

**Titre :** Propriétés garanties par la réduction de graphes et impact sur l'anonymisation de graphes

Ce stage sera effectué dans le cadre du projet ANR COREGRAPHIE est sera co-encadré par plusieurs partenaires du projet : Michel HABIB (IRIF, Paris), Sergey KIRGIZOV (LIB, Dijon), Hamida SEBA (LIRIS, Villeurbanne) et Olivier TOGNI (LIB, Dijon)

Le stagiaire recevra une gratification standard et pourra choisir le lieu du stage entre Dijon, Lyon et Paris.

Contact :

- Michel Habib [habib@irif.fr](mailto:habib@irif.fr)
- Sergey Kirgizov [Sergey.Kirgizov@u-bourgogne.fr](mailto:Sergey.Kirgizov@u-bourgogne.fr)
- Hamida Seba [hamida.seba@univ-lyon1.fr](mailto:hamida.seba@univ-lyon1.fr)
- Olivier Togni [olivier.togni@u-bourgogne.fr](mailto:olivier.togni@u-bourgogne.fr)

### **Description :**

La réduction de graphes permet d'obtenir des représentations plus simples ou plus petites de grands graphes. Il existe deux principales approches à cette réduction [3,4]: l'agrégation (coarsening or summarization) qui consiste à regrouper les nœuds et les arêtes qui partagent certaines propriétés pour capturer la structure globale du graphe au détriment de détails locaux, et la dé-densification (sparsification) qui consiste à supprimer certaines arêtes voire certains nœuds du graphe de manière à préserver une ou plusieurs propriétés du graphe initial dans le résumé obtenu. Dans ce stage, nous nous intéressons à ce qu'un résumé peut révéler sur le graphe initial en se focalisant sur les propriétés que ces résumés préservent [5]. En effet, plusieurs méthodes de réduction de graphes sont proposées dans la littérature [3,6,7] et cette approche est utilisée dans l'anonymisation de graphes. Motivées par la préservation de la confidentialité dans les réseaux sociaux, plusieurs méthodes proposent d'anonymiser le graphe d'interactions sous-jacent en modifiant sa structure par réduction [8]. L'idée ici est de cacher les informations topologiques d'un graphe qui peuvent être utilisées pour déduire l'identité des nœuds par exemple. Le but du stage est d'étudier les approches de réduction de graphes pour l'anonymisation de réseaux. L'étude des propriétés préservées par la réduction de graphes nous permet de savoir dans quelle mesure ces approches sont viables [9]. La méthodologie suivie est de considérer une ou plusieurs approches de réduction de graphes, une propriété  $P$  (degré moyen, distance, etc.) et de voir dans quelle mesure  $P$  est préservée sur le résumé. On dira qu'une méthode de compression est  $(k,r,P)$ -préservante si quel que soit le graphe, la propriété  $P$  est préservée à un ratio  $k$  (par rapport au graphe initial) sur le graphe compressé au ratio  $r$ . On s'intéressera aussi aux classes de graphes pour lesquelles de telles propriétés peuvent être vérifiées.

- [1] S. Lagraa, H. Seba, R. Khennoufa, A. MBaya, H. Kheddouci, A distance measure for large graphs based on prime graphs, *Pattern Recognition* 750 47 (9) (2014) 2993–3005. doi: <https://doi.org/10.1016/j.patcog.2014.03.014>.
- [2] Geia Bravo-Hermsdorff and Lee M. Gunderson. 2019. A unifying framework for spectrum-preserving graph sparsification and coarsening. *Proceedings of the 33rd International Conference on Neural Information Processing Systems*. Curran Associates Inc., Red Hook, NY, USA, Article 695, 7736–7747.
- [3] Abd Errahmane Kiouche, Julien Baste, Mohammed Haddad & Hamida Seba (2021). « A Neighborhood-preserving Graph Summarization ». *ArXiv* : 2101.11559. HAL : hal-03123537.
- [4] D. A. Spielman, N. Srivastava, Graph sparsification by effective resistances, *SIAM Journal on Computing* 40 (6) (2011) 1913–1926. doi: 760 10.1137/080734029.

- [5] Andreas Loukas, Pierre Vandergheynst. Spectrally Approximating Large Graphs with Smaller Graphs. Proceedings of the 35th International Conference on Machine Learning, Stockholm, Sweden, PMLR 80, 2018.
- [6] S. Navlakha, R. Rastogi, N. Shrivastava, Graph summarization with bounded error, in: Proceedings of the 2008 ACM SIGMOD International 845 Conference on Management of Data, SIGMOD '08, Association for Computing Machinery, New York, NY, USA, 2008, p. 419432. doi:10.1145/1376616.1376661.
- [7] M. Riondato, D. García-Soriano, F. Bonchi, Graph summarization with quality guarantees, Data Mining and Knowledge Discovery 31 (2) (2017) 314–349. doi:10.1007/s10618-016-0468-8.
- [8] François Rousseau, Jordi Casas-Roma, and Michalis Vazirgiannis. 2018. Community-preserving anonymization of graphs. Knowl. Inf. Syst. 54, 2 (February 2018), 315–343. <https://doi.org/10.1007/s10115-017-1064-y>
- [9] Lars Backstrom, Cynthia Dwork, and Jon Kleinberg. 2011. Wherefore art thou R3579X? anonymized social networks, hidden patterns, and structural steganography. Commun. ACM 54, 12 (December 2011), 133–141. <https://doi.org/10.1145/2043174.2043199>