OPEN ACCESS

The concept of roles in complex networks

To cite this article: K J Parousis-Orthodoxou *et al* 2013 *J. Phys.: Conf. Ser.* **410** 012080

The concept of roles in complex networks

K.J. Parousis-Orthodoxou, M. Stamos and D.S. Vlachos

Department of Computer Science and Technology,
Faculty of Sciences and Technology,
University of Peloponnese,
GR-221 00 Tripolis, Greece

E-mail: `dvlachos@uop.gr`

Abstract. This paper is an analytic process of introducing the concept of roles in a complex network. Each node has a specific connectivity pattern to the rest of the nodes in the network. There are many nodes that have the same number of connections and also the same type. In this case we say that two nodes can have the same role. The work described here, shows the process of identifying the roles of all the nodes in a network and grouping them accordingly.

1. Introduction

Numerous systems in modern day life can be assigned to complex networks and complex network's properties. From the protein network within the human body to the configuration of the Internet or the World Wide Web, all can be treated with respect to their topological configuration of nodes connected to eachother with the use of distinct edges. The variety and diversity of these networks (and many more) are what prompted the scientific community to investigate the mechanisms that explore their topology [1], [2], [3], and [4].

In modern day Computer Science as well as various different fields, the use of complex network modeling and experimentation is widespread. Within the past decade, various efforts have taken place for the modeling and analysis of complex systems. These efforts have been very fruitful, both in manufacturing networks by following certain rules (such as the preferential attachment rule) [5], as well as analyzing real data and mining emerging phenomena such as the small-world property, which implies that the average topological distance between any two nodes within a network, increases very slowly along with the evolution of the network.

By taking a closer look into a complex network, it is noticeable that each node has a specific position in the graph, not necessarily topologically, but with respect to the other nodes of the network. It is also easy to see that there are many nodes that may have the same positioning within the graph. The algorithm presented here will adopt the methodology to define the positioning of each node in the network, newly introducing roles.

2. The use of roles

The concept of *roles* is explained by analyzing the function of a node within the network, in accordance with the other nodes. Every node has a role in the network, whereas many nodes can have the same role, with regard to their own connectivity.

In order to find the role that can be assigned to every node, we must first divide the nodes into groups, and create each *connectivity vector* to the other groups. Each connectivity vector

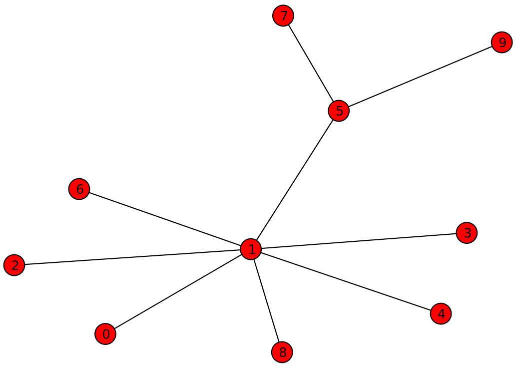

Figure 1. An example network for role identification.

records the number of connections that the group has, to all the rest (including itself) and also the number of nodes that reside in that group (the final metric of the vector).

For better understanding of the entire process, an example network is shown in **Figure 1**. The algorithm tries first to categorize the nodes into groups. The process begins by creating N distinct groups and assigning a single node to each one. Then the connectivity vectors of the network are generated for each group (initially for each node). This instance of the connectivity vectors of the example network is shown in **Table 1**.The second step compares the connectivity vectors of the groups and if it finds that two vectors are the same (apart from the last metric that shows the number of nodes in the group), it merges the two groups into one. After the merging of the two groups takes place, the algorithm recalculates the connectivity vectors and starts searching for the next groups that may share a common connectivity vector. When there can be no more merging among the groups, the algorithm has finished the grouping stage.

CV No	Grp 0	Grp 1	Grp 2	Grp 3	Grp 4	Grp 5	Grp 6	Grp 7	Grp 8	Grp 9	No
0	0	1	0	0	0	0	0	0	0	0	1
1	1	0	1	1	1	1	1	0	1	0	1
2	0	1	0	0	0	0	0	0	0	0	1
3	0	1	0	0	0	0	0	0	0	0	1
4	0	1	0	0	0	0	0	0	0	0	1
5	0	1	0	0	0	0	0	1	0	1	1
6	0	1	0	0	0	0	0	0	0	0	1
7	0	0	0	0	0	1	0	0	0	0	1
8	0	1	0	0	0	0	0	0	0	0	1
9	0	0	0	0	0	1	0	0	0	0	1

Table 1: The connectivity vectors of the example network, before the first iteration.

The example network has 10 nodes and 9 edges. The first step of the process, creates 10 groups and places a single node of the network into them. Then, it checks the connections of each group (node for now) and creates the connectivity vector for each group. The starting vectors are taken unchanged from each line of the adjacency matrix of the network, with the addition of the number of nodes for each group at the end of the vector (*one* at first). After the initial step, the algorithm identifies that the groups with number *0* and *2*, which initially

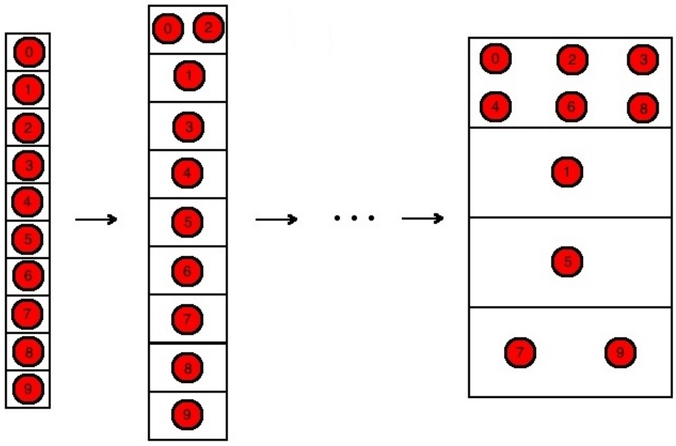

Figure 2. The process of the grouping algorithm.

contain the nodes 0 and 2 respectively, have the same connectivity vector. The two groups have only one connection with the group *1*. Then it increments the final metric of the group *0* and deletes group *2*. After that, it recalculates the connectivity vectors with one group less. On the next iteration, the groups *0* and *3* are found to have the same connectivity vector and so, the group *3* is deleted and the counter of group *0* is incremented by 1. This process is continued on the same pattern until no further grouping can be performed.

So, at the end of the final iteration, the number of groups is reduced to 4. These groups are shown in **Figure 2** and the nodes that each group contains. Note that the numbering of the nodes here is for explanatory purposes and the algorithm only focuses on the number of nodes in each group. The final connectivity vector of the example graph is shown in **Table 2**.

CV No	Grp 0	Grp 1	Grp 2	Grp 3	No
0	0	6	0	0	6
1	6	0	1	0	1
2	0	1	0	2	1
3	0	0	2	0	2

Table 2: The connectivity vectors of the example network at the end of the process.

3. Testing the algorthm

The network that will be used for the testing of the algorithm is undirected and unweighted and it is provided by Duncan Watts and Steven Strogatz [6]. The graph presents information about the topology of the Western States Power Grid of the United States of America. The network has 4941 nodes and 6594 edges. The mean degree of the graph is 2.67 whereas the maximum degree that can be found is 19.

After performing the aforementioned operation to our test network, the result was the creation of table (of similar form to **Table 2**, but significantly larger). This table contains the connectivity vectors of all the groups that can be created that would hold information about each role that might exist in the network. The specific network generated a table with 3954 groups and equivalent roles. The number of nodes that assume these roles are shown in **Figure 3**

It is fairly clear from the illustration of the connectivity vector of the specific network that

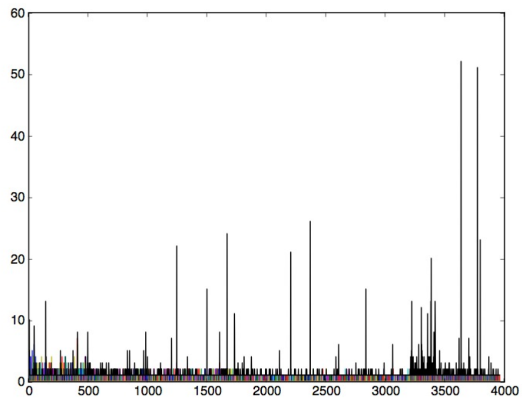

Figure 3. The role distribution of the test network.

there is a significant number of roles that are assumed by a single node (and that is why so many roles can be found in the network) and also that there are certain roles that are assumed by many nodes, illustrated by the spikes that emerge from the figure.

4. Future work

The concept of roles and their distribution over a network can provide useful information regarding the connectivity of the nodes. This algorithm allows the user to manipulate a network further when analyzing its role distribution, in several ways such as replicating or scaling the network. This part can be extensively studied in a different project an may yield very promising results.

Acknowledgment

The project is implemented under the Operational Program "Education and Lifelong Learning" and funded by the European Union (European Social Fund) and by Greek national resources.

References

[1] L.A.N. Amaral, A. Scala, M. Barthlemy, H.E. Stanley, Proc. Natl. Acad. Sci. USA, 97 (2000), p. 11149
[2] S.N. Dorogovtsev, J.F.F. Mendes, **Evolution of Networks: From Biological Nets to the Internet and WWW**, Oxford University Press, Oxford (2003)
[3] R. Pastor-Satorras, A. Vespignani, **Evolution and Structure of the Internet: A Statistical Physics Approach**, Cambridge University Press, Cambridge (2004)
[4] M.E.J. Newman, Phys. Rev. Lett., 89 (2002), p. 208701
[5] R.Albert, A.-L. Barabasi, Rev.Mod.Phys., 74(2002), p.47
[6] D.J. Watts, S.H. Strogatz, Nature, 393 (1998), p. 440