Reporte experimental - DBSCAN y Opem MP

Héctor G. T. Torres - 191589

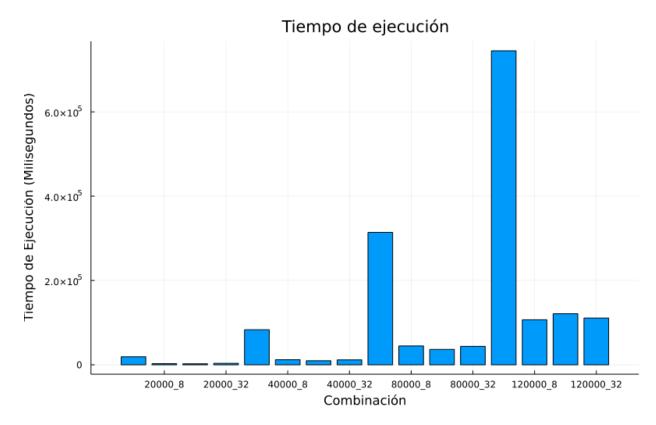
Resultados

A continuación, se muestran las gráficas con los tiempos de ejecución de cada permutación de número de hilos y cantidad de puntos. Cada gráfica se representa en formato de número de puntos_número de hilos. Por ejemplo, la barra que corresponde a la permutación de 20000 puntos y 8 hilos es la 20000_8. Las gráficas fueron generadas con Julia y los Paquetes Plots, CSV, DataFrames y Weave para exportar a un archivo PDF.

```
using DataFrames
using CSV

df = DataFrame(CSV.File("experiment_results.csv"))
```

16×2 DataFrame Combinacion Milisegundos Row String15 Float64 1 20000 1 19005.7 20000 8 2 2650.82 3 20000_16 2356.55 4 20000 32 3388.72 5 40000 1 83206.6 6 40000 8 11944.4 7 40000 16 9569.28 8 40000 32 11605.9 9 80000_1 3.14115e5 8 00008 10 44725.8 11 80000_16 36504.7 80000 32 12 43637.2 13 120000 1 7.44825e5 14 120000 8 1.06811e5 15 120000 16 1.21105e5 16 120000 32 1.10883e5



Como se puede ver en las gráficas, la mejora que representa la paralelización presenta rendimientos decrecientes: la mejora en rendimiento obtenida de pasar de 1 a 8 hilos es menor en porcentaje que la obtenida al pasar de 8 a 16 hilos. Esto se vuelve aparente al ver la diferentcia entre usar 1 y 8 hilos. En esos casos, la paralelización hizo que el algoritmo se ejecutara entre 6.5 y 7.0 veces más ráoido.

Aunado a esto, se puede observar que no hay mejora significativa al pasar de 16 (el máximo de hilos) a 32 hilos (el doble del máximo), de hecho hasta emperora un poco. Probablemente esto se debe a que la librería omp.h toma como valor predeterminado el número de hilos máximo si se ingresa un número mayor a este.

En cuanto al número de puntos, se puede observar que se mantuvo una proporción de mejora similar entre los números de puntos en cada conjunto de datos. Es decir, sin importar el número de puntos en el *dataset*, la implementación del algoritmo mejoraba en proporciones similares conforme aumentaba el número de hilos.