Community detection: comparison among clustering algorithms and application to EEG-based brain networks*

M.G. Puxeddu, *Member, EMBS*, M. Petti, *Member, EMBS*, F. Pichiorri, F. Cincotti, D. Mattia and L. Astolfi, *Member, EMBS*

Abstract— **Community structure is a feature of complex networks that can be crucial for the understanding of their internal organization. This is particularly true for brain networks, as the brain functioning is thought to be based on a modular organization. In the last decades, many clustering algorithms were developed with the aim to identify communities in networks of different nature. However, there is still no agreement about which one is the most reliable, and to test and compare these algorithms under a variety of conditions would be beneficial to potential users. In this study, we performed a comparative analysis between six different clustering algorithms, analyzing their performances on a ground-truth consisting of simulated networks with properties spanning a wide range of conditions. Results show the effect of factors like the noise level, the number of clusters, the network dimension and density on the performances of the algorithms and provide some guidelines about the use of the more appropriate algorithm according to the different conditions. The best performances under a wide range of conditions were obtained by Louvain and Leicht & Newman algorithms, while Ronhovde and Infomap proved to be more appropriate in very noisy conditions. Finally, as a proof of concept, we applied the algorithms under exam to brain functional connectivity networks obtained from EEG signals recorded during a sustained movement of the right hand, obtaining a clustering of scalp electrodes which agrees with the results of the simulation study conducted.**

I. INTRODUCTION

Many biological systems can be modeled as complex networks composed by different parts interacting together. The study of the statistical and structural properties of such networks (also called graphs) can lead to important insights into physiological and functional mechanisms. Community structure [1] is a property of complex networks that are organized in cohesive groups of nodes (communities or clusters) which can be related to specific functions of the system. Uncovering such communities has assumed an increasingly important role in the study of complex networks, including brain structural and functional ones.

Among the methods proposed in literature [2] a first family of algorithms is based on the optimization of modularity [3], a quality function that estimates the partitioning of the network in contrast with a null model (a surrogate network, equivalent to the real one in terms of degree distribution, but randomly rewired). Algorithms based on modularity criteria, and characterized by different null models, were proposed: Leicht & Newman [4], for directed networks; Danon [5], aimed to favor the detection of small communities; Ronhovde [6], which has a multi-resolution approach, and Louvain [7], which is fast on networks of any size. Other algorithms are based on network stability, in particular the one proposed by Le Martelot [8] that we will call Stability Optimization, and Infomap, an algorithm which seeks a network partition that bring to the minimum description length (defined by the so-called map equation) of a random walk taking place on the network [9].

Such large amount of clustering algorithms has required a comparative evaluation of their performances. Some benchmark graphs were suggested, namely simulated networks in which the grouping into clusters is known in advance. The one introduced in [1] generates graphs with a fixed number of clusters (CN=4) and with the same degree for each node. The one introduced in [10] is able to simulate realistic networks with nodes degree and clusters size following power laws; however the user cannot control all the features, like the clusters number or graph density, to adapt the benchmark to different applications. Also in [11] graph density is not settable by the user and clusters size is fixed (from 10 to 15 nodes for each cluster). Previous comparative analysis [12] were focused on high dimensional networks and they did not explore the dependency of the performances from the specific features of the benchmark graphs used. In this work, we present a comparative analysis of different clustering algorithms based on a simulation study exploiting benchmark graphs generated accounting for a wide range of complex network features (e.g. density, number of communities, dimension), which were systematically varied in a range of conditions, to investigate the performances of the above-listed clustering algorithms under different factors. The algorithms performances, quantified thorough an index of merit, were subjected to an analysis of variance (ANOVA). Finally, we applied all the algorithms to EEG-based brain functional networks obtained during the execution of a motor task by a healthy volunteer, and we report here the differences between the clustering achieved with different approaches.

II. METHODS

A. Graph generation

In this study, we developed a generator of benchmark networks able to simulate a variety of conditions that may be encountered in brain networks. After setting the number of

*Research partially supported by the Italian Ministry of Education, University and Research, project FIRB—Futuro in Ricerca 2013 (RBFR136E24), by Sapienza University of Rome - Progetti di Ateneo 2015 (C26A15N8LZ) and by the Italian Ministry of Healthcare (RF-2011-02348213 and Translational Research Program).

M. G. Puxeddu, M. Petti (Corresponding Autor. Tel: +39065151165. E-mail: manuela.petti@uniroma1.it), F. Cincotti and L. Astolfi are with the Department of Computer, Control and Management Engineering, University of Rome Sapienza, Rome, Italy and with IRCCS Fondazione Santa Lucia, Rome, Italy.

D. Mattia and F. Pichiorri are with Fondazione Santa Lucia, Rome, Italy.

978-1-5090-2809-2/17/$31.00 ©2017 IEEE

nodes (N), the graph density (D), the number of clusters (CN), and the ratio between intra-clusters density and inter-clusters density (dr), the algorithm starts filling an NxN empty adjacency matrix through the following steps:

1) Setting of the size of the CN communities: it is obtained by randomly choosing CN integers, with the constraint that their sum equals the number of nodes N.
2) Wiring of the network: given D and dr, the total number of links in the graph, the number of intra-clusters links (inconn) and the number of inter-clusters links (outconn) are computed. Then the matrix is filled by positioning the connections at random, provided that they fulfill the previous constraints.
3) Checking the absence of isolated nodes inside the clusters: each node must be connected with at least one node of the same community. If not (it could happen since the links are randomly distributed), a link is added to satisfy the constraint. To keep D and dr as previously set, a link between nodes belonging to the same community is deleted.
4) Checking the internal and external degree of each node: this step is computed only if dr≥2, which means having a well-defined community structure; in this case, the algorithm ensures that each node has internal degree (connections with nodes inside its cluster) greater than external degree (connections with nodes outside its cluster). If not (dr<2) weak communities are generated and check on internal and external degree is bypassed.
5) Adding noise (optional): noise is modeled as a random shifting of a certain number of connection, to account for false positives or false negatives that can be met in real complex networks. Thus, the tool allows to additionally set a noise value, corresponding to the percentage of links randomly shifted by the tool.

This procedure returns a directed and unweighted graph. In Figure 1 we show an example of a simulated network with 60 nodes and 4 clusters.

B. Simulation study

To evaluate and compare the performances of the algorithms at the state-of-the-art for the community detection, we tested them on simulated networks generated through the procedure previously described. We varied systematically all the parameters of interest (number of nodes, graph density, number of clusters, ratio between intra-clusters density and inter-clusters density, noise). The range of these factors was chosen taking as reference typical brain networks estimated from electroencephalographic (EEG) recordings. The algorithms performances were evaluated through the Normalized Mutual Information (NMI), one of the most used indices able to estimate the similarity between two objects [13], and in this case between the partitioning made by the algorithms and the imposed community structure. Maximum similarity corresponds to NMI equal to 1.

Hence, an ANOVA for repeated measures was performed to evaluate the effects of factors *Algorithm* (6 levels: Louvain, Danon, Leicht & Newman, Ronhovde, Stability Optimization, Infomap), *Clusters Number* (3 levels: 2, 4, 6) and *Noise* (4 levels: 0%, 10%, 30%, 50%), using the NMI as dependent variable. Tuckey's post-hoc tests were applied to find statistical differences between the levels of the ANOVA factors. *Nodes Number*, *Graph Density* and *dr* were not included in the ANOVA, but used as external factors. In this study we set N=[30, 60, 100], D=[0.1, 0.3, 0.5] and dr=2.

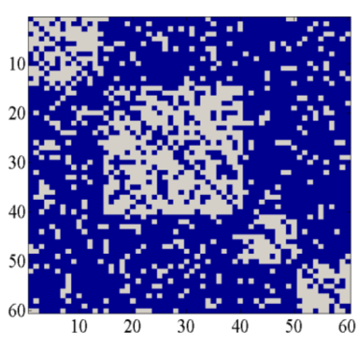

Figure 1. Example of a simulated binary graph characterized by 60 nodes, 4 clusters, graph density equal to 0.3 and ratio between density intra-cluster and inter-cluster equal to 2. The graph is represented through a squared matrix 60x60 in which grey dots are the links between nodes, and blue dots represent the absence of links.

C. Test on EEG data

EEG-based functional brain networks were estimated by means of squared Partial Directed Coherence (PDC, [14]) performed on data recorded from a healthy subject during the execution of a motor task (sustained grasping and extension of the right hand). The subject gave informed consent prior to his participation and the experiments were approved by the local Ethics Committee before data acquisition started. EEG signals were collected using 61 electrodes (according to the extended 10-20 International System). The session was composed by 30 trials of 4 seconds each, in which the subject was asked to repeatedly grasp or extend his right hand. Each trial was temporally identified by a visual stimulus. Preprocessing included band-pass filtering (1-45 Hz), artifact rejection and 4s-epochs segmentation. We computed PDC first considering all 61 channels, and then a subset (29 channels, spatially distributed on the scalp: Fp1, Fpz, Fp2, F7, F3, Fz, F4, F8, FC5, FC1, FC2, FC6, T7, C3, Cz, C4, T8, CP5, CP1, CP2, CP6, P7, P3, Pz, P4, P8, O1, Oz, O2), in four EEG frequency bands defined according to the Individual Alpha Frequency [15] (IAF=10 Hz): theta [IAF-6, IAF-3], alpha [IAF-2, IAF+2], beta [IAF+3, IAF+14] and gamma [IAF+15, IAF+30]. We considered different groups of channels to test clustering for different network sizes. Then, we applied asymptotic statistics to assess the significance of the connections [16]. Finally, for each frequency band, we obtained a binary network of dimension 61x61, a binary network of dimension 29x29, and we applied the six algorithms under analysis.

III. Results

A. Simulation study

Results of the ANOVA concerning 60 nodes networks are reported in Table 1. Tests on networks of 30 and 100 nodes led to similar results in terms of significance. The plot-of-means, together with post-hoc tests, returned significantly higher performances in networks with low levels of noise and high levels of cluster number (Fig. 2a and 2b). This is true for every value of number of nodes and density (not reported here for brevity's sake), except for networks made up of 60 nodes and highly dense, where having only 2 clusters seems to guarantee better performances. Furthermore, results show that the algorithms which better detect the simulated communities are,

on average, Leicht & Newman and Louvain (Fig. 2a and 2b). However, for very noisy networks, Ronhovde and Stability Optimization exhibit higher values of NMI compared with the other algorithms, while Infomap and Ronhovde are those less sensitive to the noise (Fig. 2b). Finally, we observed that in denser networks (irrespective of the other parameters) all the algorithms show higher performances. We don't report here the plot of means, but the significance returned by the ANOVA for different densities can be seen in Table I.

B. *Test on EEG data*

Figure 3 shows the division in clusters of the brain network, obtained as described in section II, computed by using all the six algorithms under analysis. In particular, here we present the case of 61 channels, and we consider the network relative to beta band, as of particular interest for motor tasks [17],[18]. Each panel of the figure displays a 2D model of the scalp, with electrodes divided in clusters by a different algorithm. Electrodes are represented with dots and clusters are identified by different colors. The figure shows the algorithm of Leicht & Newman (fig. 3a) and the Louvain one (fig. 3b), which are the best previously found, detecting two principal modules that can be associated with the two

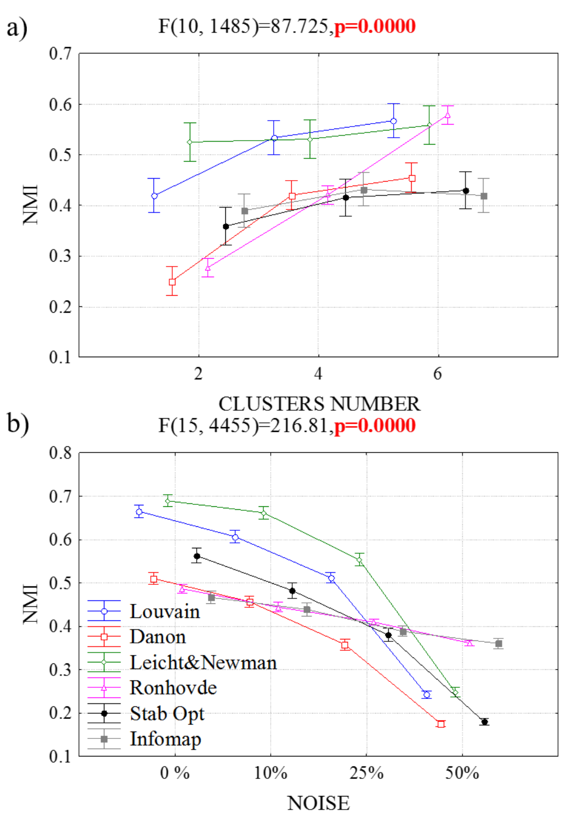

Figure 2. ANOVA performed on networks of 60 nodes and graph density equal to 0.3. a) Results obtained for the two-way interaction between *Clusters Number* (levels: 2, 4, 6) and *Algorithm* (levels: Louvain, Danon, Leicht&Newman, Ronhovde, Stability Optimization, Infomap). b) Results obtained for the two-way interaction between *Noise* (levels: 0%, 10%, 25%, 50%) and *Algorithm*. In both analysis, the dependent variable is the NMI (in ordinate).

TABLE I. RESULTS OF REPEATED MEASURES ANOVA PERFORMED ON 60 NODES NETWORKS. FACTORS: *CLUSTERS NUMBER* (CN), *NOISE* (NO), AND *ALGORITHM* (A), DEPENDENT VARIABLE: NORMALIZED MUTUAL INFORMATION. SIGNIFICANT RESULTS ($p<0.05$) ARE IN BOLD.

60 NODES	**Density 10%**		**Density 30%**		**Density 50%**	
	F	***p***	***F***	***p***	***F***	***p***
CN	**981.84**	**$<10^{-4}$**	**122.06**	**$<10^{-4}$**	0.155	0.693
A	**265.32**	**$<10^{-4}$**	**339.39**	**$<10^{-4}$**	**855.02**	**$<10^{-4}$**
No	**1009**	**$<10^{-4}$**	**2402.1**	**$<10^{-4}$**	**2280.08**	**$<10^{-4}$**
A, CN	**140.05**	**$<10^{-4}$**	**87.725**	**$<10^{-4}$**	**204.39**	**$<10^{-4}$**
No, CN	**107.63**	**$<10^{-4}$**	**40.845**	**$<10^{-4}$**	**17.097**	**$<10^{-4}$**
A, No	**109.29**	**$<10^{-4}$**	**216.81**	**$<10^{-4}$**	**282.91**	**$<10^{-4}$**
A,No,CN	**11.996**	**$<10^{-4}$**	**40.504**	**$<10^{-4}$**	**129.5**	**$<10^{-4}$**

hemispheres (red and green clusters). These two communities are clearly distinguished from the parietal and the frontal areas, which, in the case of Leicht & Newman, are further split in 2 clusters (blue and yellow, and light blue and pink, respectively). Also Stability Optimization (fig. 3c), which is the third in order of performances from fig. 2b, was able to separate the two clusters relative to the hemispheres, and it led to a clustering similar to those just commented above all for what concern the left hemisphere, which is the contralateral one to the movement. On the contrary, with the other algorithms (fig. 3d-f), that had worse performances in our simulation study, we obtained different results, and we couldn't find the same main macro-areas. In particular, the Ronhovde algorithm detected too many communities (22) some of which consisting of a single node, while Infomap and Danon joined the frontal electrodes with the cluster that we previously associated to the left hemisphere. The analysis with 29 channels led to similar results.

IV. DISCUSSION AND CONCLUSION

The aim of this study was to provide a comparative evaluation of the performances of different algorithms for the detection of community structure under the effects of different factors. To this purpose, we used a procedure for the generation of simulated networks with community structure that increases the features that can be controlled, with respect to those at the state of art [1], [10]. In fact, this procedure ensures high variability in clusters size and nodes degree and allows to set parameters such as the number of nodes of the network, the graph density, the number of clusters and the ratio between the density intra-cluster and the density inter-cluster, allowing to systematically control such factors in simulated networks. Performance evaluation has been made through NMI, chosen among the high number of available indices because of its wide use in this field, that allowed us to refer this analysis to other works. Results demonstrated that a decrease of the noise and an increase of the number of clusters induces better performances for all the algorithms in terms of community detection. One can explain these evidences respectively with the fact that noise makes the community structure more undefined, and with the fact that these algorithms have been originally thought to detect communities in big networks including many clusters. The algorithms showing best performances in most of the conditions we examined are Louvain's and Leicht and Newman's, while for noisy data Ronhovde and Infomap are more performing. We also observed that all the algorithms work best on denser networks. These results partially agree with [12], in which Infomap and Louvain are the best

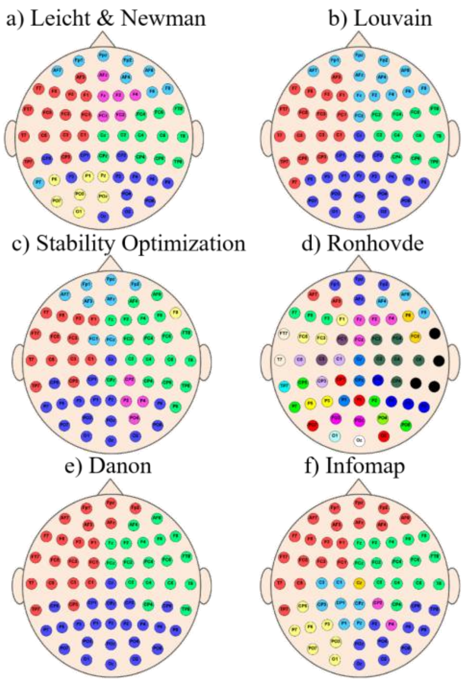

Figure 3. Clustering obtained by applying the six clustering algorithms to the brain connectivity network estimated from EEG signals of an healthy subject during the execution of a motor task. Clustering have been projected on 2D models of scalp (seen from above, triangle represent the nose) in which the dot are the 61 electrodes grouped into clusters, identified by different colors. Each panel illustrates the clustering obtained with an algorhm.

performing, with the first being in general more efficient than the second.

According to these results, we applied the six algorithms on EEG-derived brain networks estimated during a motor task (sustained grasping and extension of the right hand) in beta band. Even though the small sample size does not allow to draw general conclusion, the divisions in clusters obtained are in agreement with the statistical analysis results. In fact, we observed similar results within the two algorithms which better performed in the statistical analysis, while the other algorithms produced different divisions. In particular, the first two, Louvain and Leicht and Newman, identified two main clusters associated with the two hemispheres. The only difference is that Louvain attributed the frontal and parietal areas to two different clusters, while Leicht and Newman divided such two regions in 4 clusters. This clustering is preserved with the Stability Optimization method, especially for what concerns the hemisphere contralateral to the movement, while the application of the other algorithms, less performant, led to different results. Actually, the clustering computed by the Danon algorithm seems to differ from the first two only because it joined frontal areas with the cluster representing the left hemisphere, and this suggests that its low performances resulting from the ANOVA are due to a tendency to join different clusters. Thus, we noticed that despite the fact that the Danon algorithm was defined to find small communities, in this application it only identified 3 clusters, a smaller number with respect to the other algorithms. However, as anticipated, the application of the algorithms described in this study to one subject do not represent a validation of the simulation study on real data, but just a way to prove the usefulness of this technique to brain networks.

Further studies will extend the analysis to include an improvement of the simulated networks generation, by considering spatial correlation in the noise modeling, modeling weighted graphs, and including more conditions in the statistical analysis. Finally, future studies will address the validation of the algorithms in a group of subjects to validate the methods not only in synthetic networks, but also with real EEG-based networks.

REFERENCES

[1] M. Girvan and M. E. J. Newman, “Community structure in social and biological networks,” *Proc. Natl. Acad. Sci. U. S. A.*, vol. 99, no. 12, pp. 7821–7826, Jun. 2002.

[2] S. Fortunato, “Community detection in graphs,” *Phys. Rep.*, vol. 486, no. 3–5, pp. 75–174, Feb. 2010.

[3] M. E. J. Newman, “Modularity and community structure in networks,” *Proc. Natl. Acad. Sci.*, vol. 103, no. 23, pp. 8577–8582, Jun. 2006.

[4] E. A. Leicht and M. E. J. Newman, “Community Structure in Directed Networks,” *Phys. Rev. Lett.*, vol. 100, no. 11, p. 118703, Mar. 2008.

[5] L. Danon, A. Díaz-Guilera, and A. Arenas, “The effect of size heterogeneity on community identification in complex networks,” *J. Stat. Mech. Theory Exp.*, vol. 2006, no. 11, p. P11010, 2006.

[6] P. Ronhovde and Z. Nussinov, “Local resolution-limit-free Potts model for community detection,” *Phys. Rev. E*, vol. 81, no. 4, p. 46114, Apr. 2010.

[7] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, “Fast unfolding of communities in large networks,” *J. Stat. Mech. Theory Exp.*, vol. 2008, no. 10, p. P10008, Oct. 2008.

[8] E. L. Martelot and C. Hankin, “Multi-scale Community Detection using Stability Optimisation within Greedy Algorithms,” *ArXiv12013307 Phys.*, Jan. 2012.

[9] M. Rosvall and C. T. Bergstrom, “Maps of random walks on complex networks reveal community structure,” *Proc. Natl. Acad. Sci.*, vol. 105, no. 4, pp. 1118–1123, Jan. 2008.

[10] A. Lancichinetti, S. Fortunato, and F. Radicchi, “Benchmark graphs for testing community detection algorithms,” *Phys. Rev. E*, vol. 78, no. 4, Oct. 2008.

[11] C. Schmidt, B. Pester, N. Schmid-Hertel, H. Witte, A. Wismüller, and L. Leistritz, “A Multivariate Granger Causality Concept towards Full Brain Functional Connectivity,” *PLOS ONE*, vol. 11, no. 4, p. e0153105, Apr. 2016.

[12] A. Lancichinetti and S. Fortunato, “Community detection algorithms: A comparative analysis,” *Phys. Rev. E*, vol. 80, no. 5, p. 56117, Nov. 2009.

[13] L. Danon, J. Duch, A. Diaz-Guilera, and A. Arenas, “Comparing community structure identification,” *J. Stat. Mech. Theory Exp.*, vol. 2005, no. 9, pp. P09008–P09008, Sep. 2005.

[14] L. A. Baccalá and K. Sameshima, “Partial directed coherence: a new concept in neural structure determination,” *Biol. Cybern.*, vol. 84, no. 6, pp. 463–474, Jun. 2001.

[15] W. Klimesch, “EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis,” *Brain Res. Rev.*, vol. 29, no. 2–3, pp. 169–195, Apr. 1999.

[16] D. Y. Takahashi, L. A. Baccal, and K. Sameshima, “Connectivity Inference between Neural Structures via Partial Directed Coherence,” *J. Appl. Stat.*, vol. 34, no. 10, pp. 1259–1273, Dec. 2007.

[17] S. N. Baker, J. M. Kilner, E. M. Pinches, and R. N. Lemon, “The role of synchrony and oscillations in the motor output,” *Exp. Brain Res.*, vol. 128, no. 1–2, pp. 109–117, Sep. 1999.

[18] J. N. Sanes and J. P. Donoghue, “Oscillations in local field potentials of the primate motor cortex during voluntary movement.,” *Proc. Natl. Acad. Sci.*, vol. 90, no. 10, pp. 4470–4474, May 1993.