

## **CAIM LAB Session 5: Pagerank**

**Por:** Ferran Noguera Vall y Yeray Bosoms Blasco

### **Dificultades y elecciones al implementar el código:**

Para esta sesión hemos empezado por comprender el pseudocódigo que se nos proporcionaba para calcular el pagerank. Una vez descubierto que necesitábamos exactamente para poder calcularlo investigamos qué información nos aportan los ficheros de entrada de airports.txt y route.txt, con ello descubrimos que algunos aeropuertos mencionados en route.txt que usaban IATA code no existían en airports.txt.

En cuanto supimos que teníamos de entrada y lo que necesitábamos para el pagerank pensamos en que estructuras de datos usar. Acabamos usando diccionarios para representar el grafo, debido a la facilidad de acceso, y una lista de tuplas para el resultado final.

Para solucionar el problema de pérdida de pagerank que ocurría al haber aeropuertos sin vuelos que se originen desde ellos (sink nodes) investigamos posibles soluciones, finalmente nos acabamos decantando por extender la fórmula que se nos daba en el pseudocódigo.

### **Observaciones y comentarios de los resultados:**

El vector de pageranks que resulta tras aplicar nuestro código no suma exactamente 1, tras varias pruebas y preguntar al profesor, vimos que era debido al error producido por la falta de decimales al trabajar con tantísimos números tan pequeños.

Probando con diferentes damping factors pudimos ver que cuanto menor es el damping factor más igualados quedan los pageranks en general, ya que la parte de la fórmula que es común para todos los nodos del grafo pasa a tener más relevancia  $((1-\text{damping})/n)$ , en cambio cuanto mayor es el damping más acentuadas quedan las diferencias entre un nodo muy conectado y los nodos más alejados del grafo.

Además variando el damping se pueden observar diferentes resultados, ya que el cálculo de Q se hace en base al dumping y esto provoca que los aeropuertos no queden distribuidos igual, por ejemplo cambiando de 0.8 a 0.7 cambia el orden de los 3 primeros aeropuertos.