

17/04

- Time Varying Graphs (TVG): ver formalismos en el paper de Ziviani Wehmuth para ver como es la dinámica de los grafos y cómo cambia la detección de las características de los usuarios a partir del uso de los clusters de antenas.
- Un tema también podría ser hacer un 'survey' en machine-learning relacionada con age/gender prediction del grafo. Ver porqué no funcionan o cómo mejorar estas técnicas usuales. Y ver si de alguna manera funcionan mejor con un algoritmo más anillado.
- Ver qué propiedades tienen el conjunto de seeds y ver si mejora la predicción 'homogeneizandolo' i.e. (tomando subcortes más aleatorios).
- Ver qué pasa en un TVG si aplico el algoritmo de homofilia a un momento con un resultado y después cómo cambia la predicción si agarramos otra ventana de tiempo.
- Análisis de movilidad.
- Interfaz de datos: Usuario + pwd / conexión por ssh / cctave, python , R o scikit learn.
- Tips: citar al libro que cita wikipedia y para buscar papers buscar el deseado en google scholar y al lado del 'cited by' debería aparecer el 'all versions' como para bajar gratis el pdf. Otra alternativa es buscar directamente en arXiv.

07/05

- Fuente vs. densidad: Probaron el algoritmo de las dos maneras y no habrían habido diferencias, tal vez solamente una convergencia más rápida.
- Los beneficios de anillar el algoritmo ya están contemplados implícitamente en la propia ecuación de difusión.
- Se había probado el algoritmo con pesos $w_{i,j}$ distintos a uno pero resulto más efectivo y rápido considerarlo sin pesos.
- La ecuación general es:

$$g_{x,t} = (1-\lambda)g_{x,0} + \frac{\sum_{x \sim y} w_{x,y} * P * g_{x,t-1}}{\sum_{x \sim y} w_{x,y}}$$

- Se puede considerar softmax como parámetro 'continuo' que en cada paso tome la decisión de, o dejar $g_{x,t}$ como está o, en el otro extremo, tomar la moda de el conjunto de la probabilidad de pertenencia de $g_{x,t}$ a cada una de las clases.
- Estudiar que pasa si va cambiando el conjunto de seeds. Por ejemplo si lo 'homogeinizo' o hago más esparso al restringir el resto del grafo, etc.
- Sci-kit learn: empezar a tocarlo, usar algoritmos usuales en datasets de juguete. Regresiones, redes, Svm, y algoritmos semi-supervisados. Mirar dentro de la documentación.
- Mirar el libro de Mackay.

- Armar con Latex un technical report (buscar templates). La idea de esto sería ir armando un draft de la tesis donde se va ordenando toda la información de lo que se lee/busca y se van contando las herramientas que vamos a usar, comentando acerca de qué son, cómo funcionan cómo se relacionan con el problema, etc.
- Armar un repositorio en github.
- Está el problema: qué técnicas existen, que hace cada una, ventajas y desventajas ...
- Mirar otros papers de la bibliografía.
- Leer un paper y meterlo en el *technical report*. La idea es pensar hacer algo que podría entender yo mismo en 20