

Gestión de transporte bajo demanda en zonas despobladas

1. Introducción

En España existe una gran cantidad de localidades cuya población es muy reducida, lo que hace que en la mayoría de ocasiones el servicio público de autobús o tren no sea rentable, dado que circulan vacíos o con tan solo dos o tres pasajeros. Por esta razón, la frecuencia de paso se reduce drásticamente, provocando una ausencia de un medio de transporte colectivo barato y eficaz. Esto supone un problema para el desarrollo de dichas localidades, que tienen que hacer frente a problemas como:

- Las personas que no tienen acceso a un vehículo privado, ya sea por edad o por situación económica tienen muy difícil transportarse de un lugar a otro.
- El uso de vehículo privado es mucho más contaminante que el transporte colectivo.
- La dificultad de acceso hace que dichas localidades estén “aisladas” del resto del mundo.

GMV propone solucionar este problema mediante la prestación de un servicio de transporte bajo demanda. Los pasajeros podrán, desde una aplicación móvil, solicitar ser recogidos en un punto A para ser llevados a otro punto B en un autobús. Para evitar que dichos autobuses vayan con pocos pasajeros, se esperará a tener varias solicitudes, con las que se planificará una ruta y un horario de paso por cada uno de los puntos de recogida asignados por los viajeros. Esto conlleva algunas complicaciones:

- ¿Cuál es la mejor ruta para el autobús?
- ¿Cómo obtener los tiempos entre todos los puntos del recorrido si nunca se ha ido entre ellos? (por ejemplo, una persona mayor puede solicitar ser recogida en su casa)
- ¿Cómo ajustar los tiempos de recorrido a las condiciones meteorológicas, atascos y otros eventos?

Para obtener la mejor ruta del autobús será necesario resolver un **problema del viajante**, los tiempos entre los puntos del recorrido se pueden estimar a partir de datos históricos con métodos de **aprendizaje automático** y el ajuste de los tiempos de recorrido se puede realizar enriqueciendo el conjunto de datos históricos con **información externa**.

2. Problema y metodología para su resolución

Para la resolución del problema planteado anteriormente, es imprescindible tener claro las partes que lo componen y definir el marco de trabajo en el que nos moveremos.

El problema del viajante (Travelling Salesman Problem o TSP en inglés) consiste en encontrar la ruta de menor coste (tiempo, distancia, etc.) entre un conjunto de nodos, pasando por cada uno una sola vez y volviendo al origen. En nuestro caso va a servir para planificar la mejor ruta una vez tengamos la matriz de distancias de todos los puntos de recogida a todos los demás. En la Figura 1 se muestra un ejemplo de dimensión muy reducida de dos maneras de viajar entre 8 nodos, suponiendo las distancias de todos a todos conocidas.

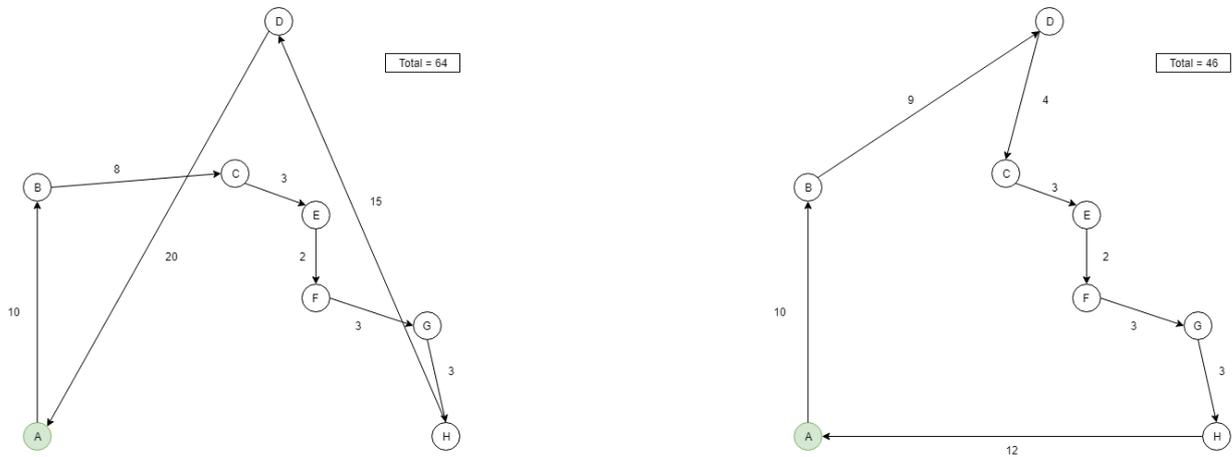


Figura 1. Problema del viajante con 8 nodos. El punto de partida está situado en el nodo A. En el panel izquierdo solución mediante un método greedy conectando cada nodo al siguiente más cercano. En el panel derecho solución mediante método heurístico. Es importante destacar que al ser ambos métodos inexactos, no tenemos la seguridad de que la solución del panel derecho sea la óptima, pero es evidente que es mucho mejor a la del panel izquierdo.

El TSP es uno de los problemas más importantes y que más han sido estudiados en investigación operativa, teoría de grafos, combinatoria y otros campos relacionados, además de ser muy importante para la resolución de problemas en la industria. Por estos motivos, existe una variedad muy grande de aproximaciones a su solución. Algunas de las vías más conocidas son:

- Algoritmos de fuerza bruta
- Algoritmos de resolución exacta mediante formulación de problemas de programación entera mixta.
- Algoritmos heurísticos y metaheurísticos

En general, debido a la complejidad del problema, no es factible resolver problemas de tamaño moderado o grande mediante algoritmos exactos ni de fuerza bruta, por lo que es necesario utilizar algoritmos inexactos basados en heurísticas para llegar a una aproximación suficientemente buena a la solución en un tiempo razonable.

Cuando los nodos son fijos y el coste de viajar entre ellos es conocido, basta con aplicar alguna de las técnicas de solución antes mencionadas, sin embargo, ¿qué pasa si no se conocen los costes entre los nodos que componen la ruta? Una solución es utilizar técnicas de aprendizaje automático.

Aprendizaje automático: El aprendizaje automático es un campo extraordinariamente extenso y de los que mayor desarrollo ha experimentado durante los últimos años. Nos vamos a centrar en los problemas de regresión, en los que a partir de unas ciertas variables de entrada denotadas por \mathbf{X} , por ejemplo latitud, longitud, tiempo atmosférico, hora de salida... Buscamos predecir una variable **continua** y, por ejemplo el tiempo empleado en realizar el recorrido. Para ello existen diversas técnicas, algunas de las más conocidas son la regresión lineal, Lasso o las máquinas vector soporte. En nuestro caso, buscaremos mediante metodologías de validación y selección de parámetros ajustar un buen modelo para predecir el tiempo que tardará nuestro autobús en recorrer la distancia de todos los nodos a todos los demás, obteniendo así la matriz de costes necesaria para resolver el TSP.

Para que el modelo sea más interesante y realice predicciones de mejor calidad, proponemos también enriquecer los datos históricos (que serán proporcionados por GMV) con datos de fuentes externas o datos derivados de las variables originales. Por ejemplo datos meteorológicos o distancias a partir de las coordenadas de partida y de destino.

3. Plan de trabajo y objetivos

Dado que el problema consta de dos partes claramente diferenciadas y la resolución completa de ambas podría suponer una inversión de tiempo y esfuerzo excesivamente grande por parte de los alumnos, se propone, en primer lugar, realizar un análisis teórico del problema del viajante y de alguno de los métodos de aproximación metaheurísticos existentes (por ejemplo, algoritmos genéticos) y continuar con un análisis práctico del problema de aprendizaje automático asociado, incluyendo la búsqueda y adquisición de datos, su cruce entre distintas fuentes y el desarrollo de modelos predictivos, culminando con una comparativa de su desempeño. El resultado final debe ser un planeador de rutas que, dada una serie de puntos de recogida, estime una matriz de distancias mediante aprendizaje automático y, después,

mediante un algoritmo de aproximación metaheurístico, resuelva el TSP. Devolviendo en última instancia la ruta que deberá seguir el conductor del autobús.

Con la propuesta de este problema se busca que los alumnos adquieran competencias en los siguientes campos:

- **Problema del viajante:** estudio teórico del problema, breve descripción de los métodos existentes para su resolución y estudio extensivo de un método metaheurístico.
- **Aprendizaje automático:** problemas de regresión, bases teóricas de los modelos más conocidos, implementación en Python de los modelos utilizando librerías específicas, técnicas de validación de los modelos, técnicas de optimización de parámetros.
- **Ingeniería de datos:** Extracción y enriquecimiento de conjuntos de datos de tamaño moderado y alto, extracción y generación de características, limpieza y calidad de datos.

Nota: La implementación del algoritmo de resolución del TSP será proporcionado por GMV.

4. Contacto

- José Carlos Baquero Triguero: jcbaquero@gmv.com
- Juan Miguel Auñón García: jmaunon@gmv.com
- Antón Makarov Samusev: amakarov@gmv.com