Your start-up has to promote its new Twitter feature. To this aim, they are considering to freely distribute the feature to a subset of at most 100 nodes in the networks, henceforth called *the seed*, and ask them to use it. Your hope is that they will influence their neighbors and, in this way, you can make the feature known to most of the network. You are in charge of deciding how many and which nodes will be part of the seed. You have to run a set of experiments in order to produce a seed of minimum size that maximizes the expected number of adopters of the new feature. The selection of the nodes should be made with respect to a centrality measure. In particular, you have to run the experiment with respect to three different measures: betweenness centrality, eigenvector centrality and **Lin's index**[1]. The spread of the influence in the network has to be modeled using the **Linear Threshold Model**[2].

Each experiment must be run on the following graphs:

- The Twitter graph (avialable on line),
- 100 random **direct** graphs,
- 100 Watts-Strogatz **direct** graphs on the 2-dimensional space,
- 100 **Preferential Attachment graphs**[3].

Each generated graph must have n nodes and m edges and average clustering c, where:

- n is the number of nodes of the Twitter graph;
- m is between 1600000 e 2000000;
- c is between 0.3 and 0.8.

[1]The Lin's index of a node u is $L(u) = \frac{|\{v:d(v,u)<\infty\}|^2}{\sum_{v:d(v,u)<\infty} d(v,u)}$.

[2]In the Linear Threshold model, to each node u is assigned a threshold θ_u chosen uniformly at random in $[0, 1]$. The Linear Threshold model consists in rounds. At beginning of each round we are given a partition $(A, \overline{A})$ of nodes, where A represents the nodes adopting the promoted feature. For the first round A coincides with the seed. The round proceeds as follows: for each node $v \in \overline{A}$ if the fraction of neighbors in A is above the threshold θ_v, then v is said to be successfully influenced. Let $N \subseteq \overline{A}$ be the set of nodes that are successfully influenced. Then, the next round will be run with partition $(A \cup N, \overline{A} \setminus N)$. The process continues as long as $N \neq \emptyset$.

[3]Preferential Attachment graphs needs three parameters: the number n of nodes, the maximum out-degree d and a probability p. The graph is then generated as follows: nodes are created in order, and named 1, 2, ldots, N. When node j is created, we produce its out-links by repeating at most d times the following procedure:

- with probability p, choose a node i uniformly at random, and creates an edge from j to i;
- with probability $1 - p$, choose a node ℓ with a probability that is proportional to the ℓ's current in-degree, and creates an edge from j to ℓ.