



## RESUMEN AMPLIADO

**Título:** Estructura dinámica del área metropolitana de Valencia a través de Big Data y SIG.

**Autores y e-mail de todos ellos:** Carmen Zornoza Gallego [carmen.zornoza@uv.es](mailto:carmen.zornoza@uv.es),  
Julia Salom Carrasco [julia.salom@uv.es](mailto:julia.salom@uv.es)

Esta comunicación se elabora en el marco del proyecto “Sostenibilidad social, conectividad global y economía creativa como estrategias de desarrollo en el Área metropolitana de Valencia” (CSO2016-74888-C4-1-R), financiado por la Agencia Estatal de Investigación (AEI) y el Fondo Europeo de Desarrollo Regional (FEDER) dentro del Programa Estatal de Investigación, Desarrollo e Innovación Orientada a los Retos de la Sociedad, incluido en el Plan Estatal de Investigación Científica y Técnica y de Innovación 2013-2016, convocatoria de 2016. Carmen Zornoza cuenta con una ayuda para contratos predoctorales para la formación de doctores (BES-2014-067846) dentro del Subprograma Estatal de Formación del Programa Estatal de Promoción del Talento y su Empleabilidad, en el marco del Plan Estatal de Investigación Científica y Técnica y de Innovación 2013-2016, financiado por el Ministerio de Economía y Competitividad y cofinanciado por el Fondo Social Europeo.

**Departamento:** Instituto Interuniversitario de Desarrollo Local (IIDL)

**Universidad:** Universitat de València

**Área Temática:** Métodos de análisis espacial, sistemas de información geográfica y big data

**Resumen:** (*mínimo 1500 palabras*)

En los últimos años, la proliferación de datos provenientes del uso de la tecnología ha supuesto un reto y una gran oportunidad para las ciencias que estudian los comportamientos humanos. El tratamiento de esta gran cantidad de información, denominada de forma genérica *BIG Data*, requiere del aprendizaje de técnicas de adquisición, análisis y validación de resultados.

En trabajo en curso emplea datos geolocalizados de la red social Twitter, correspondientes al Área Metropolitana de Valencia (AMV), para plantear un estudio



dinámico sobre la movilidad de sus residentes. La movilidad ofrece información sobre los desplazamientos, así como sobre los espacios de vida de los ciudadanos, pudiendo así estudiar la sostenibilidad ambiental y social del área.

El estudio de las ciudades actuales toma como punto de partida la indefinición de sus límites, la antigua dicotomía campo-ciudad ha dado paso a espacios intermedios que hacen compleja su delimitación. Las mejoras del sistema de comunicaciones provocan una menor fricción de la distancia, lo cual hace que las actividades urbanas se dispersen en el territorio.

Ante este nuevo modelo de ciudad, muchos autores abandonan las variables socio-económicas clásicas para definirla y optan por el estudio de los flujos individuales cotidianos de su población. Los flujos mayoritarios, definidos por los desplazamientos, muestran las interacciones de los territorios y permiten reconocer los espacios más relacionados. El estudio de los recorridos habituales de la población es de gran importancia, a la vez que complejo. La obtención de datos, debido al amplísimo tamaño de la muestra, es uno de los problemas principales. Los individuos se cuentan por miles o millones, y cada uno de ellos tiene múltiples destinos diarios. Las fuentes de datos provenientes de la tecnología se revelan como una buena alternativa. Por otro lado, la metodología debe ser propuesta para alcanzar los resultados deseados, pero sin descuidar la compleja gestión de la base de datos, incluso el software y hardware disponible.

El estudio de la movilidad en el AM de Valencia ha comenzado a ser abordada por las autoras empleando datos de Twitter.

El primer paso fue estudiar las características de la fuente, de forma que el análisis se pudiera enfocar a extraer la mayor potencialidad de los datos contenidos en Twitter. Este paso fue determinante para conocer las posibilidades que ofrece la fuente para su uso en investigaciones sociales aplicadas al territorio, así como para realizar un primer análisis sobre las características sociodemográficas de sus usuarios.

Posteriormente, se trabajaron los filtros necesarios para el tratamiento de la información. Se partió de la necesidad de reconocer los usuarios que son residentes, eliminando del análisis turistas y sensores. También se aplicó un filtro orientado a eliminar grandes cantidades de tweets realizados en un corto periodo de tiempo, que responden a un uso masivo debido a un evento determinado (deporte, política...etc). El último filtro se centró en mantener a los usuarios residentes que mostraban una cantidad mínima de lugares visitados, específicamente 5. Esto nos permitía mantener a los usuarios con información mínima para obtener sus patrones de movilidad.

La metodología para detectar las residencias de los usuarios fue el siguiente hito a desarrollar en el estudio. Se determinó a partir de la posición desde la que se realizaban más tweets a lo largo del periodo estudiado. Los resultados se validaron cuantitativa y cualitativamente, permitiendo seguir la investigación en esta dirección. Los resultados revelaron que Twitter es una fuente de datos válida para el estudio de pautas de movilidad en un área urbana.

Una vez localizada la residencia de los usuarios, fue posible realizar una primera aproximación a los patrones temporales de uso de la plataforma. Se generaron franjas horarias de uso en las que se muestran como, a partir del lugar de emisión del tweet,



existen patrones espaciales de uso de la plataforma. Aquí se revela que su utilización fuera de casa es muy superior a la de dentro, lo que implica que la fuente informa sobre una gran muestra de actividades de ocio. La clave para efectuar un estudio de movilidad se encuentra en tener en cuenta que el patrón espacial es distinto según el tipo de actividad. Una actividad de ocio tendrá muchas posibilidades de generar un tweet geolocalizado, pero pocas de repetirse, mientras que un tweet geolocalizado en un lugar frecuente tiene menos posibilidades de producirse, por lo habitual de la situación, pero muchas de repetirse en el tiempo. Este hecho se observa también al analizar los tweets emitidos desde casa, que apenas se diferencian cuantitativamente entre semana y en fin de semana. Por el contrario, los registros fuera de casa en fin de semana son muy superiores a los existentes entre semana.

Se pasa en este punto a ampliar el análisis con una cuantificación de las distancias medias recorridas. Los tweets se clasifican para diferenciar entre los lugares desde los que se produce un desplazamiento (residencia) y los lugares receptores de dicho desplazamiento, diferenciando entre las actividades frecuentes y ocasionales. Se obtiene así información sobre la longitud de desplazamientos que se recorren según el tipo de actividad a la que se acude. Los resultados muestran que el 34,3% de los lugares visitados frecuentemente se encuentran a menos de 1km de la residencia y en el caso de los ocasionales este porcentaje se reduce al 17,6%. Se observa como en los espacios de vida habitual la proximidad es una estrategia necesaria para minimizar el consumo de tiempo en los desplazamientos, mientras que, en los espacios de vida ocasional los desplazamientos responden a una menor fricción por distancia. La trasposición territorial de estos datos muestra como existe una gran concentración de lugares ocasionales en el centro de Valencia.

Obtenidas las distancias medias, se cuantifica también el impacto que la urbanización residencial discontinua tiene en las distancias recorridas por sus residentes. Los resultados muestran que las distancias medias recorridas por los residentes en tejido urbano discontinuo son 4 km superiores a las del continuo. Se constata así el impacto que los crecimientos urbanos discontinuos tienen en el aumento de la movilidad y, con ello, en la cantidad de emisiones contaminantes, el consumo de tiempo y el combustible.

El análisis de la distribución de distancias realizado en este trabajo, específicamente en lugares frecuentes, ofreció pistas sobre cómo continuar la investigación. Se observaba una cola larga, que significaba que había personas que realizaban largos desplazamientos (>30km) de forma frecuente. Al adentrarnos en la realidad de estos usuarios observamos una variable no considerada hasta el momento, y es que la residencia dentro del AMV de ciertos usuarios varía en función del tiempo. Para observar este fenómeno, que es posible gracias al BIG DATA, planteamos detectar 3 pares de residencias distintas:

- **Verano** (15 de junio-15 septiembre) **Invierno** (inversa)
- **Entre semana** (lunes-jueves y viernes hasta las 18h) **Fin de semana** (inversa)
- **Horario laboral** (lunes-viernes 9-18:00) **Horario libre** (inversa)

De esta manera se afina el cálculo de desplazamientos reales. Se recoge así el dinamismo residencial del área. Al calcular distintas residencias en función al momento



temporal, los desplazamientos se obtienen sólo sobre las posiciones que registradas en esas mismas épocas. Es decir, por ejemplo, se calculan las distancias que los usuarios realizan en verano sobre la residencia de verano.

Los resultados hasta el momento muestran las siguientes distancias medias:

Etapa		Distancia media
Verano	Frecuente	8.771,03
	Ocasional	15.362,31
Invierno	Frecuente	7.716,10
	Ocasional	13.645,39
Entre semana	Frecuente	7.748,10
	Ocasional	13.453,42
Fin de semana	Frecuente	8.295,21
	Ocasional	15.376,36
Horario laboral	Frecuente	7.471,47
	Ocasional	13.766,47
Horario libre	Frecuente	7.416,62
	Ocasional	13.888,49
Posición máxima	Frecuente	10.276,0
	Ocasional	15.643,1

La mejora de los resultados es clara al comparar los resultados obtenidos a través de la posición máxima y los demás pares. En todos los casos la distancia media de la posición máxima es superior, lo cual indica que se han eliminado desplazamientos largos no realizados desde la residencia pertinente.

El análisis de las diferencias entre pares resulta también especialmente interesante.

Entre las residencias de verano e invierno se observa que las distancias recorridas en verano son mayores. Este resultado se relaciona con el traslado de parte de la población a zonas de segunda residencia con menores servicios disponibles alrededor, haciendo necesarios desplazamientos más largos.

Igualmente ocurre entre las residencias entre semana y fines de semana. Entre semana la menor disponibilidad de tiempo hace que los lugares frecuentes y ocasionales se distribuyan en lugares más cercanos que en el fin de semana.

En el caso del horario laboral y el horario libre no se observan estas disparidades. Es más, las distancias son las menores obtenidas de todo el análisis. El motivo parece apuntar a que esta diferenciación no contiene los desplazamientos que se presumen más largos, como residencia-trabajo. Es decir, al agrupar los desplazamientos dentro del horario laboral y fuera podríamos no estar contemplando muchos de los desplazamientos entre ambas franjas. Aunque esta hipótesis queda por verificar.



Los pasos a seguir desde ese punto comienzan por la validación de resultados obtenidos. En primer lugar, se empleará el censo como guía, aunque teniendo en cuenta que las residencias que se incluyen en él no se corresponden necesariamente con las obtenidas mediante nuestra distinción.

El análisis de distancias recorridas permitirá obtener una imagen de la sostenibilidad ambiental asociada a la movilidad de la zona. Con ello, será posible reconocer y localizar en el territorio los problemas y las soluciones destinadas a disminuir su impacto.

Posteriormente se asociarán los parámetros socioeconómicos de los residentes en cada área con las movibilidades observadas. De este análisis se pretende reconocer los distintos espacios de vida y si existen o no patrones de segregación social. Esta lectura puede dar una información clara de las situaciones que se dan en el área metropolitana y, en su caso, de sus disfunciones. Será posible así relacionar la realidad observada con la deseada para, en su caso, aplicar las medidas correctoras necesarias.

**Palabras Clave:** BIG Data, análisis espacial, Áreas Metropolitanas, movilidad, dinamismo.

**Clasificación JEL:** R0