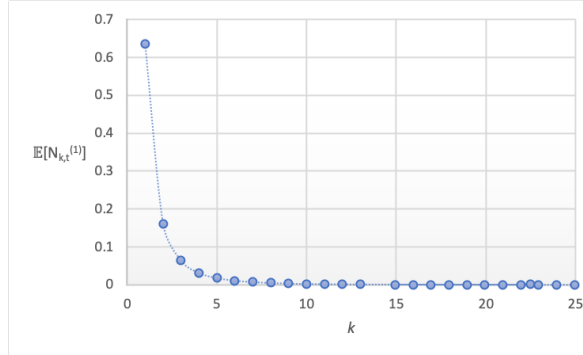


Figure 4: The expected number of neighbors of degree  $k$  of a node chosen for deletion.

## 7. Degree Distribution

Finally, Deo and Cami turn to the overall degree distribution for a graph  $G_t$  generated randomly by their model. Mathematically, they found that “the expected fraction of nodes with degree  $k$  in the graph generated by this process decreases asymptotically as  $k^{-1-(\frac{2p}{2p-1})}$ ,” and this statistic was also gathered by the new implementation described in this paper.<sup>[1]</sup> The cumulative degree distribution obtained by the new implementation

can be seen in Figure 5. These values match the experimental values produced for the original paper, and thus agree with the mathematical model.

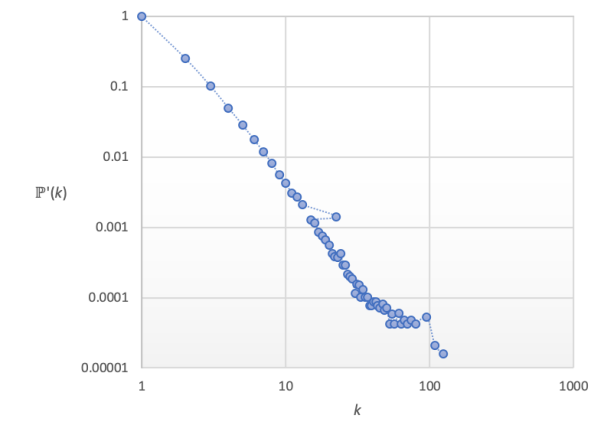


Figure 5: Cumulative degree distribution of the graph generated by the new implementation of the preferential deletion model.

## 8. Conclusion

With this paper, we can conclude that indeed the results of the original paper can be replicated, and thus gives further validity to the preferential deletion model presented by Deo and Cami. That is, this model generates graphs with asymptotically power-law degree distribution. This then further establishes the model as a very sound way of dynamically generating random graphs for further study of web-like networks. However, this model, as any model, can certainly be improved, and will be the subject of a study in the very near future.

## 9. References

- [1] N. Deo, A. Cami, Preferential deletion in dynamic models of web-like networks, in: Information Processing Letters, Volume 102, Issue 4, 2007, Pages 156-162.