4th International Conference on Industrial Engineering and Industrial Management
XIV Congreso de Ingeniería de Organización
Donostia- San Sebastián , September 8th -10th 2010

Estudio de la relación entre la asimetría de las redes de transporte por carretera, el territorio y la localización

Alejandro Rodríguez[1], Rubén Ruiz[2]

[1] Grupo de Sistemas de Optimización Aplicada, Instituto Tecnológico de Informática, Universidad Politécnica de Valencia, Pza. Ferrándiz Carbonell, 2 03801 Alcoy, España arodriguez@doe.upv.es. [2] Grupo de Sistemas de Optimización Aplicada, Instituto Tecnológico de Informática, Universidad Politécnica de Valencia, Camino de Vera s/n, 46021 Valencia, España rruiz@eio.upv.es

Resumen

En esta investigación ha cuantificado de manera significativa la relación entre el grado de asimetría en las matrices de distancias de las redes de transporte por carretera y los factores territorio y localización. El diseño de experimentos comprende 11.250 tratamientos, que combinan diferentes variantes o niveles de los factores: territorio, localización y tamaño de matriz de distancias. Cada instancia requiere de una matriz de distancias c_{ij} calculada a partir de un conjunto de n geo-localizaciones con ayuda de un sistema de información geográfica. Se proponen diferentes indicadores del grado de asimetría como variables de respuesta en diferentes análisis de la varianza (ANOVA).

Palabras clave: asimetría, redes transporte carretera

1. Introducción y objetivos

En el ámbito de la logística, y más concretamente en los problemas de rutas de vehículos (VRP), es imprescindible contar con una matriz de distancias c_{ij} entre pares de localizaciones. Esta matriz también podría estar valorada en tiempo, velocidad, y/o coste necesario para comunicar cada par de localizaciones *i-j*. Aunque en cualquier caso, el tiempo, la velocidad y el coste suelen estar calculados en función de la distancia; siendo por tanto imprescindible la obtención de dicha matriz de distancias. Sea cual fuere el problema de rutas a resolver, la naturaleza de la matriz condiciona el tipo de algoritmo, heurística o meta-heurística a utilizar: distinguiendo los problemas simétricos (donde $c_{ij}=c_{ji}$ para todo *i, j*), de los asimétricos (en caso contrario). Este trabajo defiende que es fundamental conocer la naturaleza de la red de transporte y de su matriz de distancias equivalente, para poder aplicar la mejor técnica o algoritmo no sólo según el tipo de problema a resolver, sino también según el efecto que la asimetría pueda tener sobre él.

Con este trabajo de investigación* se pretende demostrar la hipótesis de que la localización de los nodos en el mundo real y el ámbito geográfico del problema determina un diferente grado de asimetría en las matrices de distancias (tiempos o costes). Se pretende cuantificar la medida en la que el grado de asimetría de las matrices de distancias depende de factores como el territorio y la localización de los clientes. Se subrayará la importancia de la obtención de las matrices de distancias reales asimétricas, y se valorará la barrera de entrada que ello supone para las empresas, y las soluciones informáticas desarrolladas. También se desea

* Este trabajo se deriva de la participación de sus autores en un proyecto de investigación financiado por el Instituto de la Mediana y Pequeña Industria Valenciana - IMPIVA, con referencia IMIDIC/2008/137 - IMIDIC/2009/198, titulado “OSC - Optimización Integral de la Cadena de Suministro”.

proporcionar a la comunidad científica un amplio conjunto de instancias o *benchmark* de características reales que aspire en el futuro a convertirse en una nueva referencia científica. Con estos análisis se pretende valorar los diferentes grados de asimetría en su verdadero contexto, en función del ámbito real del problema (territorio y localización), así como proporcionar medidas de referencia para el testeo y la mejora de las técnicas de cálculo (u obtención) de las matrices asimétricas, ayudando en definitiva a romper la barrera computacional que en la actualidad existe.

2. Experimentos

Se plantean una serie de experimentos computacionales y estadísticos con el ánimo de alcanzar los objetivos planteados en esta investigación. Primeramente se diseña el experimento, donde se detallan todos los factores que entran en el estudio, para posteriormente comentar las variables que son objeto del mismo.

2.1. Diseño de experimentos

El diseño de experimentos (DOE) es una metodología frecuentemente usada en los más diversos ámbitos de investigación y explotación (Montgomery, 2009). En un DOE se establecen unos factores y se fija una variable respuesta. Básicamente, DOE permite estudiar si el efecto de los factores en la variable respuesta es estadísticamente relevante o si por el contrario se debe al azar. Dentro de los diseños experimentales, el más sencillo es el llamado experimento factorial completo, donde se prueban todas las combinaciones posibles de todos los niveles de los factores estudiados. Se considera como instancia (*instance*) a cada uno de los casos estudiados en esta experimentación. Cada instancia está definida por una serie de propiedades que se estudian con detalle para comprender mejor su comportamiento de manera comparativa, y probar o refutar la hipótesis planteada. Los factores estudiados han sido:

– Territorio: Es un factor cualitativo. En realidad se trata de un factor continuo, pero en este caso se discretiza a través de la elección de diferentes regiones de distinto tamaño, convirtiéndose en un factor cualitativo ordinal. Las variantes seleccionadas son: Corta distancia, Media distancia y Larga distancia. Se ha elegido la Península Ibérica como territorio, tal y como muestra la Figura 1.

Figura 1. Regiones de territorio de la Península Ibérica (e.p.).

– Localización: es un factor cualitativo nominal. Se pretende analizar el comportamiento del grado de asimetría en función de la naturaleza u ordenación de las localizaciones dentro del territorio. Se han definido tres variantes: Aleatoria, Cuadrícula o rejilla (*grid*) y

Radial. La siguiente Figura 2 se muestra un ejemplo de distribución radial de las *n* localizaciones sobre un territorio.

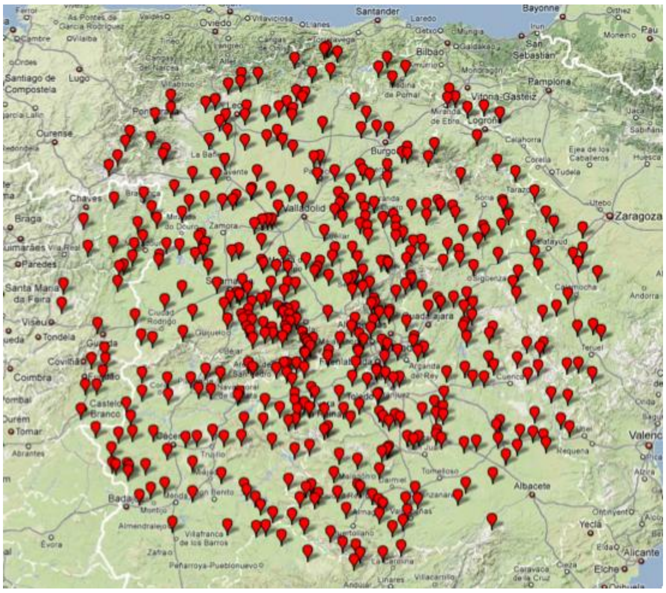

Figura 2. Instancia con localizaciones en distribución radial y territorio de largas distancias (e.p.)

– Número de nodos: es un factor cuantitativo, dependiendo del número de nodos o localizaciones cambiará el tamaño de la matriz de distancias resultante y la complejidad del problema a tratar. El rango estudiado es [50, 100, 150,…,500] nodos.

Además de los factores anteriores, también se consideró un factor réplica (5 réplicas) que proporcionará más casos de estudio para cada combinación anterior. Por tanto, habrá que construir tantas instancias de problemas a resolver como tratamientos por réplicas, en este caso: 450 tratamientos × 5 réplicas = 2250 instancias.

El reto computacional consistía en obtener y analizar: 3 territorios × 3 localizaciones × 5 réplicas ×($n \times n - 1$) arcos = 43.188.750 distancias reales mediante la realización de un igual número de consultas al sistema GIS. En cada una de estas consultas el sistema GIS ejecutaría el algoritmo de Dijkstra para el cálculo del camino mínimo entre dos localizaciones. Para valorar el orden de magnitud del problema imagine que debe calcular una matriz de distancias reales equivalente de 6572 filas × 6572 columnas, que si fuera impresa ocuparía una superficie de 2.699 m^2 (medio campo de fútbol americano, o el área de un círculo de radio igual a unos 30 metros). Para llevar a cabo el proceso de construcción de instancias con precisión, y facilitar la labor de investigación, fue necesario desarrollar un software propio. El software programado en *Ms Visual Studio .net* es compatible con el sistema operativo *Windows 7*. Este software es el encargado de realizar los cálculos necesarios para la construcción de las matrices de las instancias; ya que incorpora todo un conjunto de funciones y algoritmos (cálculos geográficos, consultas al sistema GIS, rutinas para comprobar la integridad de los datos, etc.). Esta investigación se ha llevado a cabo gracias a la más moderna y potente infraestructura* informática (redes, comunicaciones y computación) del grupo de investigación y desarrollo SOA (Sistemas de Optimización Aplicada, http://soa.iti.es) del ITI (Instituto Tecnológico de Informática). El grupo de sistemas del ITI facilita a sus investigadores un centro de computación de altas prestaciones: Cluster de computación de 30

* Financiada gracias a la Unión Europea - Fondo Europeo de Desarrollo Regional y diversos proyectos de investigación, desarrollo e innovación.

blade servers, con 16 GB de memoria RAM cada uno y dos procesadores *Intel XEON E5420* corriendo a 2,5 GHz con 4 *cores* cada uno (8 en total). Dentro de esta infraestructura, para la realización de estos experimentos se configuraron e utilizaron hasta 29 máquinas virtuales corriendo Windows XP SP3 con un núcleo a 2,5 GHz y 2 GB de memoria RAM cada una.

2.2. Medición del grado de asimetría

La simetría (o asimetría según se mire) puede ser medida en una escala continua (de completamente asimétrica a completamente simétrica) y no únicamente como una condición binaria; Saito y Yadohisa (2004). Se puede identificar el grado de simetría de las matrices de datos mediante la utilización de algún tipo de indicador matemático, tal y como se señala en Cauvin (2005). En esta investigación se utilizan los siguientes indicadores de grado de asimetría:

- Peso: el peso total de la matriz, o suma de todas las distancias c_{ij} de la misma.
- Alfa (α): es un indicador del grado de asimetría, y se expresa como el porcentaje de pares asimétricos sobre el total de pares de arcos (1). Este indicador α puede tomar valores en el intervalo [0%, 100%], desde una matriz completamente simétrica (0%), a una matriz completamente asimétrica (100%).

$$x_a(i,j) = \begin{cases} 0 & \text{si } c_{ij} = c_{ji} \\ 1 & \text{si } c_{ij} \neq c_{ji} \end{cases}$$

$$\alpha = \frac{2\sum_{a=1}^{n}(x_a)}{n^2 - n} \cdot 100 \qquad \forall i,j \in A \quad i \neq j \tag{1}$$

- Delta (δ): mide el grado de asimetría en términos de diferencias de distancias para cada par y no en términos de igualdad (2). Los valores próximos al 0% indican que los pares de arcos o la matriz es relativamente simétrica, o completamente simétrica si es igual a 0%; mientras que valores alejados del 0% subrayan un mayor grado de asimetría en cada par de arcos y en la matriz.

$$\delta_a = \frac{|\ c_{ij} - c_{ji}\ | + \min(c_{ij},c_{ji})}{\min(c_{ij},c_{ji})} \cdot 100 - 100 \qquad \forall i,j \in A \quad i \neq j \tag{2}$$

- Peso promedio: es el promedio del peso total de la matriz, esto es, la relación entre peso total de la matriz y el número total de arcos.

De los anteriores indicadores del grado de asimetría, se utilizaran δ y el peso promedio como variables de respuesta principales en diferentes análisis de la varianza, ANOVA (Fisher, 1926).

3. Análisis de los resultados

Seguidamente se presentan los análisis de los resultados obtenidos. Se analizó el tiempo de computación requerido, y la relación entre el grado de asimetría y los factores territorio y localización.

3.1. Tiempos de computación

La primera conclusión que se puede obtener de este estudio, es el altísimo coste computacional que supone obtener datos reales (matrices asimétricas) frente al ínfimo coste computacional del cálculo de las matrices ortodrómicas (matrices simétricas). Incluso mediante la utilización de 29 *blades* en su cálculo, el tiempo de computación es equivalente a 196,5 días en un sólo ordenador (4708 horas). La relación observada ha sido de 806006 a 1, entre el cálculo de las matrices asimétricas y el de las ortodrómicas. Al analizar los tiempos requeridos, se observa una gran variación de los datos debido al tamaño de la matriz, esto es, el valor de n oscila entre 50 y 500 localizaciones. Los tiempos totales oscilan entre 0,26756 horas como mínimo y 46,2064 horas de máximo para la obtención de una sola matriz. La Figura 3 ilustra en forma de histograma la distribución de frecuencias relativas (%) de los tiempos totales de obtención de las matrices asimétricas.

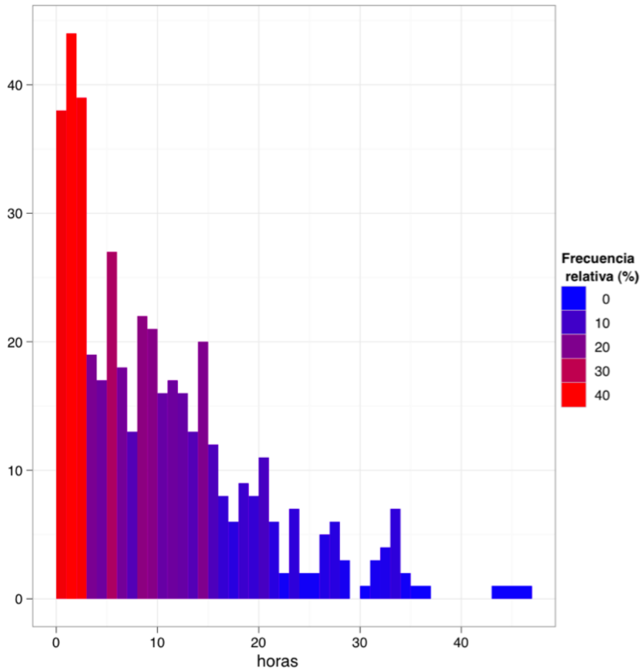

Figura 3. Histograma de tiempos totales de obtención de matrices asimétricas (e.p.).

Tras realizar diferentes estadísticas y análisis de la varianza se encontró, como era de esperar, que existe una relación significativa entre el tiempo total h y el tamaño de la matriz (n) que se puede representar como $\ln h = -6{,}45461+1{,}52892\cdot \ln n$, con un valor de $R^2 = 0{,}825348$, tal y como muestra la Figura 4.

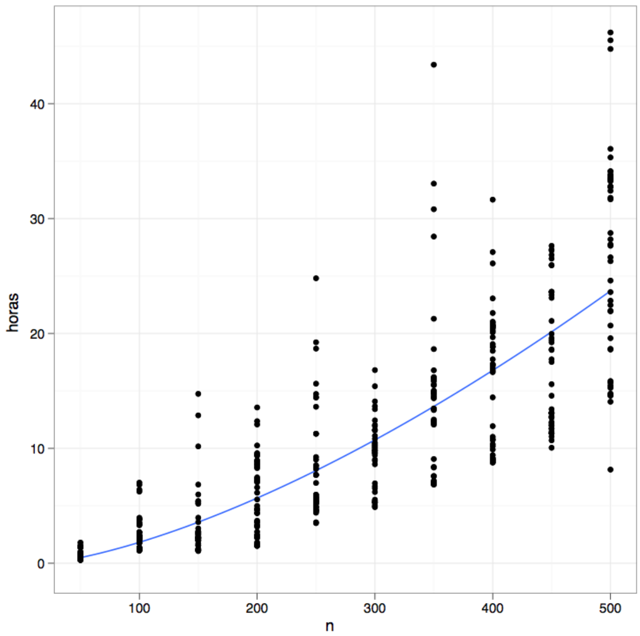

Figura 4. Gráfico de puntos: relación entre tiempo total *h* y *n* (e.p.).

3.2. Asimetría y territorio

Pasamos a corroborar estadísticamente la relación entre el grado de asimetría δ y el factor tipo de territorio (T). En la Tabla 1 se muestran los resultados de un análisis ANOVA simple δ por T para las matrices asimétricas. El valor-p de la prueba-F es menor que 0,05, así que existe una diferencia estadísticamente significativa entre la media de δ entre un nivel de T y otro, con un nivel del 95% de confianza. Nótese, que la Ratio-F tiene un valor enorme de 6618,66.

Tabla 1. Análisis ANOVA simple - Delta por T

Fuente	**Suma cuadrados**	**Grados libertad**	**Cuadrado medio**	**Ratio-F**	**Valor-*p***
Entre grupos	30 964,3	2	15 482,2	6 618,66	0,0000
Intra grupos	5 256,11	2 247	2,33917		
Total (corregido)	36 220,4	2 249			

En la Figura 5 se ha dibujado una gráfica de puntos para la relación entre el grado de asimetría δ y el indicador peso promedio, al tiempo que con un código de colores se ha señalado el factor territorio. Se puede observar la relación inversa: a mayor peso promedio, menor grado de asimetría, esto es δ menor, coincidente con el territorio de largas distancias. Claramente se observa que la dispersión en corta distancia es mayor que en el resto de casos. La función representada en la línea responde a $\delta = (0{,}745224 + 49{,}4703/PesoProm)^2$, con un valor de $R^2 = 0{,}876603$.

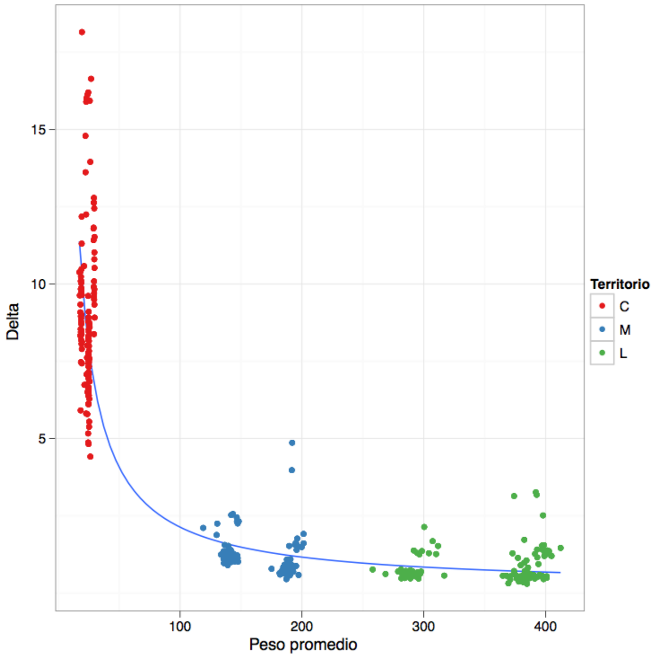

Figura 5. Gráfico de puntos: relación entre peso promedio, δ y el territorio (e.p.).

Las conclusiones más relevantes del estudio de la relación entre la asimetría y el territorio son las siguientes:

Corta distancia: en este estudio es una distancia media de 23,66 km entre cada par de localizaciones. Las matrices estudiadas para el territorio de corta distancia son altamente asimétricas con un valor promedio de α = 99,05% muy próximo al 100% (existen en promedio menos de un 1% de arcos simétricos). El indicador δ tiene un valor promedio de 8,80% de diferencia en los pares de distancias. Si se compara con las matrices ortodrómicas, el valor promedio de la distancia entre cada par de localizaciones desciende a 16,67 km (un 30 % menos que las asimétricas).

Media distancia: en este estudio se trata de una distancia media de 172,52 km entre cada par de localizaciones. Las matrices estudiadas para el territorio de media distancia son igualmente asimétricas con un valor promedio de α = 98,29% muy próximo al 100 %. Pero mucho más relevante es observar que el valor de δ desciende en comparación a la corta distancia a un valor de 1,071. Esto quiere decir que en media distancia se observan menos diferencias entre las distancias de ida y de vuelta entre cada par de arcos. Dicho de otro modo, la asimetría es un factor clave a considerar en problemas logísticos de cortas distancias (*city logistics*), y no tanto en territorios de mayor distancia. Comparativamente, el peso promedio ortodrómico es de 129,594 km (un 25% menos) frente a los 172,52 km de las matrices asimétricas.

Larga distancia: en este estudio se ha considerado una distancia media de 354,328 km entre cada par de localizaciones. Nuevamente, las matrices estudiadas para el territorio de larga distancia son asimétricas con un valor promedio de α = 99,07% muy próximo al 100%. Pero cabe señalar, que el valor del indicador δ desciende nuevamente hasta 0,797; confirmando así que a medida que aumenta la distancia promedio entre cada par de localizaciones existe una menor diferencia entre las distancias de ida y de vuelta. Es lógico pensar que en estos casos donde existen menores diferencias se podrían usar razonablemente los datos de las matrices ortodrómicas (simétricas), y no así en problemas de cortas distancias. Si se compara con las matrices ortodrómicas, el peso promedio ortodrómico es de 271,825 km (un 24% menos) frente a los 354,328 km de las asimétricas.

3.3. Asimetría y localización

Los resultados del ANOVA simple δ por L para las matrices asimétricas con un valor-p de la prueba-F igual a 0,00 y menor que 0,05, demuestran que existe una diferencia estadísticamente significativa entre la media de δ entre un nivel de L y otro, con un nivel del 95% de confianza. En la Figura 6 se comparan los factores territorio y localización en una matriz de gráficas. Dentro de cada gráfica se muestra en los ejes δ-peso promedio cada una de las matrices asimétricas. El código de colores permite identificar el factor localización (L). Esta figura resume todo el estudio de relación entre el grado de asimetría y los factores territorio y localización. Se puede observar que se al tiempo que se mantienen tres grupos de valores (que corresponden con los tres niveles del territorio), dentro de cada uno de ellos se distinguen otros dos grupos: uno que agrupa a las localizaciones aleatoria y cuadrícula, y otro para las localizaciones radiales. Nótese que las localizaciones radiales siempre tienen valores de δ superiores a las otras dos localizaciones.

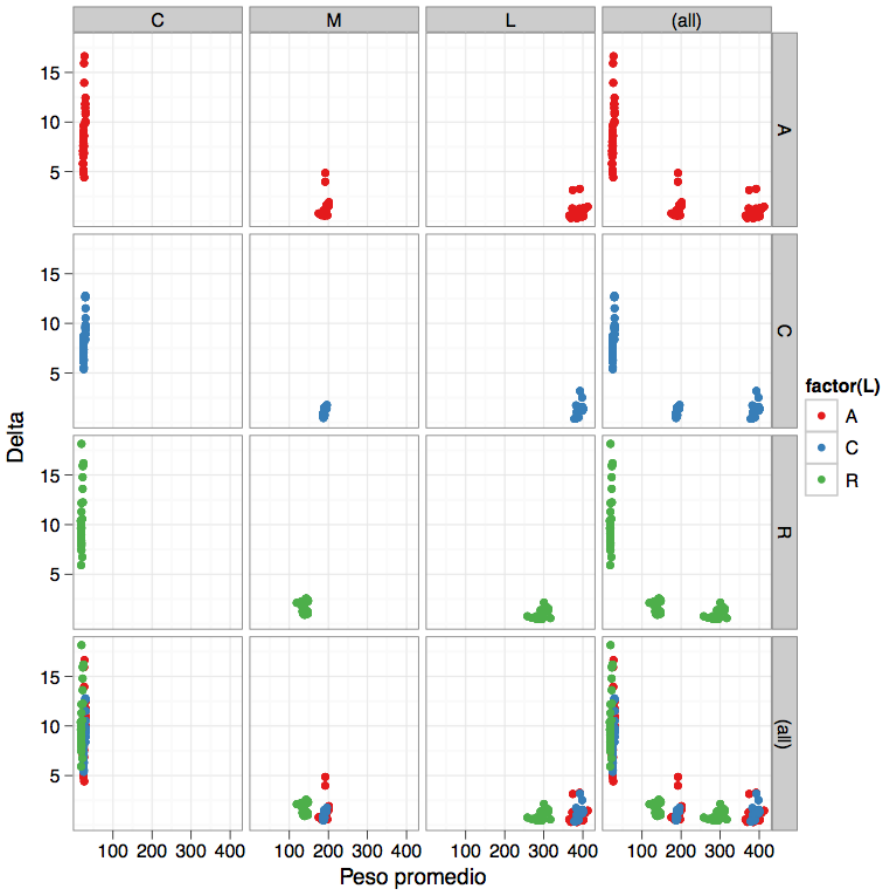

Figura 6. Relación entre el factor territorio, la localización, el grado de asimetría δ y el peso promedio (e.p.).

Tabla 2. Análisis ANOVA simple - Delta por L

Fuente	**Suma cuadrados**	**Grados libertad**	**Cuadrado medio**	**Ratio-F**	**Valor-*p***
Entre grupos	339,78	2	169,89	10,64	0,0000
Intra grupos	35880,7	2247	15,9683		
Total (corregido)	36220,4	2249			

Es curioso observar que a pesar que la red de carreteras en la Península Ibérica se caracteriza por tener una estructura radial centrada en la capital Madrid, el grado de asimetría aumenta precisamente en este tipo de localizaciones. En cualquier caso, lo relevante en este estudio es que se demuestra la relación entre el factor localización y el grado de asimetría de las matrices.

4. Conclusiones

Se ha valorado la barrera de entrada que supone para los investigadores y las empresas el alto coste computacional o tiempo necesario del cálculo y la obtención de las matrices asimétricas. En esta investigación fue necesaria la utilización de 29 *blades*, con un tiempo total de computación equivalente de 196,5 días en un solo ordenador (4.716 horas). La primera conclusión que se puede obtener es el altísimo coste computacional que supone la obtención de datos reales (matrices asimétricas). La relación observada ha sido de 806006 a 1, entre el cálculo de las matrices asimétricas y el de las ortodrómicas. Estos cálculos aportan una referencia actual sobre el coste computacional, que probablemente será útil en el futuro para valorar de manera contrastada y mejorar las técnicas de cálculo de las matrices asimétricas reales. Con todo ello, se pretende ayudar a romper la actual barrera que supone la obtención de matrices asimétricas.

Una aportación de este trabajo es que se ha confirmado estadísticamente la hipótesis de que la localización de los nodos en el mundo real (el ámbito geográfico del problema) representa diferente grado de asimetría; con lo que ello pueda suponer tanto en el tiempo de computación de las matrices, como en la resolución de los problemas de rutas por parte de diferentes algoritmos y heurísticas.

Se han utilizado con éxito diversos indicadores del grado de asimetría. Una conclusión interesante de ello, es que existen diferencias notables entre el camino de ida y el de vuelta en los pares de arcos. El indicador δ máximo ha alcanzado un valor del 18,15%, esto es, hasta una quinta parte de la distancia de diferencia. Éste es un hecho lo suficientemente importante como para desacreditar la presunción generalmente aceptada en la literatura científica de que las matrices ortodrómicas o simétricas son válidas para la resolución de problemas en entornos reales.

Ha quedado demostrado que existe una relación estadísticamente significativa entre el grado de asimetría δ y el factor territorio. Como dicho factor, a su vez está relacionado con el peso promedio, se ha calculado la función δ (peso promedio). Se puede cuantificar por tanto, la relación inversa entre el territorio y el grado de asimetría: a mayor peso promedio, menor grado de asimetría, esto es, un δ menor, coincidente con el territorio de largas distancias. Dicho de otro modo, en las largas distancias, existen menores diferencias entre el camino de ida y el de vuelta ($\delta = 0{,}798$) que en medias ($\delta = 1{,}071$), y en cortas distancias ($\delta = 8{,}801$) donde las diferencias son más notables. Esto implica que en problemas de transporte o logística en entornos de cortas distancias o ciudades, hay que prestar especial atención a la asimetría de la red, que es cuando las diferencias son mayores y las implicaciones pueden llegar a ser importantes.

Se ha demostrado estadísticamente también la relación existente entre el factor localización y el grado de asimetría. En este caso, las localizaciones radiales (para todos los territorios) hacen descender el valor del peso promedio, y esto a vez incrementa el valor del grado de asimetría δ en un 30% respecto a las localizaciones aleatorias y en cuadrícula. Es curioso observar que la infraestructura radial de la Península Ibérica favorece la asimetría en términos generales, y no al contrario como se podría pensar.

Las empresas actuales realizan sus negocios y sus operaciones logísticas en un entorno altamente competitivo y complejo. Éstas demandan soluciones eficaces y eficientes a sus problemas reales. En esta investigación se ha demostrado que incluso la realidad de los problemas más simples de transporte, es altamente compleja. El grafo de distancias (tiempos o costes) que representa la realidad del problema es sin lugar a dudas un grafo dirigido y claramente asimétrico. Obtener la matriz asimétrica de distancias de transporte incluso para problemas relativamente pequeños requiere un altísimo coste computacional, y supone una gran barrera de entrada para las empresas, y para la investigación y desarrollo de técnicas y algoritmos en entornos reales; Rodríguez y Ruiz (2009).

La asimetría es tan compleja como el mundo que nos rodea, tiene diferentes grados, y está condicionada por muchos factores; algunos de los cuales se han estudiado aquí: el territorio, la localización, y el tamaño del problema.

Referencias

Cauvin, C. (2005). A systemic approach to transport accessibility. a metho- dology developed in strasbourg: 1982-2002. *Cybergeo: European Journal of Geography*, 311:1-24.

Fisher, R. (1926). The arrangement of field experiments. Journal of the Ministry of Agriculture of Great Britain, 33:503-513.

Montgomery, D. C. (2009). Design and Analysis of Experiments. Wiley, New York, séptima edición.

Rodríguez A., Ruiz R. (2009). El impacto de la asimetría en la resolución de problemas de distribución y rutas. 3rd International Conference on Industrial Engineering and Management. XIII Congreso de Ingeniería de Organización. Barcelona-Terrassa. 1645-1654

Saito, T. y Yadohisa, H. (2004). Data analysis of asymmetric structures. CRC Press, New York.