Barabási-Albert algorithm (BA)

The Barabási-Albert algorithm of preferential attachment is a network growth procedure which was introduced to show how a simple rule can generate degree distributions similar to the ones observed in many real networks.

The algorithm starts with a small number of nodes (e.g. 5) connected between them basically as you want, usually forming a clique, i.e. all-to-all connections. Then, new nodes are added to the network in such a way that each new node forms a fixed number *m* of edges with the previously built network. The parameter *m* defines the average degree of the resulting network: in the limit of a large number of nodes, we will have $<k> = 2m$.

The only difficult part of the implementation is the selection of the nodes to which the new edges are attached. The preferential attachment rule established that the probability of selecting a node must be proportional to the degree of the destination node. For example, if the current network has 6 nodes of degrees 2, 3, 2, 4, 2 and 3 respectively, the probabilities to connect with each one are 2/16, 3/16, 2/16, 4/16, 2/16 y 3/16 respectively. How do we generate random numbers following these probabilities? We just need to generate a uniformly distributed float random between 0.0 and 16.0, and choose the proper bin:

- Between 0.0 and 2.0 (width = 2) we select the first node.
- Between 2.0 and 5.0 (width = 3) we select the second node.
- Between 5.0 and 7.0 (width = 2) we select the third node.
- Between 7.0 and 11.0 (width = 4) we select the fourth node.
- Between 11.0 and 13.0 (width = 2) we select the fifth node.
- Between 13.0 and 16.0 (width = 3) we select the sixth node.

Let us suppose we have a function *random*() which generates float random numbers between 0.0 and 1.0, and that when we have called it has returned a value 0.6789. Then we multiply it by 16, to obtain 0.6789 * 16 = 10.8624. Since this value is within the fourth bin, the first edge of the new node (the seventh one) would connect it to the fourth node. If $m > 1$, we would repeat the same procedure until the *m* edges have been formed. If the process happens to choose one of the previously selected nodes, we just generate new random numbers until a different one is selected.

Once the *m* edges have been established, bins will have to be recalculated (some of the degrees have changed and a new node has been added) to add the next node, and so on until the desired size of the network has been achieved.