

BÜYÜK MEKAN-ZAMANSAL HAREKETLİLİK VERİLERİYLE MEKANSAL ETKİLEŞİM ÖRÜNTÜLERİNİN ANALİZİ; İLİŞKİLER, KÜMELER, FONKSİYONEL BÖLGELER VE AKIŞ HARİTALARI TASARIMI

ÖZET

Varlıkların yeryüzünde bir yerden bir yere hareketleri mekânsal etkileşim dediğimiz olguyu ortaya çıkarır. Toplum ve çevreye şekil veren önemli dinamiklerden olan mekânlar arası etkileşim; insan göç hareketleri, para-mal hareketleri (ticaret), bulaşıcı hastalıkların dolaşımı, bilgi, kültür değişimi gibi nedenlerle oluşabilir.

Mekânsal ilişkilerin analizi ulaştırma, şehir ve bölge planlama, sağlık gibi birçok disiplinde kritik öneme sahiptir. Hareketlilik örüntülerinin ortaya çıkarılması, mekân ve zaman boyunca akan varlıklar arasındaki ortak noktaları anlamamızı sağlar. Sözelimi pandemik bir salgının yayılma hızı, sıklığı ve akış doğrultusunun bilinmesi epidemiyoloji uzmanları için hayati bir öneme sahiptir. Benzer şekilde kent içinde çalışma amacıyla yapılan yolculuklar şehir planlama için kentin kapasitesi, konut ve çalışma yerlerinin seçimini etkileyen bir göstergedir.

Bilgi iletişim teknolojileri ve ulaştırma alanındaki ilerlemeler günümüzde dünyayı daha önce hiç olmadığı kadar birbirine bağlı hale getirmiş ve mekânlar arasındaki insan hareketlerini artırmıştır. Bu durum coğrafyalar arası çok yönlü etkileşimlerin ortaya çıkmasına ve ilişkilerin daha karmaşık; anlaşılması güç hale gelmesine yol açmaktadır. Her alanda etkisini gösteren sayısal dönüşüm ilişkilerin de kayıt altına alınmasını sağlamaktadır. Konumsal niteliğe sahip bu verilerin analiziyle hareketlerin güdümündeki karmaşık sistemlerin anlaşılması ve gerçek dünya problemlerinin çözülmesi için karar destek araçlarının geliştirilmesi ihtiyacı vardır.

Öte yandan büyük ölçekli ve zamana bağlı hareketlilik verisinin analizi bu tür verilerin içerdiği hacim, çeşitlilik ve karmaşıklık nedeniyle oldukça güçtür. Etkileşimleri oluşturan akış verileri, mekân-zamansal ve boylamsal (belirli bir zaman dilimi için tekrar eden) nitelik taşır. Akışlara ilişkin öznitelik verilerinin çok boyutlu olması (ör; yaş, cinsiyet, eğitim durumu vb.), akış örüntülerinin zaman içinde değişmesi ve birden fazla ölçekte gerçekleşmesi (ör; ulusal, bölgesel, yerel) mekânsal etkileşimin analizini daha da güç hale getirmektedir.

Mekânsal etkileşim örüntülerinin analizi başlıca dört tema altında toplanabilir; (1) nedensellik ilişkilerine odaklanan regresyon modelleri, (2) mekânsal erişilebilirliği ortaya çıkaran istatistiksel analizler, (3) etkileşim ilişkilerini ağ yapılarıyla modelleyen çalışmalar ve (4) veri görselleştirme yeteneklerinden yararlanan etkileşimli akış haritası geliştirme araçları.

Hareketlerin nedenlerini anlamaya odaklanan çalışmalarda etkileşime yol açan etkenler sebep-sonuç ilişkileri açısından incelenir. Bu analizler belirli ekonometrik modele bağlı regresyon eşitlikleri oluşturarak bölgeler arasındaki hareketleri en iyi tahmin edecek bağımsız değişkenleri ortaya çıkarmayı amaçlar. Ancak regresyon analizleri aralarında

hareketlilik gerçekleşen her bir merkezin karakteristiğine ve çok boyutlu özelliklerine ilişkin yeterli bilgi vermez.

Erişilebilirliği inceleyen istatistiksel analizler bir merkeze gelen/giden hareketliliğin oransal büyüklüğünü tüm merkezler için karşılaştırarak erişimin makro düzeyde bir coğrafi yorumunu ortaya çıkarır. Bir bölgenin çektiği veya kaybettiği ilgiden yola çıkarak belirli bir hizmete erişim konusunda ne derece erişilebilir olduğunu değerlendirir. Öte yandan bu analizler bölge sayısı arttığında ve zaman boyutu dikkate alındığında merkezlerin karakteristiğini birbiriyle karşılaştırmakta zayıf kalmaktadır.

Etkileşim ilişkilerinin analizinde yerleşim yerlerini ve aralarındaki bağlantıları (akışları) birlikte incelemek için çizge kuramına dayanan ağ modelleri kullanılmaktadır. Ağ modeli, varlıkları birer düğüm, aralarındaki etkileşim ilişkilerini de bağlantı olarak ele alır. Mekânlar arası bağlantıların analiz edilmesiyle elde edilen parametreler ağın genel yapısı, düğümlerin karakteristiği, ağdaki kümelerin tespiti ve ağın çizimi için kullanılır.

Mekânlar arası etkileşim analizinde kullanılan diğer bir yöntem etkileşimli akış haritası tasarımı araçlarıdır. Akış haritası tasarımı; kaynak-hedef konum çiftleri arasında gerçekleşen bir hareketliliğin harita üzerinde kartografik yöntemlerle görselleştirilmesidir. Akışları harita üzerinde görselleştirmek çok değişkenli ve mekân-zamansal akışların analizi için kullanılabilir. Ancak çok boyutlu hareketlilik verilerinden harita gösterimlerinin oluşturulması bağlantı sayısının çokluğu ve harita düzlemindeki alanın sınırlı olmasından dolayı güçtür.

Bu tezin amacı; büyük ölçekli çok boyutlu mekânsal etkileşim verilerinin sistematik ve kapsamlı bir analizini yapmaktır. Bu çalışmada problemin kısıtlı bir versiyonuyla ilgilenmek yerine, bütüncül bir yaklaşımla, yukarıda belirtilen her bir tematik amaca yönelik olarak mekânsal etkileşim ilişkilerinin analizine katkıda bulunan yöntemlerin geliştirilmesi hedeflenmektedir. Bu kapsamda mekânsal etkileşim analizinde nedensellik, erişilebilirlik ve ağ modeli yaklaşımlarıyla ilgili özgün birer çalışma yapılırken akış haritaları çıkarma konusunda ise yeni bir harita tasarım aracı yazılımı geliştirilmiştir.

Tez kapsamında Türkiye genelinde iller arasında gerçekleşen hasta hareketleri mekânsal etkileşim için örnek olay olarak ele alınmaktadır. Çalışmada kullanılan veriler 2010-2013 yılları arasında 48 aylık kesitler için bir ilden diğerine tedavi amacıyla giden hasta sayılarını içermektedir. SGK'dan alınan veriler, araştırma dönemi için 1.2 milyardan fazla hastane başvurusuyla ilgili bilgiyi yansıtmaktadır. Hareketlilik verileri mekân-zamansal değişimi yansıtanın yanında kaynak il, hedef ili, hastane türü (12 tür), hizmet kademesi (3 kademe) ve tıbbi branş (120 branş) gibi öznitelik alanlarıyla çok boyutlu bir özelliğe sahiptir. Gözlem değerleri belirli bir dönem için tekrar ettiğinden araştırmada kullanılan verilerin aynı zamanda boylamsal bir niteliği vardır.

Tezde mekânsal etkileşim örüntülerinin analizinde bütüncül yaklaşımın bir parçası olarak öncelikle örnek olay hasta hareketlerinin arkasında yatan sebepler incelenmiştir. Yapılan çalışmayla Türkiye'de hasta hareketliliğine yol açan faktörler tespit edilmektedir. Panel veri ve rassal etkiler modeliyle gerçekleştirilen regresyon analiziyle daha önce farklı ülkelerdeki hasta hareketleri analizinde kullanılmış olan değişkenlerle birlikte yeni parametreler test edilmiştir. Elde edilen sonuçlar, sağlık altyapısı, demografi gibi değişkenlerin yanında Türkiye'de hedef ildeki akrabalık ilişkilerinin de etkileşim üzerinde önemli etkiye sahip olduğunu ortaya koymaktadır.

Erişilebilirliğin ele alındığı istatistiksel analizlerle benzer doğrultuda yapılan çalışmada ise, gelen/giden hareketlilik değerlerinden hizmet merkezlerinin karakteristiğini belirlemenin ötesinde, mekan-zamansal değişim örüntüleri ve grupların tespiti üzerine odaklanılmıştır. Yapılan çalışmada kullanılan sağlık hizmet bölgelerinin erişilebilirliğini değerlendiren grafik modeli (Gandy Nomogramı), zamanla değişen hareketlilik özelliklerini ve kümeleri tespit edecek şekilde ileriye taşınmıştır. Yapılan analizde erişilebilirlik metriklerine göre grafik gösterimine aktarılan gözlemlerin; (1) aylık ve yıllık zamansal değişim örüntüleri

saptanmış, (2) hiyerarşik kümeleme yöntemiyle benzer karakteristiğe sahip gruplar ortaya çıkarılmış ve (3) gruplar arası farklar istatistiksel testlerle doğrulanmıştır. Bu çalışmayla hizmet alanlarının karakteristik özelliklerini, mekân-zamansal erişilebilirlik kalıplarını ve eğilimlerini analiz eden bir yöntem geliştirilmektedir.

Ağ modelini esas alan üçüncü çalışmada, hasta hareketlilik ağı kullanılarak, mekânsal olarak sınırlandırılmış grafik bölümlene yaklaşımıyla Türkiye’de sağlık alanında işlevsel bölgeler ortaya çıkarılmaktadır. Veri güdümlü yaklaşımla tespit edilen bölgeler, Sağlık Bakanlığı’na uygulanan mevcut bölgelerle karşılaştırılmıştır. Örtüşen ve farklılaşan bölgeler sağlık hizmet planlaması hakkındaki bölgesel politikaların değerlendirilmesini sağlamaktadır. Bu çalışmada ayrıca ağ modelinden elde edilen ağırlıklı kümeleme katsayısı ve modülerlik metrikleri esas alınarak, hareketlilik ağındaki zamana göre oluşan yapısal değişiklikler tanımlanmış ve öz düzenleyici haritalar ve çok değişkenli paralel koordinat çizimi kullanılarak mekân-zamansal akış örüntüleri ortaya çıkarılmıştır. Bölgesel ve zamansal farklılıkları gözlemlediğimiz hareketlilik örüntüleri, çeşitli boyutlarda hizmet sunumu farklılıklarına ve arz-talep dengesizliklerine işaret etmektedir.

Tezin önemli bilimsel çıktılarından birisi de mekânlar arası etkileşim örüntülerinin ortaya çıkarılması amacıyla geliştirilen akış haritası görselleştirme aracıdır. Web tabanlı geliştirilen uygulama, çok boyutlu öznitelik bilgileri içeren mekânlar arası akışlardan anlam çıkarılması için gelişmiş filtreleme, sınıflandırma, akış çizgisi ve düğümleri görselleştirme (renk, büyüklük, saydamlık vb.) fonksiyonlarına sahiptir. Bu araç ile akışların başlangıç ve hedef konumlarını etkileme potansiyeli olan mekânsal dinamikleri anlamlandırma ve yorumlama mümkün hale gelmektedir. Bu ayırt edici etkiler bazı yerleşim yerlerinin özgünlüğünü ortaya koymakta ve farklı mekânsal rolleri anlamakta önemli bir unsur teşkil etmektedir.

Bu tez büyük ölçekli mekânsal etkileşim verilerinin analizini bütüncül bir yaklaşımla ele alarak politika geliştiriciler için karar aşamalarında destek sağlayan yeni yöntem ve araçlar geliştirmektedir. Araştırma; mekânlar arası etkileşimi açıklayan yeni değişkenlerin tespit edilmesi, çok sayıda hizmet merkezinin erişilebilirliğini mekân-zamansal olarak analiz eden yeni yöntemlerin geliştirilmesi, hareketlilik örüntülerinden veri güdümlü bölgelerin ortaya çıkarılarak mevcut bölgelerle karşılaştırılması, mekânlar arası akış kümelerinin tespit edilmesi, mekânsal akış haritası görselleştirme aracının geliştirilmesi gibi katkıları açısından özgün değer taşımaktadır. Ayrıca çalışma incelenen ilgi alanı (ulusal hasta hareketleri) açısından Türkiye’de, veri kapsamı (ölçek, büyüklük ve karmaşıklık) açısından da uluslararası literatürde bir ilk olma niteliğindedir

ANALYSIS OF SPATIAL INTERACTION PATTERNS USING BIG SPATIO-TEMPORAL MOBILITY DATA; RELATIONS, CLUSTERS, FUNCTIONAL REGIONS AND FLOW MAPPING

SUMMARY

Spatial interaction (SI) is a phenomenon that refers to location-to-location movements of things. Interaction between places, which is one of the essential dynamics shaping the society and the environment, can arise from human migration, commercial activities, commodity flows, infectious diseases, information or cultural exchanges.

The analysis of SI is critical in many disciplines such as transportation, urban & regional planning or health-care. Detection of mobility patterns, allows us to understand commonalities between things that move across space and time. This distinguishing feature enables SI data to be used in a wide range of research fields. For instance, it is of vital importance for epidemiologists to learn the speed of spread, frequency and flow direction of a pandemic epidemic. Similarly, commuting within cities are an indicator for city capacity planning and the choice of residential areas or work places.

The advancements in information communication technologies and transportation has made the world more interconnected than ever before and increased the human mobility between places. This causes multi-dimensional interactions between geographies and the more complex incomprehensible relationships. In other respects, digital transformations provide an opportunity to obtain information on large scale and longitudinal geographic mobility of beings. Therefore, to understand the complex systems driven by SI and solving real-world problems, there is a critical need to develop the advanced analysis methods and decision-making tools.

On the other hand, the analysis of big time-dependent mobility data is quite difficult due to its volume, variety and complexity. The flow data derived from interactions has spatial-temporal and longitudinal (repeating for a specific time slice) characteristics. The multi-dimensional attribute of mobilities (e.g. age, gender, educational status etc.) and fluctuation of flow patterns in time and scale (e.g. national, regional, local) make the analysis of spatial interaction even more difficult.

Analysis of spatial interaction patterns can be categorized under four main themes: (1) Regression models focusing on causality relationships, (2) statistical analyzes that reveal spatial accessibility, 3) studies modeling interaction relationships with network graph structures, and (4) interactive flow mapping tools benefiting data visualization capabilities.

In studies focusing on understanding the reason of mobilities, the reasons behind the interaction are examined in terms of cause-effect relationships. These researches aim to reveal independent variables that will explain the mobility between regions by creating regression equations based on a certain econometric model. However, regression analyzes

do not provide sufficient information on the features and multidimensional characteristics of all centers among which mobility occurs.

Statistical analyzes that examine accessibility reveal a geographic interpretation of access at a macro level by comparing the proportional magnitude of in-out flows of all centers. They evaluate to what extent a region has advantage in terms of accessing to a service by considering the attraction this region has gained or lost. On the other hand, these studies are weak in comparing the characteristics of centers when the number of regions increase and considering the time dimension.

To analyze simultaneously features of space and mobilities that occur between places, graph-based network models are seen as the most suitable method in which places represent by nodes, and interactions between them represent by lines. Network metrics special algorithms are used to understand the general structure of the network, the characteristics of the nodes, the detection of communities/clusters in the network and the visualization of the network.

Another method used in the analysis of SI is interactive flow map design tools. Flow map design is a visualization of mobilities happening between source-target location pairs by cartographic methods on the map. Visualizing interactions can be used for the analysis of multivariate and spatial-temporal flows. However, creating map views using multi-dimensional mobility data is difficult due to the large number of connections and limited space on the map layout.

The aim of this thesis is to make a systematic and comprehensive analysis of big multi-dimensional spatial interaction data. In this study, instead of dealing with a limited version of the problem, it is aimed to develop methods for each aforementioned thematic purpose above, to contribute to the analysis of SI with a holistic approach. In this context, we have been conducted a study on each topic; causality, accessibility and network model approaches, and we developed a new flow mapping design tool to help understanding the complex SI patterns.

Patients mobility between provinces in Turkey at national scale considered as a case for spatial interaction in the thesis. The data used in the study include the number of patients who move from one province to the other for treatment during the 48-months period between 2010 and 2013. The data obtained from Turkish Social Security Institute (SSI) reflects the information on more than 1.2 billion hospital application for the research period. In addition to reflecting spatial-temporal change, mobility data has multidimensional feature with attribute fields such as origin province, destination province, hospital type (12 types), service level (3 levels) and medical branch (120 branches). The data used in the research also have a longitudinal nature since the observation values repeat for a certain period.

In the thesis, first of all, the reasons behind the case patient mobility have been examined as a part of the holistic approach in the analysis of SI patterns. In the study conducted, factors that lead the patient mobility in Turkey are identified by performing the regression analysis in which the panel data and random effects models are employed. The new parameters were tested besides the previous variables used in patient mobility analysis in different studies. The obtained results reveal that kinship relations in target province in Turkey has a significant effect on interaction along with variables such as health infrastructure and demographics.

In line with analysis has been conducting on accessibility, the second study in this thesis, determines service centers characteristics from in-out flow values, and focuses on the determination of spatial-temporal change in flow patterns/clusters. The graphical model

(Gandy Nomogram), which evaluates the accessibility of the health service regions used in the study, has been enhanced in a way to detect the changing mobility characteristics and clusters. In the analysis, the observations were transferred to the graph layout according to the accessibility metrics, and (1) temporal variation of mobility patterns have been extracted, (2) clusters of provinces that have similar characteristics were detected by using hierarchical clustering method, and (3) differences between the clusters have been confirmed by statistical inference tests. With this study, a new method is developed to analyze characteristic features of service areas, spatio-temporal accessibility patterns and trends.

In the third study based on network model, functional regions in administrative health areas in Turkey are determined using patient mobility data by employing the spatially restricted graphical classification approach. The regions detected by the data-driven approach have been compared to the designated regions from the Turkish Ministry of Health. Overlapping and differentiated regions enable the evaluation of regional health policies on health-care planning. In this study, the structural changes in the mobility network are determined based on the weighted clustering coefficient and modularity metrics obtained from the network model, and spatio-temporal flow patterns were detected using self-organized maps and multivariate parallel coordinate plot. The patterns of mobility, which we observe regional and temporal differences, indicate differences in service provision in various dimensions and supply-demand imbalances.

One of the most important scientific outputs of the thesis is the flow mapping tool which was developed to reveal the SI patterns between places. The web-based application has advanced filtering, classification, flow line and nodes visualization functions (color, size, transparency, etc.), to gain insight from SI data which has multi-dimensional attribute. With this tool, it becomes possible to explain and interpret the spatial dynamics that have the potential to influence the starting and target positions of flows. These distinctive effects reveal the originality of some places and constitute an important element in understanding different spatial roles.

This thesis develops new methods and tools for policy-makers in decision-making processes by handling analysis of large-scale spatial interaction data with a holistic approach.

The research has unique value in terms of identifying new variables that explain the interaction between provinces in Turkey, the development of new methods in spatio-temporal analysis of accessibility for multiple service centers, uncovering data-driven regions from mobility patterns and comparison to existing regions, detection of cluster of flows between spaces, and development of spatial flow mapping tool. Furthermore, to the best of our knowledge, the research is the first study in Turkey in terms of studied field (national patient mobility) and in the international literature in terms of data scope (scale, size and complexity).