

data viz. de gros graphes

sujet d'alternance 2019-2020

Encadrement :

- François Brucker (francois.brucker@centrale-marseille.fr), LIS & ecm
- Pascal Pr  a (pascal.prea@centrale-marseille.fr), ecm & LIS

Laboratoire : Laboratoire d'Informatique & Syst  mes. Equipe ACRO (<https://www.lis-lab.fr/acro/>)

contexte

Les graphes sont une structure informatique (et de math  matiques discr  tes) d'une grande richesse, tant structurelle qu'algorithmique. Pour ne rien g  cher, ils permettent de belles et fortes applications pratiques comme la d  tection de communaut  s (*graphe facebook*), le calcul d'itin  raire (*google maps*) ou encore des probl  mes d'affectations (*parcoursup*).

Cependant,    l'air du big data, ces graphes sont *gros* (plusieurs millions de sommets). Il faut donc des algorithmes efficaces pour les traiter et en synth  tiser l'information..

Enfin pour repr  senter ces donn  es et les r  sultats obtenus de fa  on lisible, intelligible et interpr  table, il est n  cessaire de mettre en   uvres des techniques   labor  es et visuellement   l  gantes.

sujet d'alternance

Le sujet consiste      tudier, proposer et impl  menter une m  thode innovante de r  sum   et de visualisation de gros graphes.

Les techniques de r  sum   ou de d  cision actuelles utilisent essentiellement des partitions ou des hi  rarchies. Si ces structures permettent un traitement rapide, elles ne permettent pas de repr  senter, par exemple dans le *graphe facebook*, les interactions entre groupes et la multiplicit   des appartenances.

L'originalit   de ce travail consiste    utiliser des structures discr  tes (les *totally balanced hypergraphs*) qui autorisent le chevauchement des diff  rentes classes, et permettent par l   de rendre compte de l'appartenance d'un individu    plusieurs communaut  s.

Pour cela, on pourra s'inspirer de la *m  thode de Louvain* (qui partitionne des gros graphes) et de techniques issues de l'apprentissage automatique (pour conserver le maximum d'information). Enfin, on s'inspirera d'algorithmes de repr  sentation des arbres phylog  n  tique pour la visualisation.