Open Mathematics

Research Article

M.T. Trobajo*, J. Cifuentes-Rodríguez, and M.V. Carriegos

On dynamic network security: A random decentering algorithm on graphs

https://doi.org/10.1515/math-2018-0059
Received September 29, 2017; accepted April 23, 2018.

Abstract: Random Decentering Algorithm (RDA) on a undirected unweighted graph is defined and tested over several concrete scale-free networks. RDA introduces ancillary nodes to the given network following basic principles of minimal cost, density preservation, centrality reduction and randomness. First simulations over scale-free networks show that RDA gives a significant decreasing of both betweenness centrality and closeness centrality and hence topological protection of network is improved. On the other hand, the procedure is performed without significant change of the density of connections of the given network. Thus ancillae are not distinguible from real nodes (in a straightforward way) and hence network is obfuscated to potential adversaries by our manipulation.

Keywords: Network analysis, Cybersecurity, Scale-free graph, Centrality, Algorithms

MSC: 68R10, 94C15

1 Introduction

Network analysis [1] is a research field related to a wide range of scientific areas having roots in seminal papers on social sciences [2–5] but also of increasing interest in Biology [6, 7], Engineering [8, 9], Finance [10] or Computer Science and Security [11–13]. Networks are structures collecting a fixed number of objects or entities called nodes together with relations between them called links. Hence algebraic structures like graphs, directed graphs and their weighted versions [14] are fundamental tools to describe networks in order to perform simulations and qualitative or quantitative analysis by means of statistical or numerical procedures.

In addition to social or physical networks, like social work environments [15], transportation networks [8, 9], supply systems [16, 17], we have technological networks, like the Internet [18] or data networks, whose nodes are computers or virtual machines or email clients [11] and links represent the data exchange between nodes.

Distributed computing over telecommunication networks, peer-to-peer file sharing, online communities, social networks or even virtual networks in the cloud are emerging topics of increasing interest in several branches of applied mathematics which are essentially graph structures. This applies also to privacy concerns related to data-analysis over graphs and networks including attack or fraud auditing techniques in cybersecurity research [11].

***Corresponding Author: M.T. Trobajo:** Departamento de Matemáticas, Universidad de León, 24071 León, Spain, E-mail: m.trobajo@unileon.es

J. Cifuentes-Rodrıguez: Departamento de Tecnología Minera, Topográfica y de Estructuras, Universidad de León, 24071 León, Spain, E-mail: jcifr@unileon.es

M.V. Carriegos: Departamento de Matemáticas, Universidad de León, 24071 León, Spain, E-mail: miguel.carriegos@unileon.es

Privacy breaches in a network can be grouped into three categories [19]: 1) *identity disclosure*: the identity of an individual who is associated with a node is revealed; 2) *link disclosure*: the sensitive relationships between two individuals are disclosed; and 3) *content disclosure*: the sensitive data associated with each node is compromised, e.g. the email message sent and/or received by the individuals in a email communication network. A privacy-preserving system over graphs and networks should consider all of these issues.

Those issues imply the following challenges [19]:

Challenge 1. *To model the knowledge and the capability of an adversary/attacker because any topological structures of the graph can be exploited by the attacker to derive private information.*

Challenge 2. *To quantify the information as a function of several different measures associated to the graph: degree, centrality, betweenness, average path length, diameter, clustering coefficient etc. Should we attempt to modify these metrics? How?*

Challenge 3. *To define graph-modification algorithms that balance privacy and data utility. The nodes and links of a graph are all correlated. Thus, the impact of a single change of an edge or a node can spread across the whole network.*

Challenge 4. *To model the behavior of the participants involved in a network-based collaborative computing environment.*

Several privacy models, adversaries and graph-modification algorithms have been proposed recently [20–22] in order to face the challenges above. Note also that network recovery can be approached by similar methods (see [23–25]). Unfortunately, it is unlikely to solve all problems in a single shot as protection against each type of privacy branch requires different techniques or even a combination of them [19].

In this work we focus on topological properties related to centrality in networks. Structural central features of technological and communication networks are associated with their capacity to share and broadcast information in a secure way.

To be specific, when the flow of the information is condensed into a few amount of nodes, it is easier to crash the network by removing or infecting some of those nodes before the system administrator detects and stops the attack. Thus, networks with a few of such *central* nodes, called hubs, are more vulnerable to targeted attacks [26, 27]. Our goal is to obtain suitable dynamic extensions to hide the central structure of the network, so that it becomes less vulnerable to attacks.

Therefore, it's necessary to know how principal or central a node is, and moreover, if a network has a high level of centralization or not. Concepts of centrality and centralization measures are discussed in the following sections. Note that there have been studies on hiding nodes in a network [28] but the focus therein is not on centrality but on percolation threshold.

Section 2 is devoted to some topics on centrality and centralization measures. Then section 3 describes the effect on the underlying graph of introducing ancillary nodes to a given network; afterwards, in section 4 we propose our Random Decentralized Algorithm (RDA) to perform those dynamic extensions on networks. In 5 we provide several numerical simulations, test RDA and analyze numerical and graphical results in order to establish some final conclusions and comments.

2 Graph centrality measures

From a computer science perspective, a network can be identified with a graph $G = (V, E)$, where $V = \{v_1, v_2, \dots, v_n\}$ is the set of vertices or nodes and $E \subset V \times V$ is the set of links or edges; a pair $(v_i, v_j) \in E$ if and only if vertices v_i and v_j are connected. In this work we are interested in network structures with data exchanges, where nodes are terminals or other connected computational devices. Therefore, two nodes are linked if and only if they are connected and there is information exchange between them. Several approaches to technological communication networks could be modeled by graphs, but it is not within the scope of this paper to cover all cases, so we are considering binary relations uniquely, which implies that underlying graphs are undirected and unweighted. Also the graph will be assumed to have no isolated nodes nor loops. We will call such networks simple networks. Finally, the networks will be assumed to be connected. We shall refer to Brandes and Erlebach [1] for terminology and basic properties on Network Analysis.

An undirected graph $G = (V, E)$ consists of two sets V, and E, such that $V \neq \varnothing$ and E is a set of unordered pairs of elements of V. A graph G is completely determined by its adjacency matrix $A(G) = A = (a_{ij})$, where $a_{ij} = 1 \Leftrightarrow (v_i, v_j) \in E$ and $a_{ij} = 0 \Leftrightarrow (v_i, v_j) \notin E$. Symmetric 0-1 matrix $A(G)$ contains relevant information about the topology of the network, such as centralization, number of triangles, diameter of the network, presence of cohesive clusters, bipartite character, randomness, etc. ([1, 5, 29]).

We briefly discuss in this section some measures of centrality of a single node within a graph -point centrality measures- and Freeman's normalization procedure to measure the centrality of the graph taken as a whole -network centralization measures-. Much work has been done on centrality properties and their applications [30–32]. In this paper we follow Freeman [33] who in the seventies collected and formalized point centrality and network centrality concepts.

Definition 1. *Degree point centrality computes the total amount of adjacencies (direct neighbors) of a node v_k, which corresponds to the sum of terms of row (or column) k of the adjacency matrix A.*

$$C_D(v_k) = \sum_{i=1}^{n} a_{k,i}$$

Therefore, a node with high degree centrality is an important node of communication, due to its capacity to have direct contact with a great number of nodes in the network. Nevertheless, a node could have high degree, but be disconnected to other nodes if their neighbors have no links to others in the network. So, it is central in a “local sense". To solve this limitation, we consider betweenness and closeness point centrality measures, which take into account the connections of all the nodes in the network.

Betweenness point centrality quantifies the ratio of geodesics (shortest paths) linking two nodes passing through a third point v_k with respect to all geodesics between them.

Definition 2. *If $g_{i,j}$ denotes the number of geodesics connecting v_i, v_j and $g_{i,j}(v_k)$ the number of those shortest paths passing trough v_k, the betweenness index of v_k is given by*

$$C_B(v_k) = \sum_{i<j}^{n} \sum_{j=1}^{n} \frac{g_{i,j}(v_k)}{g_{i,j}}$$

A node with high value of betweenness is also a communication hub, having the capacity to control a significant part of the flow of information in the network due to the great proportion of nodes communicated through it.

Definition 3. *The closeness index of a node v_k is given by*

$$C_C(v_k) = \frac{1}{\sum_{i=1}^{n} d_{i,k}}$$

where $d_{i,k}$ denotes the distance between nodes v_i, and v_k.

A node with high closeness is an important communication node, related to efficiency [34] and minimal cost in communication, due its proximity to other nodes in the network.

The three structural point centrality measures described above strongly depend on the network size and have normalized versions. In section 5, we compute and plot that measures for simulated networks and apply Freeman's normalization in order to compare centrality in networks removing network size effect.

Centrality network indexes (or centralization indexes) quantify the homogeneity of point centrality of all the nodes in the network. Networks with high level of centralization have great differences between point centrality value of the most central point and the others. Networks with low level of centralization have homogeneous point centrality values, which are around the value of the most central node.

Definition 4 (cf. Freeman [33]). *Freeman's normalized indexes of network centralization measures are defined as follows:*

$$C_X = \frac{\sum_{i=1}^{n} C_X^\star - C_X(v_i)}{\max_{G\in G_n} \sum_{i=1}^{n} C_X^\star - C_X(v_i)}$$

where $C_X(-)$ is a point centrality measure, $C_X^\star$ its value at the most central node and G_n is the set of networks of size to n. Note that C_X takes values between 0 and 1.

It is evident that from the point of view of degree, betweenness and closeness point centrality, the center of the star is the most central node in all three cases. The star is the network in which the maximum value in the denominator is reached.

Theorem 1. *Degree-based centrality measures are computed as follows:*

(i) Degree network centrality

$$C_D = \frac{\sum_{i=1}^{n} \left(C_D^\star - C_D(v_i)\right)}{(n-1)(n-2)}$$

(ii) Betweenness network centrality

$$C_B = \frac{2\sum_{i=1}^{n} \left(C_B^\star - C_B(v_i)\right)}{(n-1)^2(n-2)}$$

(iii) Closeness network centrality

$$C_C = \frac{(2n-3)\sum_{i=1}^{n} \left(C_C^\star - C_C(v_i)\right)}{(n-2)}$$

Proof. (i) and (ii) see Freeman [33]. (iii) is a slight modification with respect to its counterpart Freeman's coefficient. Considering point closeness centrality as $C_C(v_i) = 1/\sum d_{i,j}$ it's easy to proof that the maximum value of $\sum_{i=1}^{n} C_C^\star - C_C(v_i)$ is $\frac{n-2}{2n-3}$. □

3 Dynamic extensions

Hiding network topology can be important in order to protect node's identity or other confidencial parameters [35]. Our goal in this paper is to hide information about the topology of a given network by adding ancillary nodes. To be concise, given a network (fig. 1), we add a new node A_6 and its connections to the remaining nodes (fig. 2).

Note that this manipulation is cheap and possible in virtual environments where nodes are virtual machines in a cloud.

Some topological properties may be directly shifted by means of such an operation because in particular the characteristic polynomial of the graph and hence its roots (spectrum) are manipulated.

Note also that in such a dynamic enlargement both the number of ancillary nodes added as well as their connections have to be decided.

Definition 5. *A dynamic enlargement of a graph $G = (V, E)$ is a new graph $\Gamma(G) = (V', E')$ where $V' \supseteq V$ and $E' \supseteq E$. We say that a dynamic enlargement is connected if $\Gamma(G)$ is a connected graph.*

We deal with adjacency matrices of simple and connected networks. Consequently, this matrices are binary and symmetric, their diagonal entries are equal to zero and there does not exist a node permutation to get a block decomposition of the adjacency matrix. In particular, matrices of above examples are:

Fig. 1. Original Network

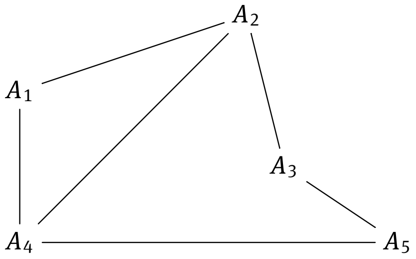
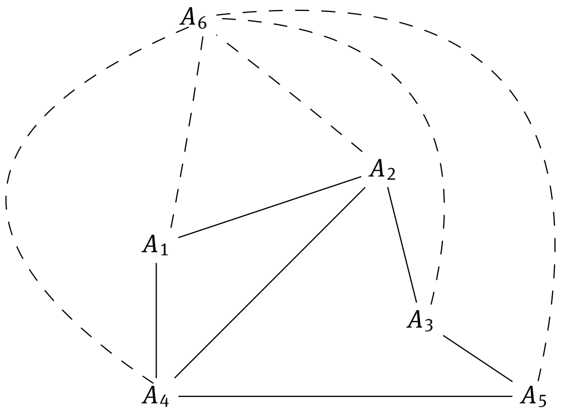


Fig. 2. Extended Network



$$A(G) = \begin{pmatrix} 0\,1\,0\,1\,0 \\ 1\,0\,1\,1\,0 \\ 0\,1\,0\,0\,1 \\ 1\,1\,0\,0\,1 \\ 0\,0\,1\,1\,0 \end{pmatrix}$$

$$A(\Gamma(G)) = \left(\begin{array}{c|c} 0\,1\,0\,1\,0 & \star \\ 1\,0\,1\,1\,0 & \star \\ 0\,1\,0\,0\,1 & \star \\ 1\,1\,0\,0\,1 & \star \\ 0\,0\,1\,1\,0 & \star \\ \hline \star\,\star\,\star\,\star\,\star & 0 \end{array}\right)$$

Dynamic enlargement with one ancilla gives the incidence matrix $A(\Gamma(G))$ where connections $\star$ have to be decided.

4 The dynamic Random Decentering Algorithm (RDA)

Now we propose an algorithm transforming a given graph G into a dynamic enlargement $\Gamma(G)$ of graph G with the following principles:

Principle 1. ***Minimal cost.*** *Algorithm does add one single ancillary node in each step. This minimizes the manipulation cost in physical networks. However we can design a significant number of ancillae by means of running algorithm in a loop.*

Principle 2. ***Topological properties preservation. Density case.*** *Enlargements should have similar topological features to original networks. Nevertheless, topological measures are related one with another. It is not possible to preserve more than one of them with the extension process. For example, average degree $\langle k \rangle$ and density d are widely used to analyze the connectivity of a network. Nevertheless $\langle k \rangle = d(n-1)$. Thus an algorithm preserving $\langle k \rangle$ returns sparse networks with small density values, while if d is preserved, $\langle k \rangle$ increases proportionally to $(n-1)$. In this work, we propose dynamic enlargements $\Gamma(G)$ having similar density of connections to G.*

Principle 3. ***Centrality and centralization reduction.*** *Distribution of centralities in network $\Gamma(G)$ should be more homogeneous than in network G hence centralization of network decreases each time algorithm is performed.*

Principle 4. ***Algorithm protection.*** *Dynamic enlargement links should be obtained by a random process.*

Notations used in the sequel are detailed next: A is the adjacency matrix of original graph $G = (V, E)$, $n = \#V$ is the size of matrix A, $m = \frac{1}{2}\|A\|_1$ is the number of edges, and parameter p is the proportion of ancillae we add.

We will obtain a succession of dynamic extensions $A = A(0), A(1), \ldots, A(pn)$ of the original graph to reach the number of required ancillae; parameters $n(t), m(t)$ are referred to the correspondent extension with $n = n(0)$, $m = m(0)$. We also use the notation $\Delta m(t) = m(t) - m(t-1)$ for the number of links we add in the step $A(t-1) \to A(t)$. Requirements we stated for our algorithm yield the following equalities related to the parameters:

Principle 1 implies

$$n(t) = n(t-1) + 1$$

then, by recurrence

$$n(t) = n(0) + t \tag{1}$$

Principle 2 establishes that

$$\frac{2m(0)}{n(0)\,(n(0)-1)} = \frac{2m(t)}{n(t)\,(n(t)-1)}$$

It follows that

$$m(t) = m(0)\frac{(n(0)+t)\,(n(0)+t-1)}{n(0)\,(n(0)-1)}$$

thus, the number of links added in each step $\Delta m(t) = m(t) - m(t-1)$ can be approximated by the nearest integer to the quotient

$$\frac{2m(0)\,(n(0)+t-1)}{n(0)\,(n(0)-1)} \tag{2}$$

On the other hand, to follow principles 3 and 4, a multinomial distribution law is applied. A new ancillary node links to a previous node v_i with probability

$$p_i = \frac{1/d_i}{\sum_j 1/d_j} \qquad i = 1, 2, \ldots, n(t-1) \tag{3}$$

at each step t.

In order to avoid multilinks we apply this probability model step by step, recalculating probabilities p_i after each selection (no replacement probability model).

We provide a self-explain pseudocode of RDA (see Algorithm 1). After performing RDA algorithm we obtain all the adjacency matrices of the successive dynamic extensions and its basic topological properties (size, number or links and node degrees).

The following sections are devoted to obtaining experimental results. The centrality measures of the enlarged networks will be computed and plotted as functions of the number of added nodes. It is worth to remark here that original graphs are chosen to be scale-free graphs randomly obtained by BA algorithm.

5 Experiments

In this section section we carry out dynamic extensions over simulated networks and analyze numerical changes on centrality and centralization measures and their relations. We have conducted the experiments on simulated data that closely model technological data networks we are interested in.

5.1 Real world networks. Scale-free

Several models to generate real world networks have been proposed since the 50's [36–41]. Paul Erdós and Alfred Rényi modeled random networks, characterized by having nodes with approximately the same degree, showing low level of centralization and being robust against target attacks but vulnerable to random attacks.

Algorithm 1. ***Random Decentering Algorithm RDA***

```
Require: $A = A(0)$ incidence matrix
Require: $p > 0$ proportion of ancillary nodes
Ensure: $t$ number of steps and $A(t)$ dynamic enlargement of $A(0)$ following principles 1-4.
n(0) ← n and t ← 0
while $i \leq n(0)$ do
    $d(0; i) = \sum_{k=1}^{n(0)} A(0; k, i)$
end while
$m(0) = \frac{1}{2} \sum_{i=1}^{n(0)} d(0; i)$
while $t < pn$ do
    $t \leftarrow t + 1$
    $n(t) \leftarrow n(t) + 1$
    $\Delta m(t) = \left[ \frac{2m(0)n(t-1)}{n(0)(n(0)-1)} \right]$
    choose $\Delta m(t)$ elements without replacement $\{k_1, \ldots, k_{\Delta m(t)}\} \subset \{1, \ldots, n(t-1)\}$ with probabilities
    $p(t, i) = \frac{1/d(t-1,i)}{\sum_{i=1}^{n+t-1} 1/d(t-1,i)}$
    $A(t; i, j) = A(t-1, i, j)$ for $i, j = 1, \ldots, n(t-1)$
    if $k \in \{k_1, \ldots, k_{\Delta m(t)}\}$ then
        $A(t; n(t), k) = A(t; k, n(t)) = 1$
        $d(t, i) = d(t-1, i) + 1$
    else
        $A(t; n(t), k) = A(t; k, n(t)) = 0$
        $d(t, i) = d(t-1, i)$
    end if
    $d(t, n(t)) = \Delta m(t)$
end while
return $A(t); d(t, i); n(t); m(t)$    $i = 1, \ldots, n(t)$    $t = 0, \ldots, np$
```

On the other hand, scale free networks have a few nodes with a great number of connections (called hubs), whereas most nodes have small degree. Networks with this characteristics might appear in complex and computational networks, have high level of centralization and are more vulnerable to coordinated attacks than to random threats.

5.2 Simulation data

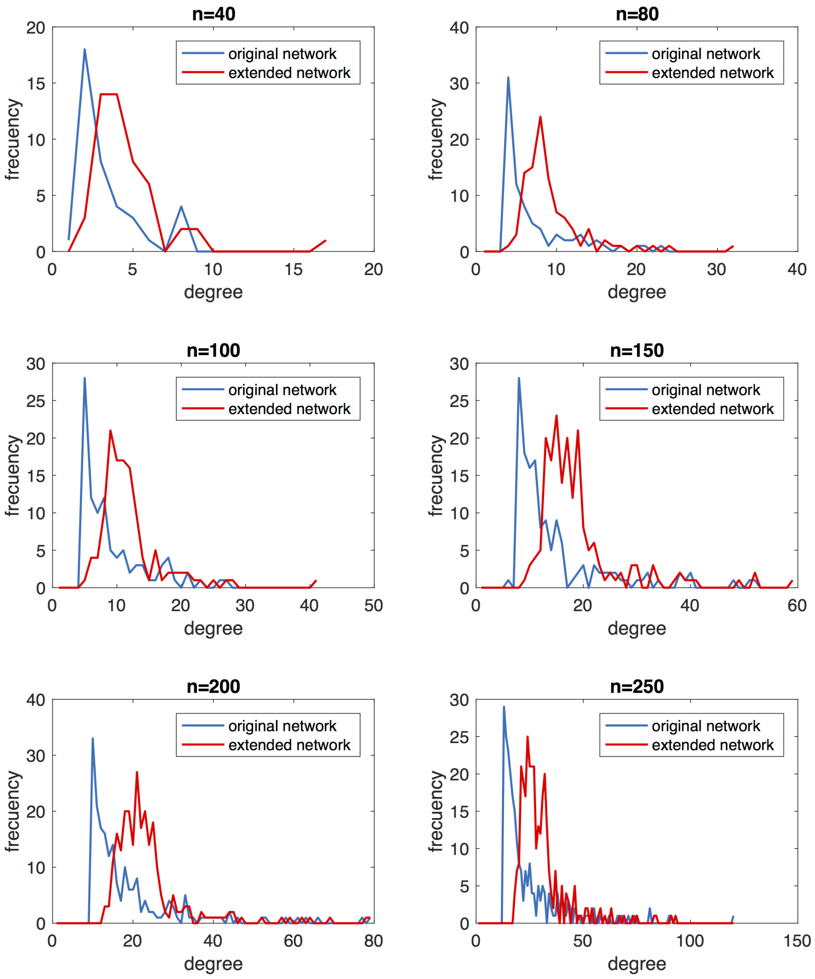
Simulated data of our experiment were eventually obtained by BA scale free algorithm [37] with SFNG m-file of MATLAB. We carried out BA simulations and obtained networks with different size but similar density we call “original networks". Random Decentralized Algorithm (RDA) was applied with different percentage p of ancillary points added. We shall refer to the dynamic extensions as “extended networks" at each algorithm step t. Degree point centrality index was obtained for single nodes in all the networks, original and extended and degree distribution plots were also performed. Finally, network centrality indexes were obtained together with their evolution graphs over all the algorithm steps from 0% to 100% of ancillary nodes added.

Six original networks were randomly obtained by BA algorithm and several sizes n=40, 80, 100, 150, 200, 250. Parameters were selected in order to get networks with similar density $d \approx 0.1$. RDA algorithm was performed for each original matrix A with proportion $p = 1$ (100% of ancillary nodes added). Finally, sequences of matrices $A(t)$ were stored from $t = 0$ (adjacency matrix of the original network) to $t = [p * n(0)] = n(0)$.

5.3 Results

In order to compare the point degree distribution for original and extended networks, frequency plots were performed for 25% and %75 of ancillary nodes added (figs. 3, 4). While original scale-free networks fit to decreasing exponential shapes, extended networks show right displacement and tendency to converge towards bell shapes, typical for random networks, witch are robust to target attacks.

Fig. 3. Frequency plots for point degree with 25% of nodes added



Centralization measures were computed for all original and extended matrices $A(t)$. As we have mentioned before, all measures have been normalized. Consequently, potential network-size effect has been removed. This reveals that centralization decreases as t incerases. Table 1 contains absolute values of degree, betweenness and closeness centralization measures respectively.

As we remarked above, original simulated networks were randomly obtained by BA algorithm. In spite of the random character of BA algorithm, networks obtained by RDA have similar centralization values of degree (varying between 0.304 and 0.387) and closeness (between 0.272 and 0.417). Nevertheless, betweenness network centrality tends to decrease when n increases (from 0.071 and 0.448). On the one hand, the range of variation of betweenness values is greater in general for this index [4] than the others. Furthermore, the inherent nature of BA algorithm tends to get short paths between nodes if seeds have small size compared to the final size of the scale free network. This fact implies that nodes have low point betweenness centrality and the same occurs with betweenness network index.

Fig. 4. Frequency plots for point degree with 75% of nodes added

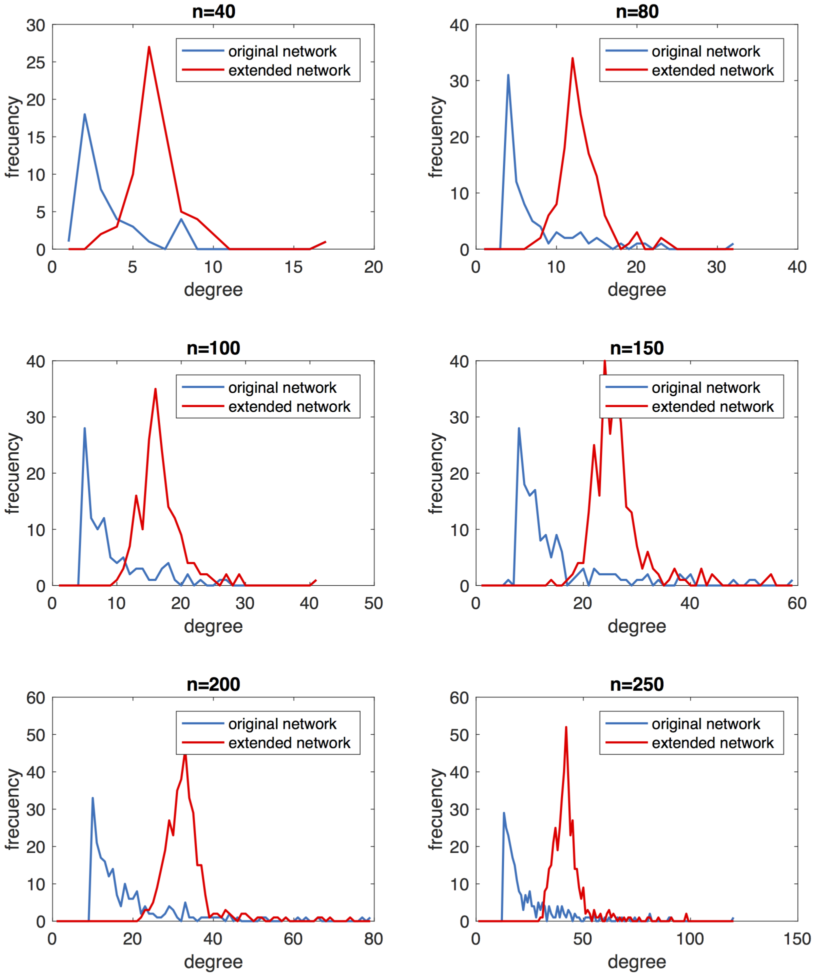


Table 1. Absolute values of degree network centrality (C_D), betweenness network centrality (C_B) and closeness network centrality (C_C) for the original simulated networks and extended networks with 0%, 25%, 50%, 75% and 100% of ancillary points.

	C_D					C_B					C_C				
n	0%	25%	50%	75%	100%	0%	25%	50%	75%	100%	0%	25%	50%	75%	100%
40	0.359	0.270	0.205	0.160	0.126	0.448	0.327	0.208	0.141	0.094	0.417	0.360	0.304	0.255	0.211
80	0.319	0.237	0.188	0.147	0.122	0.177	0.105	0.061	0.036	0.026	0.371	0.309	0.255	0.208	0.177
100	0.324	0.241	0.183	0.142	0.117	0.160	0.086	0.047	0.027	0.018	0.361	0.276	0.218	0.171	0.137
150	0.299	0.221	0.171	0.144	0.123	0.073	0.037	0.022	0.014	0.010	0.284	0.210	0.156	0.122	0.097
200	0.304	0.225	0.174	0.141	0.116	0.063	0.031	0.017	0.011	0.007	0.272	0.193	0.138	0.104	0.079
250	0.387	0.294	0.228	0.184	0.150	0.071	0.036	0.019	0.012	0.008	0.314	0.220	0.159	0.120	0.094

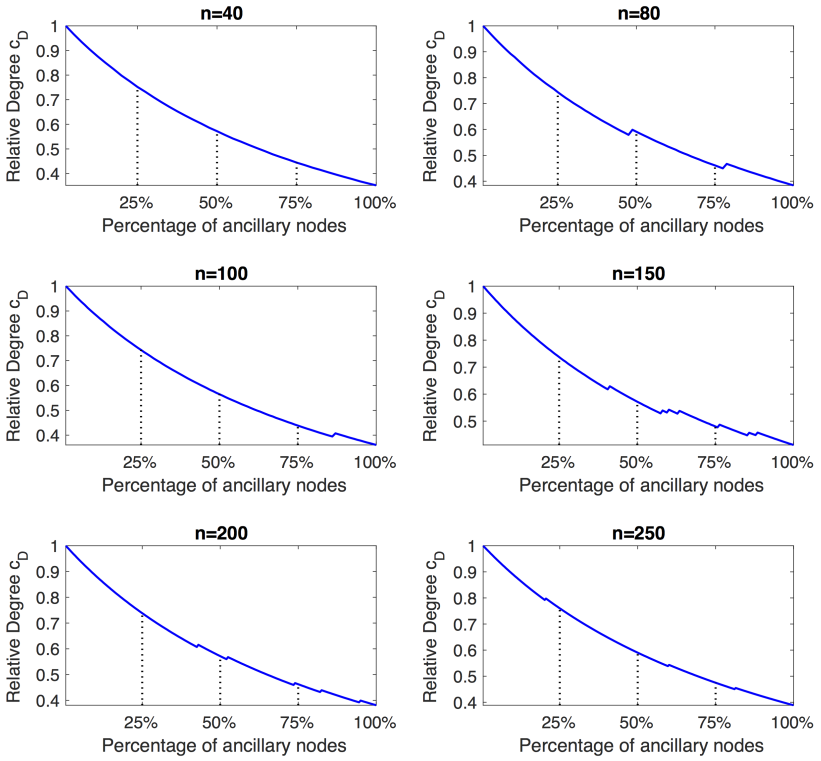
In order to analyze variations of centralization measures as a function of t, relative values of centralization indexes were obtained (table 2) and evolution plots were obtained in each case (figs. 5, 6 and 7). Values with 0% of ancillary points added are reference values and hence omitted in this table.

RDA gets significant increasing values of centralization indexes in all cases. C_C and C_B tend to decrease faster for original networks with a large size while C_D does not. On the other hand, RDA shows greater variation of relative values of C_B. With 25% of nodes added, C_D and C_C have a decrease between 70% and 86%, and between 50% and 72%. If 50% of ancillary nodes are added to the networks with RDA, the percentage of decrease varies between 50% and 72% for C_D and C_C, while C_B decreases between 27% and 46%. The

Table 2. Relative values of degree network centrality (C_D), betweenness network centrality (C_B) and closeness network centrality (C_C) for the original simulated networks and extended networks with 25%, 50%, 75% and 100% of ancillary points.

	C_D				C_B				CC_C			
n	**25%**	**50%**	**75%**	**100%**	**25%**	**50%**	**75%**	**100%**	**25%**	**50%**	**75%**	**100%**
40	75.199	57.214	44.452	35.204	72.894	46.382	31.351	21.031	86.366	72.771	61.099	50.517
80	74.399	59.116	46.125	38.406	59.126	34.445	20.133	14.405	83.275	68.708	56.011	47.798
100	74.316	56.491	43.864	36.056	53.971	29.660	17.188	11.477	76.319	60.404	47.387	37.978
150	73.734	57.178	48.012	41.164	51.381	29.629	19.740	14.104	73.950	55.104	42.982	34.000
200	73.833	57.170	46.260	38.111	70.937	50.933	38.404	29.053	70.937	50.933	38.404	29.053
250	76.000	59.011	47.501	38.869	70.073	50.478	38.297	29.960	70.073	50.478	38.297	29.960

Fig. 5. Relative Degree network centrality



greatest variation is observed with 100% of nodes added, 36% to 50% for C_D and C_C and 11% to 21% for betweenness C_B.

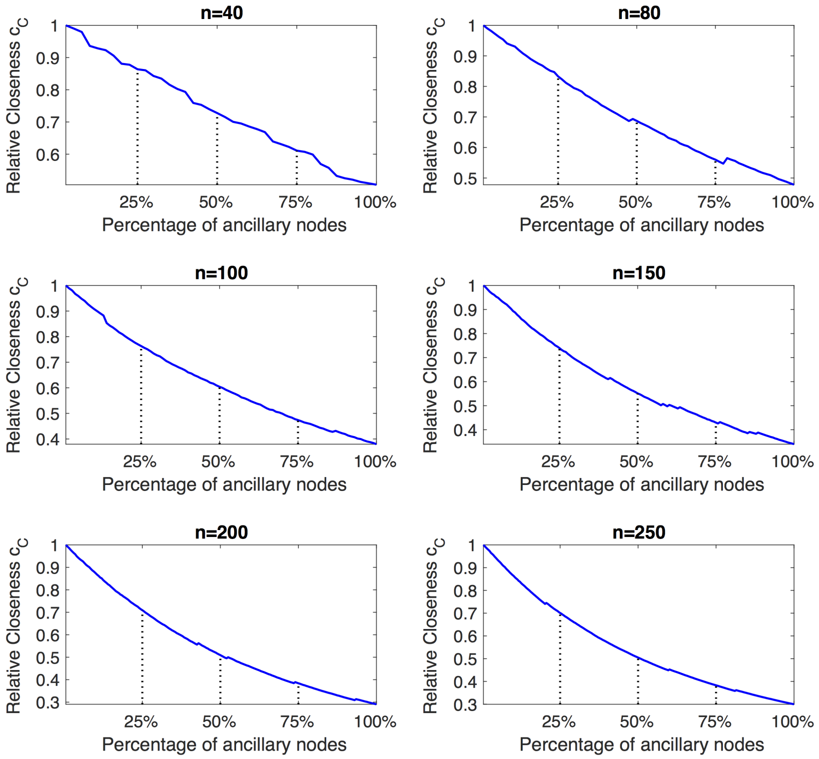
6 Conclusions

In this paper we give an algorithm that modifies the character of a given scale-free network by means of adding ancillary nodes.

First experimental results show decreasing values of network centralization measures. This provides a possible tool for protecting communication networks from vulnerabilities exploiting topology of underlying graph, and consequently RDA provides an answer to Challenge 1 and Challenge 2.

It remains to face Challenge 3 by estimating the consequences of each single change in terms of data preserving. Note that RDA algorithm improve node‘s privacy becouse centrality decreases and therefore principal nodes are hidden within new ancillary nodes. However, more experiments are needed in the future.

Fig. 6. Relative Closeness network centrality



We note once again that a dynamic treatment like RDA is cheap to implement on networks of virtual machines in the cloud, but its implementation on cyber-physical environments needs previous study in order to evaluate the cost of introducing each ancillary node and each ancillary link.

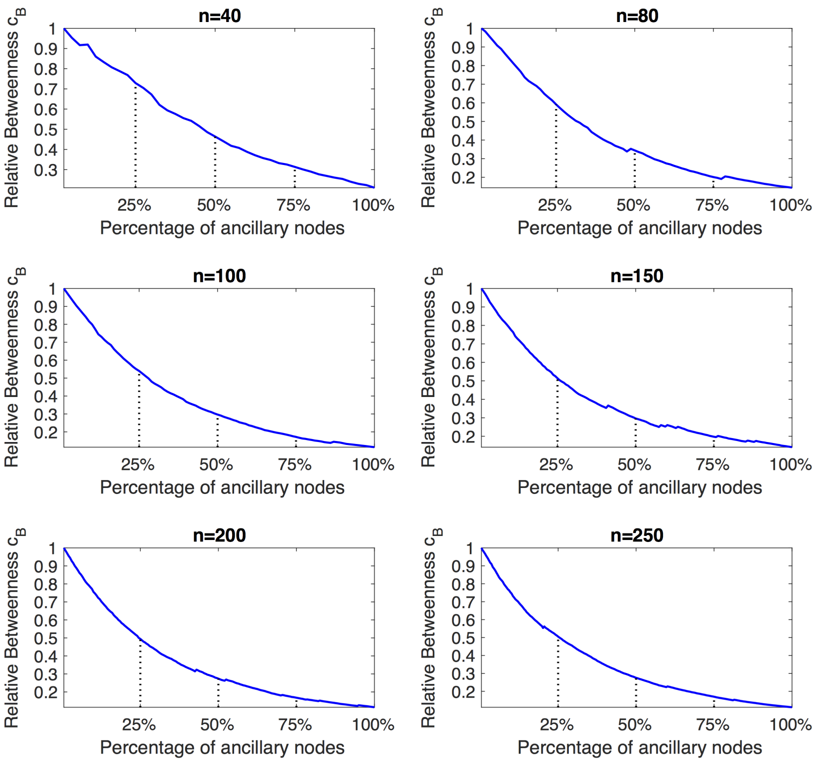
Some pending tasks are (1) to get, if possible, a true random network from a scale-free one by means of adding ancillary nodes; (2) to compute effectively how many ancillary nodes are necessary to hide main properties of original network; (3) to perform experiments of order 10^n nodes, $n = 4, 5, ..$ in order to check our procedure in real world networks.

Acknowledgement: This research was partially supported by INCIBE (Spanish Institute of Cybersecurity) under grant X47.

References

[1] Brandes U., Erlebach T., Network Analysis, Lecture Notes in Computer Science, 1989, 3418, Springer.
[2] Borgatti S. P., Dynamic Social Network Modeling and Analysis, Workshop Summary and Papers, 2003.
[3] Borgatti S. P., Ajay M., Brass D. J., Labianca G., Network Analysis in the Social Sciences, Science, 2009, 892-895.
[4] Freeman L., The development of social network analysis, A Study in the Sociology of Science, 2004, 1.
[5] Seary A. J., Richards W. D., Dynamic Social Network Modeling and Analysis: Workshop Summary and Papers, Spectral methods for analyzing and visualizing networks: an introduction, 2003.
[6] Lozano M., García-Martínez C., Rodríguez F. J., Trujillo H. M., Optimizing network attacks by artificial bee colony, Information Sciences, 2017, 377, 30-50.
[7] Zhang B., Horvath S., A general framework for weighted gene coexpression network analysis, Statistical Applications in Genetics and Molecular Biology, 2005, 4.
[8] Guida M., Funaro M., Topology of the Italian airport network: A scale-free small-world network with a fractal structure?, Chaos, Solitons and Fractals, 2007, 31, 3, 527-536.

Fig. 7. Relative Betweenness network centrality



[9] Pedroche F., Romance M., Criado R., Some rankings based on PageRank applied to the Valencia metro, International Journal of Complex Systems in Science, 2016.

[10] Balci M. A., Fractional virus epidemic model on financial networks, Open Mathematics, 2016, 14, 1074-1086.

[11] Chapanond A., Krishnamoorthy M. S., Yener B., Graph theoretic and spectral analysis of Enron email data, Computational and Mathematical Organization Theory, 2005, 11, 265-281.

[12] Cvetović D., Simić S., Graph spectra in Computer Science, Linear Algebra and its Applications, 2011, 434, 6, 1545-1562.

[13] Dang-Pham D., Pittayachawan S., Bruno V., Applications of social network analysis in behavioural information security research: Concepts and empirical analysis, Computers & Security, 2017, 68, 1-15.

[14] Diestel R., Graph Theory. Graduate Texts in Mathematics, Springer-Verlag, 2005.

[15] Benito del Pozo P., Serrano N., Marqués-Sánchez P., Social networks and healthy cities: spreading good practices based on a spanish case study, Geographical Review, 2016.

[16] Pasqualetti F., Bicchi A., Bullo F., A graph-theoretical characterization of power network vulnerabilities, 2011, 3918-3923.

[17] Sridhar S., Hahn A., Govindarasu M., Cyber-Physical System Security for the Electric Power Grid, Proceedings of the IEEE, 2012, 100, 1, 210-224.

[18] Puzis R., Yagil D., Elovici Y., Braha D., Collaborative attack on Internet users' anonymity, Internet Research, 2009, 19, 1, 60-77.

[19] Liu K., Das K., Grandison T., Kargupta H., Privacy-Preserving Data Analysis on Graphs and Social Networks, Next Generation of Data Minning, 2008.

[20] Arsič B., Cvetović D., Simić S., Škarić M., Graph spectral techniques in computer sciences, Applicable Analysis and Discrete Mathematics, 2012, 6, 1, 1-30.

[21] Cvetković D., Rowlinson P., Simić S., Eigenspaces of graphs, Eigenspaces of graphs, 1997, 66.

[22] Simic S. Andelic M., DaFonseca C. M., Zivkovic D., On the Multiplicities of Eigenvalues of Graphs and Their Vertex Deleted Subgraphs: Old and New Results, Electronic Journal of Linear Algebra, 2015, 30, 85-105 66.

[23] Shang, Y., Impact of self-healing capability on network robustness, Phys Rev E Stat Nonlin Soft Matter Phys., 2015, 91, 4, 042804.

[24] Shang, Y., Effect of link oriented self-healing on resilience of networks, Journal of Statistical Mechanics: Theory and Experiment, 2016, 8, 083403.

[25] Shang, Y., Localized recovery of complex networks against failure, Scient. Rep., 2016, 6, 30521 EP -.

[26] Holme P., Kim B. J., Yoon C., Han S. K., Attack vulnerability of complex networks, Phys. Rev. E, 2002, 65, 056109.

[27] Iyer S., Killingback T. and Sundaram B. Wang Z., Attack Robustness and Centrality of Complex Networks, PLoS ONE, 2013, 8, 4, e59613.

[28] Shang Y., Robustness of scale-free networks under attack with tunable grey information, EPL (Europhysics Letters), 2011, 95, 28005.

[29] West, D. B., Introduction to Graph Theory, Prentice Hall, 2001.

[30] Borgatti S. P., Everett M. G., A Graph-theoretic perspective on centrality, ESocial Networks, 2006, 28, 466-484.

[31] Bounova G. de Weck O., Overview of metrics and their correlation patterns for multiple-metric topology analysis on heterogeneous graph ensembles, Phys. Rev. E, 2012, 85, 016117.

[32] Canright G. S., Engø-Monsen K., Some Relevant Aspects of Network Analysis and Graph Theory, Elsevier, 2008, 361-424.

[33] Freeman L. C., Centrality in social networks conceptual clarification, Social Networks, 1978, 215-239.

[34] Latora V. Marchiori M., Efficient Behavior of Small-World Networks, Phys. Rev. Lett., 2001, 87, 19, 198701.

[35] LaVigne R. Zhang C. D. L., Maurer U. Moran T., Mularczyk M., Tschudi D., Topology-Hiding Computation Beyond Semi-Honest Adversaries, IACR Cryptology ePrint Archive, 2018, 255.

[36] Barabási A. L., Scale-Free Networks: A Decade and Beyond, Science, 2009, 325, 5939, 412-413.

[37] Barabási A. L., Réka A., Emergence of Scaling in Random Networks, Science, 1999, 286, 5439, 509-512.

[38] Barabási A. L., Réka A., Hawoong J., Mean-field theory for scale-free random networks, Physica A: Statistical Mechanics and its Applications, 1999, 272, 1-2, 173-187.

[39] Erdös P., Rényi A., On random graphs I, Publicationes Mathematicae (Debrecen), 1959, 6, 290-297.

[40] Meghanathan N., A Model for Generating Random Networks with Clustering Coefficient Corresponding to Real-World Network Graphs, International Journal of Control and Automation, 2016, 9, 163-176.

[41] Pedroche F., Criado R., Garcia E., Romance M., Matrix growth models based on centrality measures: a first analysis, International Journal of Complex Systems in Science, 2011, 1, 124-128.