

LIMOS – axe MAAD, thème AGC

Directeur de thèse : Florent Foucaud (MCF), florent.foucaud@uca.fr

Titre du sujet de thèse : Mesures de complexité et identification dans les graphes: structure et algorithmes

Résumé du sujet de thèse : L'objectif de cette thèse est d'étudier les mesures de complexité des données pour les structures discrètes (avec un accent sur les graphes) et les problèmes associés, principalement d'un point de vue théorique, selon les points de vue structurel et algorithmique.

Mesurer la complexité des données est crucial dans les applications de traitement de données, d'autant plus que les jeux de données croissent exponentiellement en taille, usage et diversité. Parmi les mesures les plus utilisées figurent la dimension de Vapnik-Chervonenkis (VC-dimension) [1], la dimension d'enseignement [4] et la complexité de voisinage [2]. Un domaine connexe est celui des problèmes d'identification, où l'on dispose d'une structure combinatoire, telle qu'un graphe, hypergraphe ou configuration géométrique. On souhaite sélectionner un (petit) ensemble d'éléments de la structure, de sorte que tous les autres éléments soient uniquement déterminés par leurs relations avec cet ensemble solution. Le but de cette thèse est d'étudier le domaine des mesures de complexité de graphes et son lien avec les problèmes d'identification.

Du côté structurel/combinatoire, plusieurs conjectures et problèmes ouverts méritent d'être approfondis. L'une de ces conjectures est celle de Simon-Zilles [4], portant sur la relation optimale entre la dimension d'enseignement récursive et la VC-dimension, liée à la célèbre conjecture de compression d'échantillons en apprentissage automatique théorique [6]. Une conjecture connexe est celle de localisation-domination $n/2$ [7], un problème ouvert majeur en identification sur les graphes. Ces problèmes ouverts sont liés et nous espérons que des progrès pourront être accomplis en les étudiant sous un angle commun. Nous souhaitons également étudier la complexité de voisinage de classes de graphes importantes, telles que les graphes planaires et les graphes de largeur arborescente bornée, pour lesquels les valeurs optimales restent inconnues [2].

Ces aspects structurels peuvent être complétés par des aspects algorithmiques. En effet, bien que plusieurs articles s'intéressent aux aspects algorithmiques des mesures de complexité des données [1] et des problèmes d'identification [3], de nombreuses questions intéressantes restent ouvertes, comme la complexité algorithmique de la reconnaissance des graphes à n sommets admettant un code identifiant de taille exactement $\log_2(n)$ (ou, de façon équivalente, de VC-dimension exactement $\log_2(n)$), ou la complexité paramétrée des mesures de complexité des données par rapport à des paramètres structurels. Une autre direction de recherche intéressante consiste à explorer les applications algorithmiques de la complexité de voisinage, sujet dont l'investigation a débuté récemment [5] mais qui reste à approfondir.

Références

- [1] C. Bazgan, F. Foucaud, and F. Sikora. Parameterized and approximation complexity of Partial VC Dimension. *Theoretical Computer Science*, 766 :1–15, 2019.
- [2] L. Beaudou, J. Bok, F. Foucaud, D. A. Quiroz, and J. Raymond. Profile and neighbourhood complexity of graphs with excluded minors and tree-structured graphs. *ArXiv preprint 2501.08895*, 2025.
- [3] D. Chakraborty, F. Foucaud, D. Majumdar, and P. Tale. Tight (double) exponential bounds for identification problems: Locating-dominating set and test cover. *Proceedings of ISAAC 2024*.
- [4] T. Doliwa, G. Fan, H. U. Simon, and S. Zilles. Recursive teaching dimension, VC-dimension and sample compression. *Journal of Machine Learning Research*, 15(1):3107–3131, 2014.
- [5] J. Dreier, I. Eleftheriadis, N. Mählmann, R. McCarty, M. Pilipczuk, and S. Torunczyk. First-order model checking on monadically stable graph classes. *Proceedings of FOCS 2024*.
- [6] S. Floyd and M. K. Warmuth. Sample compression, learnability, and the vapnik-chervonenkis dimension. *Machine Learning*, 21(3):269–304, 1995.

[7] D. Garijo, A. González, and A. Márquez. The difference between the metric dimension and the determining number of a graph. Applied Mathematics and Computation, 249:487–501, 2014.