

Coğrafi Kodlama için Adres Standardizasyonu

Batuhan Kılıç^{1,*}, Fatih Gülgen²

^{1,2}Yıldız Teknik Üniversitesi, İnşaat Fakültesi, Harita Mühendisliği Bölümü, 24220, İstanbul

Özet

Adres, kentsel, sağlık, endüstriyel amaçlara bakılmaksızın herhangi bir toprak parçasının en temel tanımıdır ve coğrafi bir detayın harita üzerindeki konum bilgisidir. Metinsel odaklı tanımlamalarda, adres verisi doğrudan haritalar üzerinde gösterilemediğinden adres bilgisini enlem ve boylam açıları ile ifade edilen coğrafi koordinatlara dönüştürmek için coğrafi kodlama işlemine ihtiyaç duyulur. Coğrafi kodlamanın sağlıklı bir şekilde yapılabilmesi için yerel yönetimlerin ve diğer kurum ve kuruluşların ortak bir adres formatı kullanması gereklidir. Bu çalışmada ilk olarak ABD ve Türkiye'deki ilgi noktalarına ait adres bilgileri bir çevrimiçi coğrafi kodlama servisi üzerinden coğrafi koordinatlara dönüştürülmüştür. Elde edilen sonuçlar incelendiğinde ABD'deki adres içerikleri arasında farklılıklar olmasına karşın temelde aynı adres formatını kullandığından coğrafi kodlamanın oldukça başarılı olduğu gözlemlenmiştir. Aynı çalışma Türkiye'deki adres verileri üzerinde gerçekleştirildiğinde, adres içerikleri arasında farklılıkların yanında benzer bir adres formatının kullanılmadığı gözlenmiştir. Bu kapsamda, Türkiye'de adres verilerindeki standardizasyon eksiklikleri göz önünde bulundurularak sistemin daha sağlıklı çalışabilmesi için birtakım iyileştirmeler önerilmektedir.

Anahtar Sözcükler

Adres, Coğrafi kodlama, İlgi noktaları, Adres Standardizasyonu

Abstract

Address is the main definition of a piece of land regardless of its urban, health, industrial purposes and it is the positional information of a geographical detail on the map. In textual-oriented definitions, geocoding is required to transform an address into geographic coordinates expressed in terms of latitude and longitude, since the address data can not be displayed directly on the maps. Local governments and other agencies and organizations need to use a common address format in order for geocoding to be carried out in a healthy way. In this study, address information of points of interest in USA and Turkey was firstly converted into geographical coordinates through an online geocoding service. When the results obtained are examined, it is observed that the geocoding is very successful because it basically uses the same address format although there are differences between the address contents in the USA. When the same study was conducted on address data in Turkey, it was observed that similar address formats were not used in addition to differences between address contents. In this scope, some improvements are proposed in order to make the system work more reliable, considering the lack of standardization in addressing in Turkey.

Keywords

Address, Geocoding, Points of Interest, Address Standardization

1. Giriş

Günümüzde mobil cihazların ve sosyal ağların artması kullanıcıların coğrafi bilgiye hızlı ve ayrıntılı şekilde ulaşabilmesini sağlamaktadır. İnsanların öğrenmek istedikleri nokta detaylarına ait konum bilgisini gösteren ve onları uğrak yerlere sevk etmekte kullanılan mekânlar ilgi noktaları ("Points of Interest – POIs") olarak gösterilmektedir. İlgi noktalarına ait bilgiler coğrafi veritabanlarında ya da haritalarda tanımlı bir koordinat sistemi içinde koordinat çiftleriyle konumlandırılmaktadır. Buna karşın ilgi noktalarına ait adres verilerini temsil eden koordinat çiftlerini tespit edebilmek için adres bilgisi ile gerçek dünyadaki konumu arasında bir köprüünün kurulması gereklidir. Coğrafi kodlama, bir adres bilgisini enlem ve boylam açıları ile ifade edilen coğrafi koordinatlara çeviren bir işlem aracıdır. Coğrafi kodlama ile ilgili literatürde birçok çalışma mevcuttur (Rushton G. vd. 2006, Zhan F. Benjamin vd. 2006, Roongpiboonsopit ve Karimi 2010, Zandbergen 2008, Tian vd. 2016). Adresler üzerinden coğrafi kodlama çalışmalarının sağlıklı bir şekilde gerçekleştirilebilmesi için ilk olarak adres veritabanı kavramı temel bir bileşen olarak dikkate alınmalıdır (Davis ve Fonseca 2007, Vieira vd. 2010, Wey vd. 2009). Adres; bir toplumda yaşayan insanlar ile yaşadıkları yere ait konumları arasında bir bağ kuran önemli bir yapı olarak düşünülmektedir. Bu bağlamda, standart bir adres altyapısının oluşturulması, yapılacak mekânsal çalışmalarda doğru sonuçlara ulaşabilmeyi arttıracak, kentsel alanlarda etkili bir yönetimi sağlayabilecek ve aynı zamanda zaman ve maddi kayıpları da önleyecektir (Yomralıoğlu ve Yıldırım 2005).

* Sorumlu Yazar E-posta: batuhank@yildiz.edu.tr (Batuhan Kılıç)

Bu çalışmada, gelişmiş ülkelerden ABD’de kullanılan adres formatı coğrafi kodlama açısından incelenmiştir. Test bölgesi olarak seçilen ABD’nin Brooklyn bölgesindeki konaklama verileri ele alınarak adres bilgileri bir scraping yazılımı ile web üzerinden otomatik olarak toplanmıştır. Elde edilen adresler ile Google Geocoding servisi kullanılarak coğrafi kodlama gerçekleştirilmiştir. Benzer bir çalışma İstanbul Laleli bölgesi için de denenmiştir. İki test bölgesi arasındaki coğrafi kodlama sonuçları incelendiğinde; ülkemizde standart bir adres formatının kullanılmamasından kaynaklanan sorunlar ortaya çıktığı gözlenmiştir. Bu nedenle, coğrafi kodlama çalışmalarındaki yetersizlikler gösterilerek birtakım öneriler ortaya konulmuştur.

2. Uygulama

Brooklyn bölgesinde yer alan konaklama tesislerinin adres verilerinin toplanabileceği Booking.com ve Hotels.com isimli iki farklı web sitesi seçilmiştir. Siteler üzerinden bir Web scraping yazılımı ile alınan adres verileri karşılaştırılmış ve daha sonra bu adres bilgileri üzerinden coğrafi kodlama işlemi yapılarak adres bilgilerine ait coğrafi koordinatlar ile eşlenmiş adres verisi elde edilmiştir. Daha sonra adres veri yapısı ayrıştırılarak bir standart form haline getirilmiştir. Booking web sitesinden 60, Hotels web sitesinden ise 41 konaklama tesisi elde edilmiştir. Yapılan uygulamaya ait verilerin bir kısmı Tablo 1 ve 2’de görülmektedir.

Tablo 1. Booking.com web sitesinden alınan Brooklyn bölgesi konaklama verileri ile yapılan coğrafi kodlama sonuçları

ID	İsim	Adres	Eşleşme Durumu	Eşleşmiş Adres Formatı	Enlem	Boylam
1	McCarren Hotel & Pool	160 North 12th Street Brooklyn NY 11249	Ok	160 N 12th St Brooklyn NY 11249 USA	40.7209693	-73.9557681
2	Sleep Inn Brooklyn Coney Island	2586 Stillwell Ave Brooklyn NY 11223	Ok	2586 Stillwell Ave Brooklyn NY 11223 USA	40.5845305	-73.9832481
3	Sleep Inn Prospect Park South	1118 36th Street Brooklyn NY 11218	Ok	1118 36th St Brooklyn NY 11218 USA	40.644045	-73.987903
4	Two Bedroom Apartment - Brooklyn	Highland Place & Arlington Avenue Brooklyn NY 11208	Ok	Arlington Ave & Highland Pl Brooklyn NY 11208 USA	40.6813492	-73.8811681

Tablo 2. Hotels.com web sitesinden alınan Brooklyn bölgesi konaklama verileri ile yapılan coğrafi kodlama sonuçları

ID	İsim	Adres	Eşleşme Durumu	Eşleşmiş Adres Formatı	Enlem	Boylam
1	EVEN Hotel Brooklyn	46 Nevins Street, Brooklyn, NY, 11217	Ok	46 Nevins St, Brooklyn, NY 11217, USA	40.6876143	-73.9817682
2	Phoenix Hotel	517 39th Street, Brooklyn, NY, 11232	Ok	517 39th St, Brooklyn, NY 11232, USA	40.6510124	-74.0032181
3	Comfort Inn Red Hook	17 Seabring Street, Brooklyn, NY, 11231	Ok	17 Seabring St, Brooklyn, NY 11231, USA	40.6799958	-74.0059675
4	Days Inn Brooklyn	437 39th Street, Brooklyn, NY, 11232	Ok	437 39th St, Brooklyn, NY 11232, USA	40.652231	-74.0050354

Tablolar incelendiğinde farklı iki web sitesinden alınan konaklama tesisleri üzerinden gerçekleştirilen coğrafi kodlama işleminde verilerin tamamında bir eşleşme gerçekleşmiş ve eşleşme sonrası standart hale getirilmiş bir adres formatı görülmektedir.

Tablo 3 ve 4 ise aynı web siteleri kullanılarak Laleli bölgesinden toplanan konaklama tesislerini göstermektedir.

Tablo 3. Booking.com web sitesinden alınan Laleli bölgesi konaklama verileri ile yapılan coğrafi kodlama sonuçları

ID	İsim	Adres	Eşleşme Durumu	Eşleşmiş Adres Formatı	Enlem	Boylam
1	Hotel Bulvar Palas	Atatürk Bulvarı No:36 Sarachane Beyazıt, Fatih, 34134 İstanbul, Türkiye	Ok, partial match	Tabaklar, Atatürk Blv. No:36, 14100 Bolu Merkez/Bolu, Turkey	40.7314849	31.5998322
2	Polatdemir Hotel	Aksaray Cad. No:7 Laleli, Fatih, 34096 İstanbul, Türkiye	Ok, partial match	Mesih Paşa, Laleli Cd. No:7, 34130 Fatih/İstanbul, Turkey	41.009138	28.956438
3	Hotel Zurich Istanbul	Vidinli Tevfik Pasa Cad. No:14 Laleli, Fatih, 34470 İstanbul, Türkiye	Ok, partial match	Emin Sinan, Pertev Paşa Sk. No:14, 34130 Fatih/İstanbul, Turkey	41.0076482	28.969738
4	Best Western Premier Senator Hotel Istanbul - Old City	Gencturk Cad. Sirvanizade Sok. No 5, Vezneciler, Fatih, 34134 İstanbul, Türkiye	Ok, partial match	Zeyrek Mh., Fatih Anıt Parkı, 34083 Fatih/İstanbul, Turkey	41.0157168	28.9546731

Tablo 4. Hotels.com web sitesinden alınan Laleli bölgesi konaklama verileri ile yapılan coğrafi kodlama sonuçları

ID	İsim	Adres	Eşleşme Durumu	Eşleşmiş Adres Formatı	Enlem	Boylam
1	Hotel Bulvar Palas	Atatürk Blv. No: 36, Sarachane 34134 İstanbul, Türkiye	Ok, partial match	Tabaklar, Atatürk Blv. No:36, 14100 Bolu Merkez/Bolu, Turkey	40.7314849	31.5998322
2	Matiat Hotel İstanbul	Langa Hisari Sok. No: 19 Yenikapı - Fatih 34130 Fatih İstanbul, Türkiye	Ok, partial match	Zeyrek Mh., Fatih Anıt Parkı, 34083 Fatih/İstanbul, Turkey	41.0157255	28.9542249
3	Samir Deluxe Hotel	Büyük Resitpasa Cad. Ahmet sınıayip Sk. 22 34470 İstanbul, Türkiye	Error : Maalesef bir sunucu hatası oluştu.			
4	Hotel Monaco	Ordu Cad. sair Fitna Sk. No: 20 Laleli İstanbul, Türkiye	Error : Maalesef bir sunucu hatası oluştu.			

Tablo 3 incelendiğinde eşleşmenin tam olarak yapılamadığı ve eşleşme sonrasında ortaya çıkan standart adres formatlarında yanlış adres tayini, farklı bir şehirdeki nokta ile eşleşme vb. çeşitli hatalar olduğu görülmektedir. Tablo 4 incelendiğinde eşleşmenin neredeyse hiç gerçekleşmediği ve coğrafi kodlama işleminin yapılamadığı görülmektedir.

Her iki bölgeye ait oluşturulan standart adres yapılarına bakıldığında, Brooklyn bölgesindeki iki ayrı web sitesinden toplanan adres içerikleri arasında farklılıklar olmasına karşın temelde aynı standart adres formatını kullandıkları belirlenmiştir. Bu adresler ortak bir sistem üzerinde ayrıştırıldıklarında standart bir formata kolaylıkla getirilebilmektedir. Buna karşın Laleli bölgesinde yapılan aynı tür çalışma göz önüne alındığında Brooklyn bölgesinde olduğu gibi adres formatlarında farklılıklar olmasına rağmen sistem üzerinde ayrıştırılmadığı ve standart bir form haline getirilemediği görülmektedir.

İrdelemeler

Adres referans verileri coğrafi kodlama çalışmalarının en önemli bileşenidir. Yol isimleri, yol türleri, kısaltmalar, caddelerin çift yönlü başlangıç ve bitiş numaralarının olması ve yol merkez ekseninin doğruluğu doğrudan coğrafi kodlama çalışmalarının güvenilirliğini etkilemektedir. Bu bağlamda, ülkemizde kullanılan adres içeriklerine bakıldığında farklı bileşen türlerine sahip olduğu görülmektedir.

Bu bileşenler dışında adres üzerinden yapılacak mekânsal çalışmalarda ülkemize özgü numarataj altyapısındaki eksiklikler, yol başlangıç ve bitişlerinin yanlış tespitleri, ülkemizdeki kurum ve kuruluşların farklı adres formatlarını kullanması, idari birim sınırlarının yol ağıyla olan bağlantılarının kesinleştirilememesi ve özellikle standart bir adres formatının kullanılmaması gibi durumlar sonuçları doğrudan etkileyebilecek önemli faktörler olarak göze çarpmaktadır.

Sonuç ve Öneriler

Coğrafi Kodlama işlemi, adres eşleme ve adres standardizasyonu gibi çalışmalarda sıkça kullanılan bir yöntemdir. Ülkemizde de bu tip çalışmaların gerçekleştirilebilmesi için öncelikle standart bir adres yapısına sahip olmamız gerekmektedir. Böyle bir adres sisteminin yerleşmiş olması bundan sonraki süreçte yapılacak olan adres sistemi ile ilgili çalışmalarda kolaylıklar sağlayacağı öngörülmektedir. Aynı zamanda bu format ile ülkemize ait bir standart adres yapısının oturtulması hem zaman hem de maddi olarak olumsuzlukların yaşanmasını önleyecektir.

Coğrafi kodlama işlemi için kullanılan çevrimiçi veya bilgisayar destekli yazılımlar maalesef ülkemizin adres standart yapısına uymamaktadır. Bu nedenle adres yapımızın bu yazılımlara uygun standart hale getirilmesi veya ülkemize özgü bir yazılımın oluşturulması bundan sonraki birçok akademik ve ticari çalışmalarda kolaylık sağlayacaktır.

Kaynaklar

- Davis, C. A., & Fonseca, F. T. (2007). Assessing the certainty of locations produced by an address geocoding system. *Geoinformatica*, 11(1), 103-129.
- Roongpiboonsopit, D., & Karimi, H. A. (2010). Comparative evaluation and analysis of online geocoding services. *International Journal of Geographical Information Science*, 24(7), 1081-1100.
- Rushton, G., Armstrong, M. P., Gittler, J., Greene, B. R., Pavlik, C. E., West, M. M., & Zimmerman, D. L. (2006). Geocoding in cancer research: a review. *American journal of preventive medicine*, 30(2), S16-S24.
- Tian, Q., Ren, F., Hu, T., Liu, J., Li, R., & Du, Q. (2016). Using an Optimized Chinese Address Matching Method to Develop a Geocoding Service: A Case Study of Shenzhen, China. *ISPRS International Journal of Geo-Information*, 5(5), 65.
- Vieira, V. M., Howard, G. J., Gallagher, L. G., & Fletcher, T. (2010). Geocoding rural addresses in a community contaminated by PFOA: a comparison of methods. *Environmental Health*, 9(1), 18.

- Wey, C. L., Griesse, J., Kightlinger, L., & Wimberly, M. C. (2009). Geographic variability in geocoding success for West Nile virus cases in South Dakota. *Health & place*, 15(4), 1108-1114.
- Yomralıođlu T. ve Yıldırım V., (2005), "Trabzon Belediyesi Kent Bilgi Sistemi Projesi Kapsamında-Türkiye`de Adres Bilgisinin Kullanımı ve Sorunları Hakkında Rapor", İçişleri Bakanlığı, Trabzon.
- Zandbergen, P. A. (2008). A comparison of address point, parcel and street geocoding techniques. *Computers, Environment and Urban Systems*, 32(3), 214-232.
- Zhan, F. B., Brender, J. D., Lima, I. D., Suarez, L., & Langlois, P. H. (2006). Match rate and positional accuracy of two geocoding methods for epidemiologic research. *Annals of epidemiology*, 16(11), 842-849.