

CBS'DE KULLANILMAK ÜZERE SAYISALLAŞTIRILAN VERİLERİN GEOMETRİK/TPOLOJİK HATALARININ OTOMATİK OLARAK DÜZELTİLMESİ

İ. R. Karas¹, F. Batuk²

¹Gebze Yüksek Teknoloji Enstitüsü, Jeodezi ve Fotogrametri Mühendisliği Bölümü, Gebze, Kocaeli, ragib@gyte.edu.tr

²Yıldız Teknik Üniversitesi, Jeodezi ve Fotogrametri Mühendisliği Bölümü, Beşiktaş, İstanbul, batuk@yildiz.edu.tr

ÖZET

Coğrafi Bilgi Sistemi kurulumunda en külfetli ve pahalı aşama veri üretimidir. Bazı araştırmalara göre, bir CBS için veri toplamanın toplam maliyete oranı % 60-80 düzeyindedir. Bu nedenle gereksinilen vektör yapıdaki veriyi mevcut verilerden yararlanarak elde etmek sıkça başvurulan bir yöntemdir. Mevcut grafik verilerden yararlanmada kullanılan en yaygın yöntemlerden biri, çeşitli altlıklar üzerine çizili olan haritaların tarayıcılar yardımıyla sayısallaştırılması ve ardından vektörize edilmesidir. Bununla birlikte vektörizasyon sonucu elde edilmiş olan veri, bu haliyle bir CBS'de kullanılmaya uygun değildir. Verinin öncelikle geometrik ve topolojik olarak düzeltilmesi gerekmektedir. Bu bildiride, genel olarak topoloji kavramı üzerinde durularak sayısallaştırma sonucu ortaya çıkabilecek topolojik hatalar ortaya konulacak, bu hataların otomatik olarak düzeltilmesi amacıyla geliştirilmiş olan bir yazılım tanıtılacaktır. Yazılımda kullanılan matematiksel ve algoritmik temeller detaylandırılarak, verinin hatalardan nasıl arındırıldığı ve CBS'e uygun hale getirildiği gösterilecektir.

Anahtar Sözcükler: Sayısallaştırma, Coğrafi Bilgi Sistemleri, Topoloji, Vektörizasyon.

ABSTRACT

AUTOMATICALLY EDITING OF GEOMETRICAL AND TOPOLOGICAL ERRORS OF DIGITIZED DATA WHICH IS SUITABLE FOR GIS

The most troublesome and expensive stage in developing a Geographic Information System (GIS) is the data generation. Previous researches indicate that the proportion of data generation cost is about 60-80% of the total GIS project cost. In order to reduce total project cost, data generation method of extracting data from existing archives has been widely applied. By using scanning method, analogue format data from the archives can be transformed into digital format data, which are called raster. However, the data which has been obtained after the vectorization process needs post processing. It has to edit geometrically and topologically. In this paper, topology concept and topological errors which has been occurred after the vectorization process are described. And also an algorithm which has been developed to remove these errors is presented.

Keywords: Digitization, vectorization, GIS, topology.

1. GİRİŞ

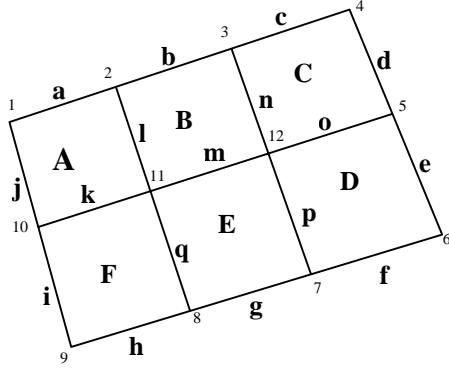
Bilgi sistemleri, organizasyonların yönetsel fonksiyonlarını desteklemek amacı ile bilgi toplayan, depolayan üreten ve dağıtan bir mekanizma olarak tanımlanır (Yomraloğlu ve Çelik, 199). Her bilgi sisteminin felsefesinde veriden bilgiye dönüştürme vardır (Sarbanoğlu, 1997). Dolayısı ile bilgiye temel teşkil eden veri olmadan bir bilgi sisteminin varlığından söz edilemez. Buna karşılık bir Coğrafi Bilgi Sistemi (CBS) kurulumunda en çok zaman alan, külfetli, pahalı ve rutin aşama veri üretimidir. Bazı araştırmacılara göre, bir CBS için veri toplamanın toplam maliyete oranı %60-80 düzeyindedir (Dickinson and Calkins 1988). Bu yüzden ihtiyaç duyulan veriyi ilk elden toplamak yerine, halihazırdaki verilerinden yararlanmak çok daha ekonomik, hızlı ve verimli bir yöntemdir. Mevcut grafik verilerden yararlanmada kullanılan en yaygın yöntemlerden biri ise çeşitli altlıklar üzerine çizili olan haritaların sayısallaştırılarak raster hale getirilmesi ve ardından vektörizasyonudur. Bununla birlikte vektörizasyon sonucu elde edilmiş olan veri, bu haliyle bir CBS'de kullanılmaya uygun değildir. Çünkü bir CBS'nin en önemli özelliği coğrafi analiz, diğer bir ifade ile konumsal analitik işlemleri gerçekleştirilme yeteneğidir (Maguire, 1992) Analiz fonksiyonları ile oluşturulacak sonuçların doğruluğu, veri ve seçilen kriterlere bağlıdır. Dolayısı ile sayısallaştırma sonucu elde edilen verinin bir CBS'de kullanılabilmesi için, öncelikle geometrik ve topolojik olarak düzeltilmesi ve modellenmesi gerekmektedir.

2. COĞRAFİ BİLGİ SİSTEMLERİ VE TOPOLOJİ

Topoloji, varlıkların metrik özelliklerinden çok birbirleriyle olan ilişkileri ile ilgilenen bir matematik dalı olarak tarif edilmiştir ve şekillerin büyüklük ve biçim özellikleri ile değil, şekil bozulmaları karşısında değişmeden kalan özellikleri ile ilgilenir. Dolayısı ile bir topolojik dönüşümde metrik özelliklerin değil, ilişkilerin ve bağlantıların

korunması esastır. Buna göre; noktaların sırası korunur, bağlantılar korunur, aynı noktadan aynı hatlar çıkar, aynı hatlar aynı noktaya birleşir.

Konuyu bir örnek üzerinde detaylandırarak olursak, Şekil 1'deki imar adası bilgisayar ekranında görüntülenmek istendiğinde, Tablo 1 ve Tablo 2'deki veriler şekli çizmek için yeterli olacaktır. Dolayısı ile bir haritayı bilgisayara çizebilmek için gerekli olan asgari veri, köşe noktalarına ait koordinatlar ve bu noktaları birleştirecek olan doğrulara ait bilgilerdir. Buna göre çizgilerin başlangıç ve bitiş noktalarını Tablo 2'den, koordinatlarını ise Tablo 1'den alarak imar adası görselleştirilebilir.



Şekil 1: İmar adası

NN	Y	X
1
2
3

Tablo 1: Koordinatlar

Çizgi No	Başlangıç NN	Bitiş NN
a	1	2
b	2	3
c	3	4
...		

Tablo 2: Doğrular

Parsel No	Çevreleyen Çizgiler
A	a, l, k, j
B	b, n, m, l
C	c, d, o, n
..	...

Tablo 3: Çokgenler (Poligonlar)

Kullanıcı, ekranda bu çizime baktığında “Bu imar adasında kaç parsel var?”, “Şu parselin komşuları hangi parsellerdir?” gibi soruların cevaplarını kolayca verebilir. Fakat şekli görüntüleyebildiği halde, Tablo 1 ve 2 deki bilgileri kullanarak yukarıda geçen soruları, bilgisayarın cevaplayabilmesi mümkün değildir. Cevaplayabilmesi için, Tablo 3'de görülen ilave verilerin de bilgisayara girilmesi gerekir. Böylece, Tablo 3'teki poligonlardan, sınır çizgileri ortak olanları belirleyerek komşu parseller hakkında bilgi verebilir. Yada “Bu adada kaç parsel var?”, “Şu parselin sınır çizgileri hangileridir?” gibi analizleri sonuçlandırabilir. Yine, parsellerin çevre ve alanlarını hesaplayabilmek için Tablo 3 deki verilerin girilmiş olması zorunludur (Karaş ve Batuk, 2005).

Yukarıdaki örnek dikkatle incelenirse Tablo 1'deki verilerin metrik bilgiler, Tablo 2 ve Tablo 3'deki verilerin ise topolojik özellikler olduğu görülecektir. Gerçekten de, şekil elastik bir yüzeye (rubber sheet) çizilip, esnetilerek, deformasyona uğratılırsa, koordinatların değiştiği, fakat Tablo 2 ve Tablo 3'te belirlenen topolojik ilişkilerin korunduğu görülecektir. Alansal detaylar (çokgenler), yine aynı çizgilerle çevrili kapalı detaylar olarak kalacak ve çizgilerin uç noktaları, dolayısı ile bağlantılar kesinlikle değişmeyecektir.

Görüldüğü gibi topoloji, aslında şekle bakan kimse tarafından “aşıkâr” olan şeylerin matematiksel olarak ifadesinden başka bir şey değildir. Sonuç olarak Coğrafi Bilgi Sistemleri açısından bir tanım yapmak gerekirse: “Topoloji, coğrafi varlıkların birbiriyle nasıl ve ne şekilde ilişkilendirildiğini geometriden bağımsız şekilde gösterme biçimidir.” denebilir (Yomraloğlu, 2000). Dolayısı ile bir CBS'de gerçekleştirilecek coğrafi analizler ancak bu ilişkilerin ortaya konması ile mümkün olabilmektedir.

3. ÇİZGİ TEMİZLEME UYGULAMASI

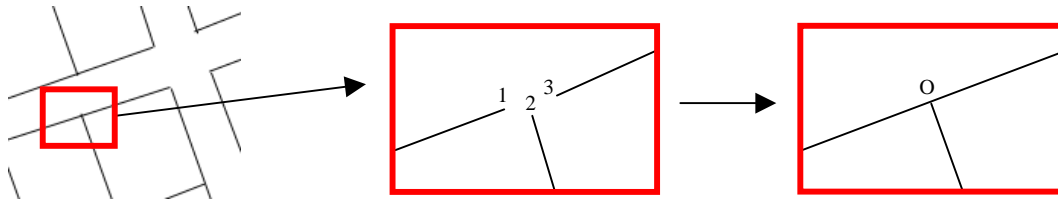
Yukarıdaki bilgilerden de anlaşılacağı üzere bir CBS'de verilerin topolojisinin sağlam bir şekilde kurulmuş olması gereklidir. Topolojide ise bağlantılılık (connectivity) esastır. Dolayısı ile ardarda gelen doğruların aynı noktada birleşmeleri olmazsa olmaz şarttır. Bununla birlikte, sayısallaştırma işleminin ardından elde edilen doğrular genellikle bağlantılılık şartını sağlamazlar. Birbirinden kopuk ve gereksiz sayıda elde edilen bu doğrular bir çok hatalı durum meydana getirirler ve sayısallaştırma sonrasında verinin işlenmesi ve bu hataların düzeltilmesi gerekir.

Geliştirdiğimiz çizgi temizleme amaçlı yazılımla raster veriseti üzerinde gerçekleştirilen sayısallaştırma işleminin ardından, elde edilen doğruların birleşim noktalarında oluşan hatalar otomatik olarak tespit edilmekte ve giderilmektedir. Bu hatalar ve algoritmanın dayandığı matematiksel temeller aşağıdaki gibidir.

3.1 Doğru parçalarının uç noktalarının bir araya getirilmesi

Vektörizasyonun ardından elde edilen çizgilerin birleşme noktalarında ortaya çıkan bu olumsuz durum, belirli bir kriterin altında birbirine yakın mesafede bulunan doğru uçlarının düzeltilmesi ile giderilmektedir. Buna göre bir arada olması gereken iki yada daha fazla doğru ucuna ait koordinatların ortalaması alınarak, bu uçlar aynı noktaya taşınmaktadır (Şekil 2).

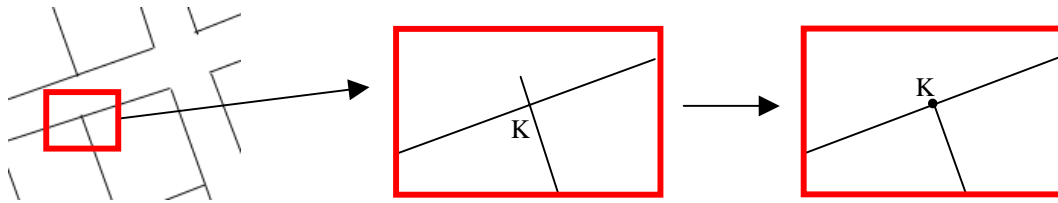
$$X_o = (X_1 + X_2 + \dots + X_n) / n \quad , \quad Y_o = (Y_1 + Y_2 + \dots + Y_n) / n \quad (1)$$



Şekil 2: Doğru parçalarının uç noktalarının bir araya getirilmesi

3.2 Sarkan doğru uçlarının düzeltilmesi

“Overshoot” hatası olarak bilinen bu hata, belirli bir limit içinde kalan sarkan doğru uçlarının olması gereken yere taşınması ile düzeltilmektedir (Şekil 3).



Şekil 3: Sarkan doğru uçlarının düzeltilmesi

Bunun için önce doğruların kesişim noktasının koordinatları hesaplanmaktadır. $A(x_a, y_a)$ ve $B(x_b, y_b)$ gibi iki uç noktası ile belirli bir doğru parçasının denklemi;

$$Y = \underbrace{\frac{(y_a - y_b)}{(x_a - x_b)}}_{m_1} X + \underbrace{\frac{(y_b x_a - x_b y_a)}{(x_a - x_b)}}_{n_1} \quad (2)$$

şekindedir. Uç noktaları $C(x_c, y_c)$ ve $D(x_d, y_d)$ olan bir doğrunun denklemi ise;

$$Y = \underbrace{\frac{(y_c - y_d)}{(x_c - x_d)}}_{m_2} X + \underbrace{\frac{(y_d x_c - x_d y_c)}{(x_c - x_d)}}_{n_2} \quad (3)$$

şeklinde olacaktır. $y = m_1 x + n_1$ ve $y = m_2 x + n_2$ şeklindeki bu iki doğru denklemini birbirine eşitlersek kesim noktasının koordinatları için;

$Y_k = \frac{(n_2 - n_1)}{(m_1 - m_2)}$, $X_k = \frac{(m_2 n_1 - m_1 n_2)}{(m_1 - m_2)}$ formülüne ulaşırız. Yukarıdaki m_1 , m_2 , n_1 ve n_2 değerlerini yerine koyarsak,

$$Y_k = \frac{((y_d x_c - x_d y_c) / (x_c - x_d) - (y_b x_a - x_b y_a) / (x_a - x_b)) / ((y_a - y_b) / (x_a - x_b) - (y_c - y_d) / (x_c - x_d))}{((y_a - y_b) / (x_a - x_b) - (y_c - y_d) / (x_c - x_d))} + \frac{((y_b x_a - x_b y_a) / (x_a - x_b))}{((y_a - y_b) / (x_a - x_b) - (y_c - y_d) / (x_c - x_d))} \quad (4)$$

$$X_k = \frac{(y_d x_c - x_d y_c) / (x_c - x_d) - (y_b x_a - x_b y_a) / (x_a - x_b)}{(y_a - y_b) / (x_a - x_b) - (y_c - y_d) / (x_c - x_d)}$$

eşitlikleri ortaya çıkacaktır. Kesişim noktasının koordinatları bu şekilde hesaplandıktan sonra aşağıdaki formül kullanılarak, bu noktanın her iki doğrunun uç noktalarına olan mesafeleri hesaplanmaktadır:

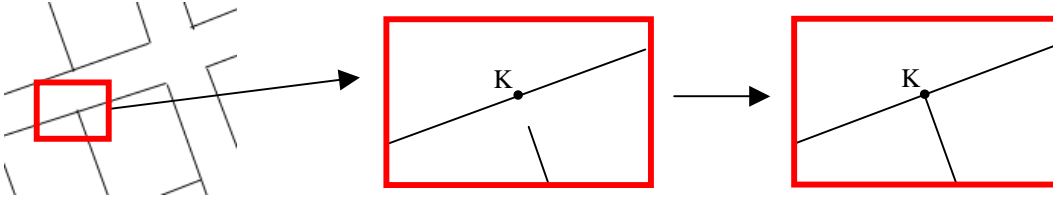
$$S = \sqrt{(x_a - x_k)^2 + (y_a - y_k)^2} \quad (5)$$

Eğer bu mesafelerden biri belirlenen kriterin altında kalıyorsa, söz konusu doğru ucunun sarkan nokta olduğu ortaya çıkarılmakta, ardından bu noktanın yeri kesişim noktasına taşınarak sarkan nokta düzeltilmektedir (Şekil 3).

3.3 Erişmeyen doğru uçlarının düzeltilmesi

“Undershoot” hatası olarak bilinen bu hata, belirli bir limit içinde kalan erişmemiş doğru uçlarının olması gereken yere taşınması ile düzeltilmektedir. Bir önceki aşamada koordinatları hesaplanmış olan kesişim noktası, doğru parçalarının dışında ise, yani doğru parçasının, uzantısı diğer doğruyu kesiyorsa ve de doğrunun uç noktası ile

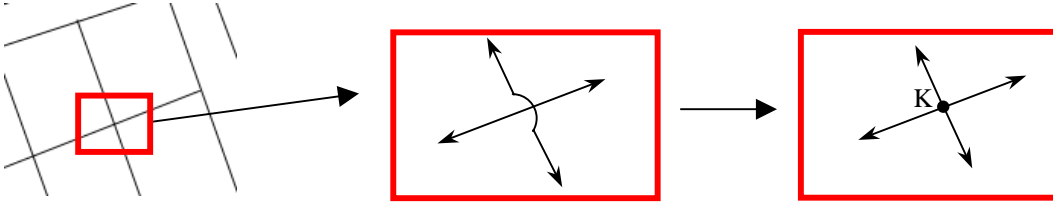
kesişim noktası arasındaki mesafe belirlenen kriterin altında kalıyorsa, sözkonusu ucun erişmeyen nokta olduğu anlaşılabilir ve kesişim noktasına taşınarak düzeltilmektedir (Şekil 4).



Şekil 4: Erişmeyen doğru uçlarının düzeltilmesi

3.4 Kesişim Noktalarının Tanımlanması

Vektörizasyon sonrasında görsel olarak doğrular kesişiyor gibi görünse de aslında kesişmeyip birbirlerinin üzerinden geçmektedir. Bu aşamada bu türden noktalar tespit edilerek, kesişim noktalarına bağlı olarak doğruların başlangıç ve bitiş koordinatları yeniden belirlenmektedir. Bunun için ilk aşamada belirtildiği şekilde koordinatları hesaplanmış olan kesişim noktasının, doğru uç noktalarına olan mesafesi, belirlenen kriterin altında kalmıyorsa, yani sarkan yada erişmeyen nokta yoksa, doğru parçaları kesişim noktası esas alınarak yeniden düzenlenmektedir. Buna göre birbirinin üstünden geçen 2 doğru parçası, kesişim noktası bunların birleşim noktası olmak üzere 4 adet doğru parçasına dönüştürülmektedir (Şekil 5).



Şekil 5: Kesişim Noktalarının Tanımlanması

4. SONUÇLAR

Sorunsuz ve sağlıklı coğrafi analiz ve sorgulamalar, sayısallaştırma hatalarından arındırılmış verilerle yapılabilir. Sayısallaştırma ne kadar dikkatli yapılırsa yapılsın hatasız olamaz. (Bildirici, 2003) Sayısallaştırma hatalarının giderilmesi için çizgi temizleme işlemlerinin gerçekleştirilmesi gereklidir. Söz konusu uygulamada, sayısallaştırma esnasında ortaya çıkan hataların otomatik olarak tespiti ve giderilmesine yönelik olarak bir çizgi temizleme yazılımı geliştirilmiş ve elde edilen sonuçlarla, sayısallaştırılan verilerin CBS'de sorunsuzca kullanılabilirliği gösterilmiştir.

KAYNAKLAR

Bildirici, İ. Ö., 2003. *Coğrafi Bilgi Sistemlerinde Ağ Topolojisi ve Genelleştirme*, 9. Türkiye Harita Bilimsel ve Teknik Kurultayı 31 Mart- 4Nisan 2003, 337-347, Ankara.

Dickinson, H., Calkins, H. W., 1988. *The economic evaluation of implementing a GIS*. International Journal of Geographic Information Systems, 2, 4, 307-327.

Karaş, İ. R., Batuk, F., 2005. *Coğrafi Bilgi Sistemlerinde Topoloji Kavramı*, 10. Türkiye Harita Bilimsel ve Teknik Kurultayı, Ankara.

Maguire, D. J., 1992. *An Overview and Definition of GIS*, in maguire D. J., Goodchild M, Rhind d (eds.), GIS Principles and Applications, Vol.1, Longman, London.

Sarbanoğlu, H., 1997. *Coğrafi Bilgi Sistemleri*, HKMO Harita bülteni, sayı 33.

Yomrahoğlu, T. ve Çelik, K., 1994. *GIS?*, 21. Ulusal CBS sempozyumu, bildirileri, Karadeniz Teknik Üniversitesi, Trabzon.

Yomrahoğlu T., 2000. *Coğrafi Bilgi Sistemleri*, Temel Kavramlar ve Uygulamaları, Seçil Ofset, 480 s.