

Identification of active modules in interaction networks using node2vec network embedding

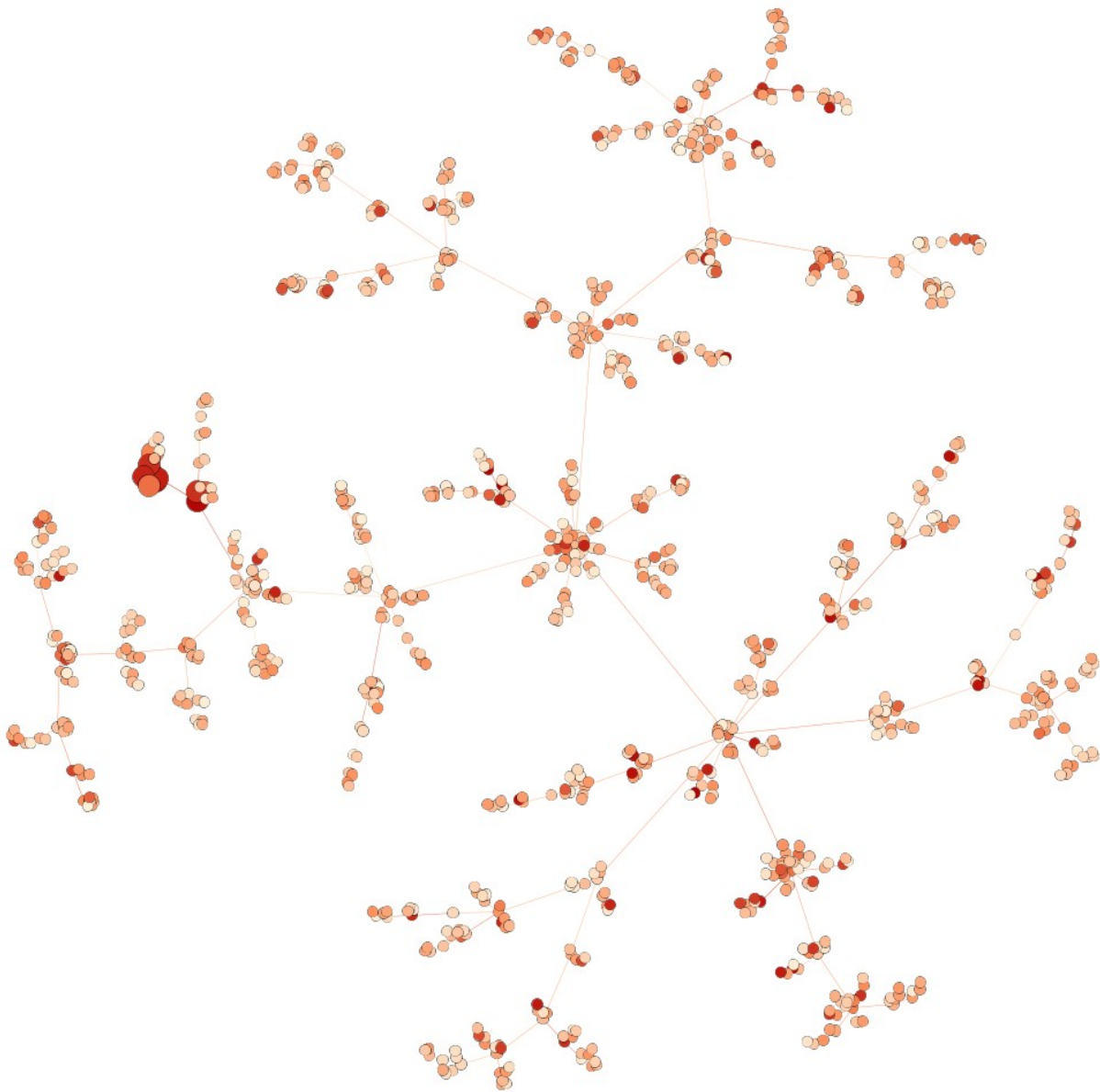
Claude Pasquier^{1*}, Vincent Guerlais¹, Denis Pallez¹, Raphaël Rapetti-Mauss², Olivier Soriani²

¹ Université Côte d'Azur, CNRS, I3S, France.

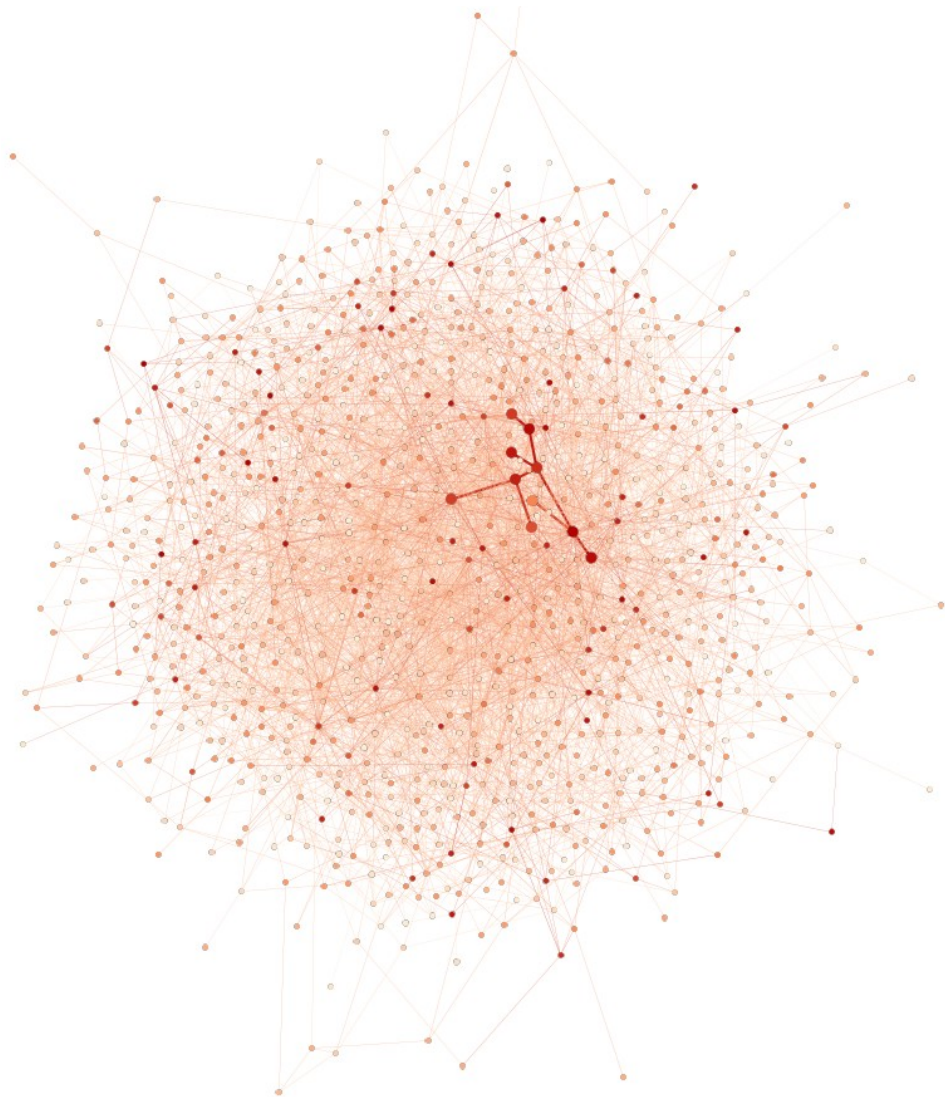
² Université Côte d'Azur, CNRS, INSERM, iBV, France.

* Corresponding author (claude.pasquier@univ-cotedazur.fr).

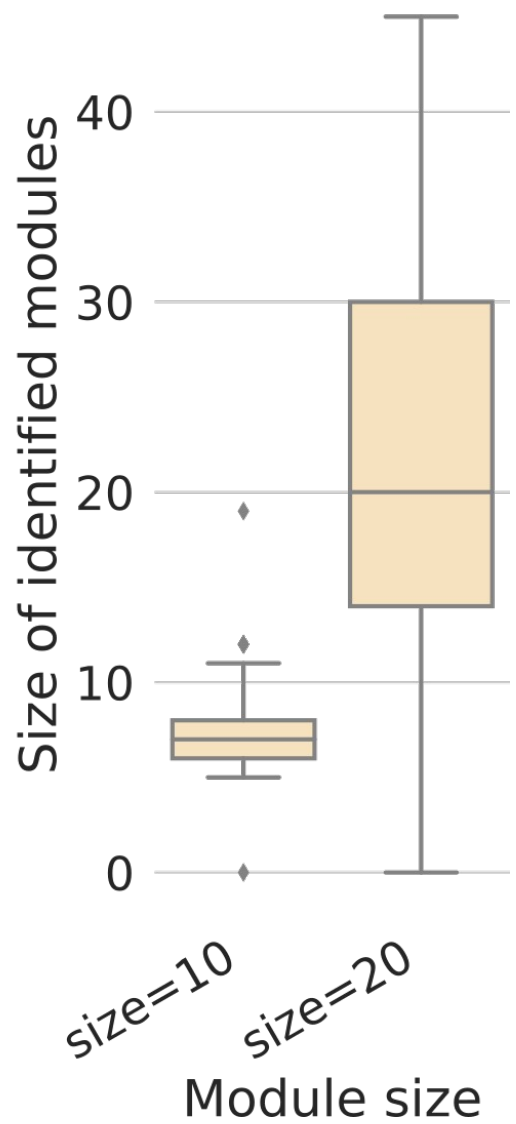
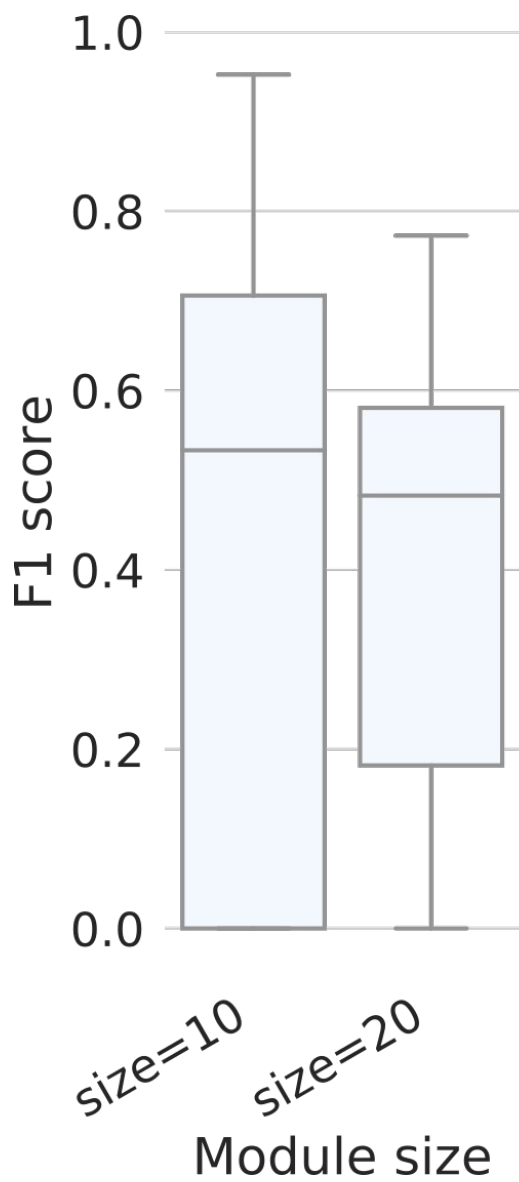
Supplementary materials



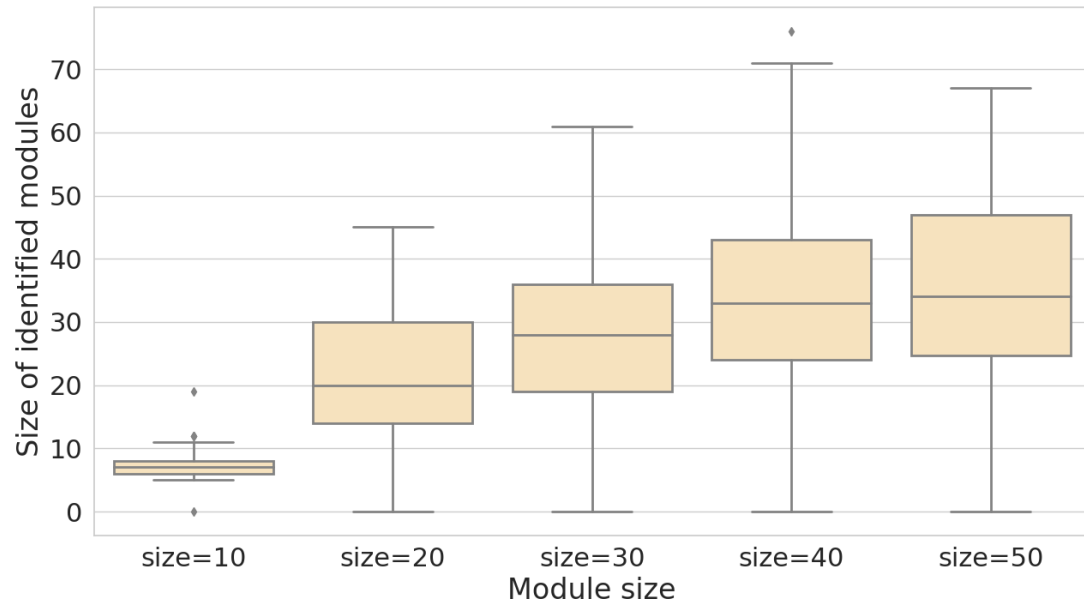
Supplementary Fig. S1. Example of a scale free graph of 1000 vertices generated with the Barabasi-Albert model of preferential attachment ([Albert and Barabási, 2000](#)). Node colors represents the values associated with nodes. Larger nodes constitutes the module to discover.



Supplementary Fig. S2. Example of a scale free graph of 1000 vertices generated with the extended version of Barabasi-Albert model of preferential attachment ([Albert and Barabási, 2000](#)). Node colors represents the values associated with nodes. Larger nodes, connected by wider edges, constitute the module to discover.



Supplementary Fig. S3. Boxplots representing the distribution of F1 scores and identified module size obtained by our method on dense artificial networks of 10000 vertices with modules of size 10 and 20 to discover.



Supplementary Fig. S4. Boxplots representing the distribution of the size of the modules identified by our method on dense artificial networks of 10000 vertices for several theoretical module sizes.

The screenshot shows a web browser window with the URL `amine.i3s.unice.fr`. The page header includes logos for I3S and CNRS, and the text "AMINE". Navigation links for "Home", "About", and "Contact" are visible. The main content area is titled "Active module detection with AMINE algorithm" and provides a brief description of the method. Below the description is a "Module detection" form with the following fields:

- Job title:** A text input field.
- Differential expression data:** A file upload button labeled "Parcourir..." with the text "Aucun fichier sélectionné."
- Column with gene names:** A dropdown menu with the value "1" selected.
- Column with p-values:** A dropdown menu with the value "2" selected.
- Column with log2 fold change values:** A dropdown menu with the value "0" selected.
- Specie:** A dropdown menu with "Homo sapiens" selected.
- Favorite genes (optional):** A file upload button labeled "Parcourir..." with the text "Aucun fichier sélectionné."
- E-mail (optional):** A text input field.

A blue "Submit" button is located at the bottom of the form. The footer of the page contains the text "I3S 2020".

Supplementary Fig. S5. Input form of the AMINE method. P-values, representing the evidence of change in gene expression across experimental conditions, are specified in a file uploaded via the "Differential expression data" input field. At a minimum, the data consists of a tabular file with two columns: one containing the name of the genes and the other, the associated p-value. Another way to indicate the data to be used is to simply specify the result file produced by differential expression analysis methods such as DeSeq2 or edgeR. The input file can be in csv (comma-separated values) or xlsx format. The following two input fields are used to specify the columns containing the required data. The index of the column that contains the gene names must be specified in "Column with gene names" while that of the column with p-values must be specified in "Column with p-values". The index of the column with the log₂ fold change can optionally be specified in the next field. Then, a drop-down list allows to choose the organism. It is possible to focus the search on modules containing specific genes. A file containing the list of these "favorite" genes (one gene name per line) can be specified with the "Favorite genes" input field. The e-mail address is used to notify the user of the completion of the task and to provide the address where the results are available.