Titre : Introduction aux arbres de plus proches voisins et quelques résultats nouveaux.

Orateur : Jérôme Casse

Résumé : Prenons un espace E muni d'une métrique δ et d'une mesure de probabilité P. À partir de P, on construit une suite de points aléatoires i.i.d. de loi P. Au temps n, le $n$ième point de cette suite s'accroche au point qui lui est le plus proche pour la métrique δ parmi les $n-1$ premiers. En cas d'égalité, on prend la loi uniforme sur l'ensemble des points les plus proches. Si on regarde les liens entre les points, on obtient au final un arbre appelé arbre de plus proche voisin sur l'espace métrique probabilisé (E, δ, P).

Une question naturelle est de se demander comment l'espace (E, δ, P) influence les propriétés de l'arbre de plus proche voisin.

Dans cet exposé, nous considérons le cas le plus commun qui est quand E est la sphère de dimension d, muni de la distance euclidienne et de la mesure de Lebesgue/Haar. Nous verrons que les arbres de plus proches voisins sur ces espaces possèdent quelques propriétés communes. Une des questions est alors de trouver et de prouver des propriétés qui dépendent de la dimension d.

Dans un premier temps, nous présentons des résultats nouveaux pour le cas $d = 1$ que nous comparons à ceux (bien connus) du Random Recursive Tree qui correspond à $d = \infty$.

Dans un second temps, nous calculons le nombre moyen de frères/soeurs d'un noeud quelconque de l'arbre de plus proche voisin en dimension d et montrons que ce nombre dépend de la dimension.

Quelques questions ouvertes :

- Comment généraliser à la dimension d quelconque, certains résultats obtenus uniquement en dim 1 ?
- La statistique du nombre moyen de frères/soeurs est-elle optimale pour distinguer la dimension d ?
- Peut-on distinguer un arbre de plus proche voisin construit dans l'hypercube de dim d de celui construit sur la sphère de dimension d ?