

La centralité dans les réseaux de collaboration étudiant ; une comparaison des similitudes avec les réseaux écologiques

Philippe Bourassa, Simon Bourgeois, Justine Labelle, Kayla Trempe-Kay

^a Université de Sherbrooke Département de biologie 2500 Boulevard de l'Université Sherbrooke Québec J1K 2R1

Abstract

This is the abstract.

It consists of two paragraphs.

Keywords: Réseau écologique, Collaboration, Centralité,

Introduction

Les écologistes ont comme pratique l'étude des interactions entre les espèces qui peut être traduite et analysée à l'aide de réseaux écologiques. Les différentes propriétés de ces réseaux peuvent fournir des indications sur l'organisation des processus écologiques, mais aussi sur la résilience du réseau face aux changements et aux perturbations (Delmas et al., 2019). L'étude des réseaux écologiques peut se faire autant à l'échelle des écosystèmes qu'au niveau des espèces. Dans le dernier cas, ils permettent d'observer le rôle et l'importance de celles-ci à travers leurs interactions au sein du réseau (Delmas et al., 2019). Plusieurs facteurs peuvent influencer ces interactions tels que l'abondance locale des espèces, leurs traits ainsi que les conditions environnementales (Delmas et al., 2019). En tant qu'écologistes si bien formés par l'Université de Sherbrooke, nous avons tenté de déterminer si les propriétés d'un réseau écologique sont similaires à ceux d'un réseau d'interactions entre les étudiants du cours BIO500 à l'hiver 2023. Au cours de leur parcours universitaire, les étudiants ont été confrontés à réaliser de nombreux travaux d'équipes qui représentent des collaborations avec différents individus. Ces collaborations permettent d'analyser les interactions entre les étudiants et les facteurs qui peuvent les influencer. L'objectif est donc d'observer l'importance des étudiants au sein du réseau de collaboration à l'aide de la centralité, mais aussi de comprendre comment l'année du début du baccalauréat et la formation préalable influencent cette centralité.

Méthode

Échantillonnage

Durant l'hiver 2023, les étudiants du cours de BIO500 ont été divisés en 10 équipes de 4 personnes. Chaque équipe a complété trois fichiers csv concernant les cours auxquels ils ont réalisé des travaux d'équipes au courant de leur baccalauréat. Le premier fichier permet d'identifier les collaborations effectuées, c'est-à-dire les liens d'interactions avec leurs différents coéquipiers en fonction du cours. Le deuxième fichier représente des informations relatives aux étudiants de l'équipe. Plus spécifiquement, leur région administrative, leur participation au régime coopératif, leur année de début du baccalauréat, leur formation préalable et leur numéro du programme. Puis, le dernier identifie les cours dans lesquels des travaux ont été réalisés et d'autres données pertinentes.

*Corresponding author

Email addresses: Philippe.Bourassa4@USherbrooke.ca (Philippe Bourassa), Simon.Bourgeois@USherbrooke.ca (Simon Bourgeois), Justine.Labelle@USherbrooke.ca (Justine Labelle), Kayla.Trempe-Kay@USherbrooke.ca (Kayla Trempe-Kay)

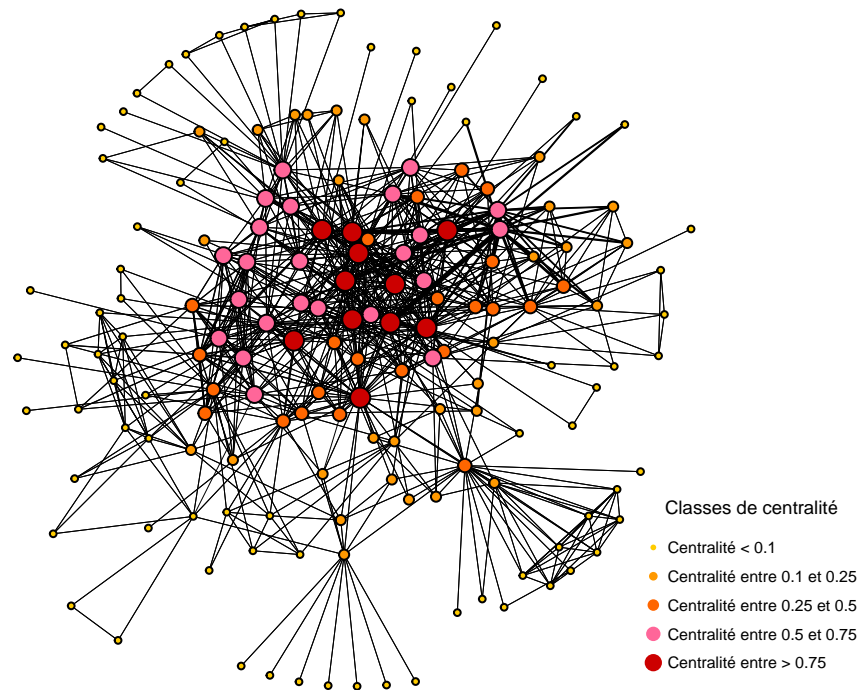


Figure 1: Distribution de la centralité selon la session de début de programme des individus du réseau. La taille des nœuds et leur couleur varient selon la centralité des étudiants tel qu'expliqué par la légende. L'épaisseur des connexions entre les nœuds reflètent l'occurrence d'interactions entre deux étudiants. Plus les lignes sont épaisses, plus les étudiants ont collaboré un grand nombre de fois.

Traitement et analyse des données

Une fois la compilation des données réalisée par chaque équipe, les données ont été partagées à l'ensemble du groupe et celles-ci ont été fusionnées à l'aide du logiciel R sous forme de trois fichiers csv. Par la suite, les données des différents fichiers ont été nettoyées et standardisées afin d'obtenir une conformité au sein des différentes tables. Ces données ont ensuite été injectées à l'aide du logiciel R dans le système de gestion de données SQLite3 (Müller et al., 2023). Ce système nous a permis d'extraire les données d'intérêt via des requêtes qui sont nécessaires à la création des figures. Il est à noter que ce n'est pas toutes les données qui ont été utilisées lors des analyses. Les représentations visuelles du réseau d'interaction, ainsi que la centralité ont été effectuées grâce au package "Igraph", ainsi qu'à l'aide du package "Scales" du logiciel R (Csárdi et al., 2023; Wickham et al., 2022). Le package "Vioplot" du logiciel R a aussi été utilisé pour la création du diagramme en violon (Adler et al., 2022). Ensuite, le package "targets" a été utilisé afin d'automatiser l'ensemble du processus et d'augmenter la reproductibilité de cette étude (Landau et al., 2023a,b). Finalement, le gabarit PNAS dans Rmarkdown a été utilisé grâce au package "Rticles" (Allaire et al., 2022).

Résultats

La Figure 1 représente le réseau de collaborations des individus du cours BIO500 donné à l'hiver 2023 et leurs collègues. On y observe une variation au niveau de l'importance des étudiants grâce au score

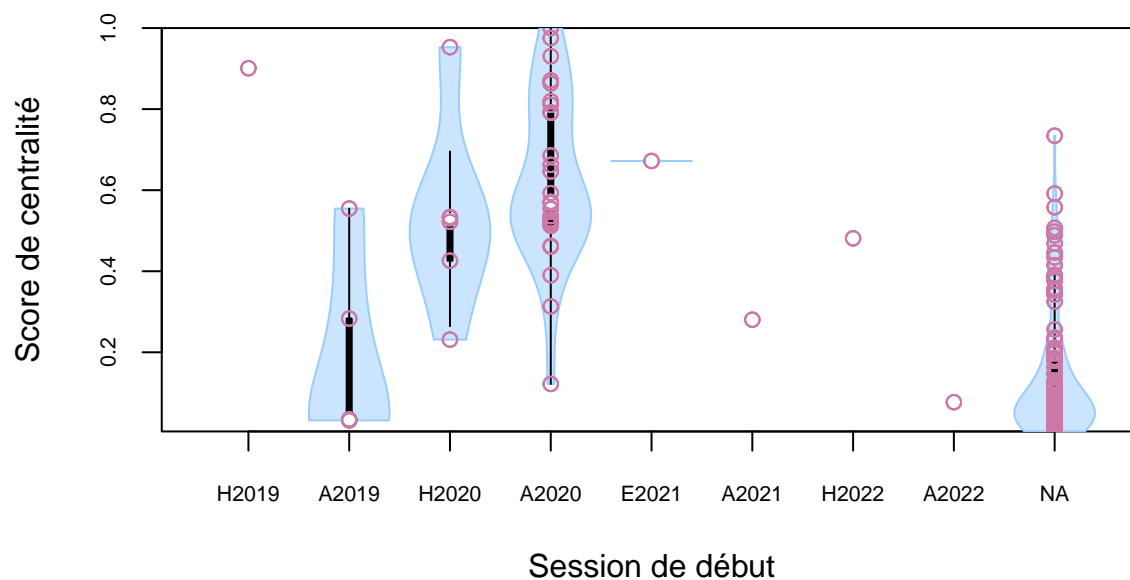


Figure 2: Distribution de la centralité selon la session de début de programme des individus du réseau. Les points mauves dans le graphique correspondent au score de centralité de chaque individu. Le point blanc représente la médiane du groupe et les boîtes noires correspondent à l'intervalle entre les quartiles. La ligne noir plus mince montre l'étendu où 95% des données se situent. Les côtés montrent la distribution des données.

de centralité. En effet, la taille des nœuds, qui représente chaque étudiant, varie de manière à ce qu'elle augmente plus le score de centralité est haut, soit proche de 1. Les individus de faible centralité sont retrouvés en marge du réseau alors que les individus présentant une forte centralité sont retrouvés plus au centre. On y décèle également plusieurs occurrences d'interactions entre des étudiants. Effectivement, des lignes plus épaisses sont dénotées dans le réseau.

Discussion

Salut
#Conclusion

Remerciement

Nous aimerions prendre le temps de remercier Hugo Morin Brassard et Félix-Olivier Dufour qui nous ont grandement aidé lors de nos nombreuses difficultés rencontrées principalement celles dans notre code R.

References

Daniel Adler, S. Thomas Kelly, Tom M. Elliott, and Jordan Adamson. vioplot: Violin Plot, December 2022. URL <https://CRAN.R-project.org/package=vioplot>.

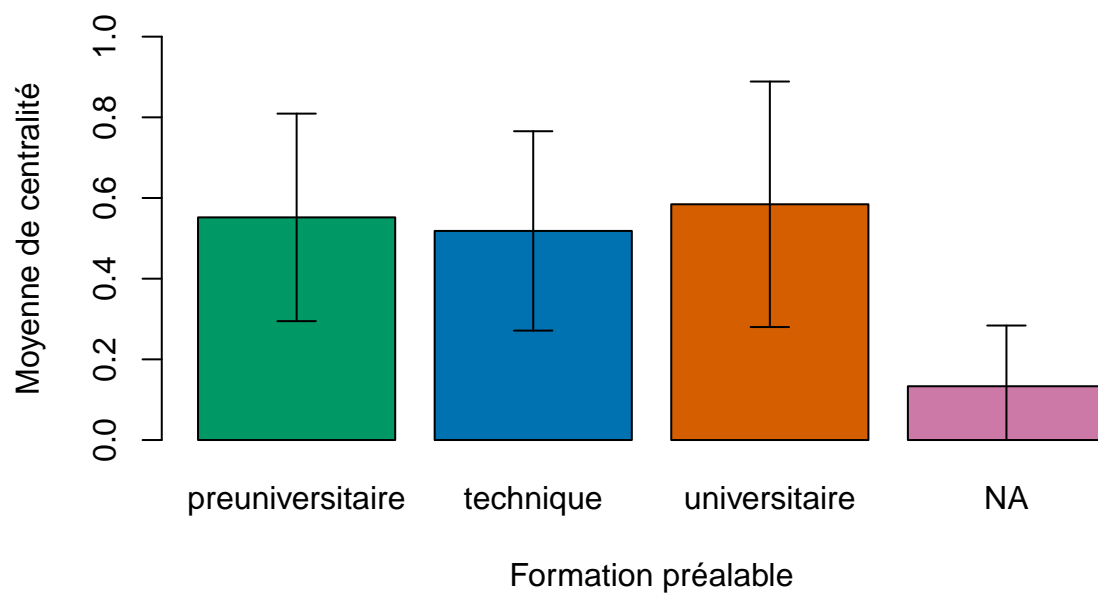


Figure 3: Moyenne de centralité selon la formation préalable des individus du réseau. Les barres d'erreurs correspondent aux écart-types associés à la moyenne de chaque classe soit : préuniversitaire, technique, universitaire et NA (individus hors cours BIO500 hiver 2023).

J. J. Allaire, Yihui Xie, Christophe Dervieux, RStudio, R. Foundation, Hadley Wickham, Journal of Statistical Software, Ramnath Vaidyanathan, Association for Computing Machinery, Carl Boettiger, Elsevier, Karl Broman, Kirill Mueller, Bastiaan Quast, Randall Pruim, Ben Marwick, Charlotte Wickham, Oliver Keyes, Miao Yu, Daniel Emaasit, Thierry Onkelinx, Alessandro Gasparini, Marc-Andre Desautels, Dominik Leutnant, MDPI, Taylor Francis, , Oğuzhan Ögreden, Dalton Hance, Daniel Nüst [aut, cph, Petter Uvesten, Elio Campitelli, John Muschelli, Alex Hayes, Zhian N. Kamvar, Noam Ross [aut, cph, Robrecht Cannoodt (rcannoodt), Duncan Luguern, David M. Kaplan (dmkaplan2000), Sebastian Kreutzer, Shixiang Wang, Jay Hesselberth, Alfredo Hernández, Stefano Coretta (stefanocoretta), Greg Macfarlane (gregmacfarlane), Matthias Templ (matthias da), Alvaro Uzaheta (auzaheta), JooYoung Seo, Callum Arnold (arnold c), Rob Hyndman, Dmytro Perepolkin (dmi3kno), and Tom Palmer (remlapmot). *rticles: Article Formats for R Markdown*, August 2022. URL <https://CRAN.R-project.org/package=rticles>.

Gábor Csárdi, Tamás Nepusz (See also AUTHORS file.), Vincent Traag, Szabolcs Horvát, Fabio Zanini, Daniel Noom, Kirill Müller, Maëlle Salmon, and Chan Zuckerberg Initiative *igraph* author details. *igraph: Network Analysis and Visualization*, April 2023. URL <https://CRAN.R-project.org/package=igraph>.

Eva Delmas, Mathilde Besson, Marie-Hélène Brice, Laura A Burkle, Giulio V Dalla Riva, Marie-Josée Fortin, Dominique Gravel, Paulo R Guimarães Jr, David H Hembry, Erica A Newman, et al. *Analysing ecological networks of species interactions. Biological Reviews*, 94(1):16–36, 2019.

William Michael Landau, Samantha Oliver, Tristan Mahr, and Eli Lilly and Company. *tarchetypes: Archetypes for Targets*, March 2023a. URL <https://CRAN.R-project.org/package=tarchetypes>.

William Michael Landau, Matthew T. Warkentin, Mark Edmondson, Samantha Oliver, Tristan Mahr, and Eli Lilly and Company. *targets: Dynamic Function-Oriented 'Make'-Like Declarative Pipelines*, March 2023b. URL <https://CRAN.R-project.org/package=targets>.

Kirill Müller, Hadley Wickham, David A. James, Seth Falcon, D. Richard Hipp (for the included SQLite sources), Dan Kennedy (for the included SQLite sources), Joe Mistachkin (for the included SQLite sources), SQLite Authors (for the included SQLite sources), Liam Healy (for the included SQLite sources), R. Consortium, and RStudio. *RSQLite: SQLite Interface for R*, April 2023. URL <https://CRAN.R-project.org/package=RSQLite>.

Hadley Wickham, Dana Seidel, and RStudio. *scales: Scale Functions for Visualization*, August 2022. URL <https://CRAN.R-project.org/package=scales>.