

**T.C.
YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

MEKÂNSAL VERİ ALTYAPILARINDA GEOMETRİK ENTEGRASYON

MÜSLÜM HACAR

**YÜKSEK LİSANS TEZİ
HARİTA MÜHENDİSLİĞİ ANABİLİM DALI
UZAKTAN ALGILAMA VE CBS PROGRAMI**

**DANIŞMAN
DOÇ. DR. TÜRKAY GÖKGÖZ**

İSTANBUL, 2015

T.C.
YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

MEKÂNSAL VERİ ALTYAPILARINDA GEOMETRİK ENTEGRASYON

Müslüm Hacı tarafından hazırlanan tez çalışması 06.08.2015 tarihinde aşağıdaki jüri tarafından Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü Harita Mühendisliği Anabilim Dalı'nda **YÜKSEK LİSANS TEZİ** olarak kabul edilmiştir.

Tez Danışmanı

Doç.Dr. Türcay GÖKGÖZ
Yıldız Teknik Üniversitesi

Jüri Üyeleri

Doç.Dr. Türcay GÖKGÖZ
Yıldız Teknik Üniversitesi

Doç.Dr. Fatih GÜLGEN
Yıldız Teknik Üniversitesi

Yrd.Doç.Dr. Ahmet Özgür DOĞRU
İstanbul Teknik Üniversitesi

ÖNSÖZ

Tez çalışması süresince gösterdikleri anlayış ve verdikleri destek için annem Fatma, babam Mehmet Tevfik, ablalarım Ayşe ve Sara ve abim Abdullah’a teşekkür ederim.

Yüksek lisans eğitimim ve tez çalışması süresince bilgi ve tecrübelerinden faydalandığım, tez danışmanım ve değerli hocam sayın Doç.Dr. Türkay GÖKGÖZ’e, derslerinde edindiğim bilgi ve yönlendirmeleri için değerli hocalarım Doç.Dr. Fatih GÜLGEN, Doç.Dr. Melih BAŞARANER, Yrd.Doç.Dr. Ahmet Özgür DOĞRU’ya teşekkürlerimi sunarım.

Tez kapsamındaki uygulamada; kullandığım verisetleri için İstanbul Büyükşehir Belediyesi Coğrafi Bilgi Sistemi Müdürlüğü’ne ve Başarsoft’a, kullandığım açık kaynak eklenti (MatchingPlugin) ve optimizasyon modeli için, ve ayrıca sorularıma verdikleri yanıtlar için, Michaël MICHAUD ve Linna LI’ye teşekkür ederim.

Akademik yayın hazırlama ile ilgili bilgi ve tecrübelerinden faydalandığım Doç.Dr. Türkay GÖKGÖZ ve Yrd.Doç.Dr. Alper ŞEN’e teşekkür ederim.

Tez yazım döneminde verdikleri destek ve gösterdikleri anlayış için çalışma arkadaşlarım; Dr. Sinan ÇETİNKAYA, Abdulkadir MEMDUHOĞLU, Batuhan KILIÇ, Kadir ŞAHBAZ ve ev arkadaşım Zeynel Abidin POLAT’a teşekkür ederim.

Ayrıca dostlarım; Mehmet, Sinan, Kürşat ve Murat’a sağladıkları motivasyon için teşekkür ederim.

Ağustos, 2015

Müslüm HACAR

İÇİNDEKİLER

	Sayfa
SİMGE LİSTESİ.....	vi
KISALTMA LİSTESİ.....	vii
ŞEKİL LİSTESİ.....	viii
ÇİZELGE LİSTESİ	ix
ÖZET	x
ABSTRACT.....	xii
BÖLÜM 1	
GİRİŞ.....	1
1.1 Literatür Özeti	2
1.2 Tezin Amacı	7
1.3 Hipotez	7
BÖLÜM 2	
MEKÂNSAL VERİ ALTYAPISI	8
2.1 Ortaya Çıkışı	8
2.2 Bir İhtiyaç Olarak UMVA	8
2.3 UMVA Örnekleri.....	11
2.3.1 Dünyada UMVA	11
2.3.1.1 Amerika Birleşik Devletleri (NSDI).....	11
2.3.1.2 Avustralya (ASDI).....	12
2.3.1.3 Güney Afrika Cumhuriyeti (NSIF)	12
2.3.2 Avrupa’da UMVA ve INSPIRE Direktifi	12
2.3.3 Ülkemizde UMVA.....	14
2.4 Değerlendirme	16

BÖLÜM 3

MEKÂNSAL VERİLERİN ENTEGRASYONU.....	18
3.1 Kapsam	18
3.2 UMVA Kapsamında Veri Entegrasyonu	19
3.3 Mekânsal Veri Entegrasyonunun Çeşitleri.....	20
3.4 Mekânsal Verilerin Eşlenmesi	20
3.4.1 Öznitelik Eşlemesi	21
3.4.2 Topoloji Eşlemesi	23
3.4.3 Geometri Eşlemesi.....	23

BÖLÜM 4

UYGULAMA: YOL ORTA EKSENLERİNİN EŞLENMESİ	26
4.1 Giriş	26
4.1.1 Uygulama Bölgesi ve Kullanılan Veriler	28
4.2 Yöntemler	30
4.2.1 Açık Kaynak Yazılım (OpenJump) Eklentilerinde Eşleme Yöntemleri 31	
4.2.1.1 MatchingPlugin ile Eşleme	32
4.2.1.2 RoadMatcher ile Eşleme	40
4.2.2 Bir Optimizasyon Modeli ile Eşleme	42
4.3 Uygulama Sonuçları	45
4.3.1 MatchingPlugin Eklentisi	45
4.3.2 RoadMatcher Eklentisi.....	49
4.3.3 Optimizasyon Modeli.....	51
4.4 Sonuçların Değerlendirilmesi.....	54

BÖLÜM 5

SONUÇ VE ÖNERİLER	57
KAYNAKLAR.....	60
EK-A	
UYGULAMA SONUÇLARI	64
ÖZGEÇMİŞ	70

SİMGE LİSTESİ

β	Optimizasyon modeli çizgi uzunluğu katsayısı
D_H	Hausdorff mesafesi
d_i	i numaralı çizgi
D_{maks}	Maksimum mesafe
$D_{\text{Ö}}$	Öklid mesafesi
S_{i-j}	i nesnesi ile j nesnesi arasındaki benzerlik değeri
T_e	Eşik mesafe değeri
X_g	Çizginin ağırlık merkezinin X koordinat değeri (apsisi)
x_i	Segmentlerin orta noktalarının X koordinat değerleri
Y_g	Çizginin ağırlık merkezinin Y koordinat değeri (ordinatı)
y_i	Segmentlerin orta noktalarının Y koordinat değerleri

KISALTMA LİSTESİ

ABD	Amerika Birleşik Devletleri
ASDI	Australian Spatial Data Infrastructure
CBS	Coğrafi Bilgi Sistemi
CGDI	Canadian Geospatial Data Infrastructure
DCGF	Digital China Geospatial Framework
DEM	Digital Elevation Model
DLM	Digital Land Cover Model
FGDC	Federal Geographic Data Committee
GDI-DE	Geodateninfrastruktur Deutschland
GML	Geography Markup Language
GPL	General Public Licence
HGK	Harita Genel Komutanlığı
INSPIRE	Infrastructure for Spatial Information in Europe
ISO	International Organization for Standardization
JUMP	JAVA Unified Mapping Platform
NSDI	National Spatial Data Infrastructure
NSIF	National Spatial Information Framework
OGC	Open Geospatial Consortium
TAKBIS	Tapu ve Kadastro Bilgi Sistemi
TUCBS	Türkiye Ulusal Coğrafi Bilgi Sistemleri
UMVA	Ulusal Mekânsal Veri Altyapısı
VPF	Vector Product Format
W3C	World Wide Web Consortium
XML	Extensible Markup Language

ŞEKİL LİSTESİ

	Sayfa
Şekil 2.1	INSPIRE veri temaları [40] 14
Şekil 2.2	Standart hiyerarşisi [43] 16
Şekil 3.1	Mekânsal veri entegrasyonu: eşleme ve veri transferi..... 21
Şekil 3.2	Levenshtein mesafesi için örnek matrisler 22
Şekil 3.3	Topolojiyi kullanarak eşleme (Zhang ve Meng [16])..... 23
Şekil 3.4	Detay türüne göre eşleme işlemi (Yuan ve Tao [11]) 24
Şekil 4.1	(a) Kamu ve (b) özel şirketin ürettiği yol orta eksen verileri ile (c) uydu görüntüsü 27
Şekil 4.2	Uygulama alanının alt bölgelere ayrılması..... 28
Şekil 4.3	Bölge 1’de Veriseti 1 (yeşil) ve Veriseti 2 (kırmızı)’nin üst üste görünümü ve yolların birebir çakışmadığını gösteren örnekler 29
Şekil 4.4	Veriseti 1 (solda) ve Veriseti 2 (sağda) verisetlerinde öznitelik verileri..... 30
Şekil 4.5	Uygulamada kullanılan eşleme yöntemleri 31
Şekil 4.6	Aynı (a ve c çizgileri eşleme adayları) ve farklı sayıda segmente sahip çizgiler 33
Şekil 4.7	Eşleme adayı çizgilerin noktaları arasındaki Öklid mesafeleri 34
Şekil 4.8	Noktalar arası en kısa mesafeler 35
Şekil 4.9	Eşleme adayı çizgiler arasındaki seçilmiş mesafe D_{maks} 35
Şekil 4.10	Koordinata dayalı eşleme yöntemi ile eşlenmiş a ve b çizgileri 36
Şekil 4.11	(a) Segmentlerin ve (b) çizgilerin ağırlık merkezleri (siyah nokta)..... 37
Şekil 4.12	Noktalar ile çizgiler arasındaki en kısa mesafeler..... 38
Şekil 4.13	Seçilen en kısa mesafe 38
Şekil 4.14	Seçilen en uzun mesafe 39
Şekil 4.15	b çizgisinden a çizgisine Kısmi-Hausdorff mesafesi 40
Şekil 4.16	Eşlenen çiftler arası yakınlığın belirlenmesi [48] 42
Şekil 4.17	Kırılmış Hausdorff mesafesi [48]..... 42
Şekil 4.18	Yönlendirilmiş Hausdorff mesafesi (Li 2010) 43
Şekil 4.19	Yöntemlere ait istatistiksel verilerin sütun ve çizgi grafik ile gösterimi 55
Şekil 4.20	Bölge 1’de tüm yöntemlerin yanlış eşlediği yollar (kırmızı nokta işaretler).56
Şekil 4.21	Veriseti 1 (yeşil) ve Veriseti 2’de (kırmızı) kavşağın farklı temsil edilmesi .. 56

ÇİZELGE LİSTESİ

	Sayfa
Çizelge 3.1	Hamming mesafesi örnekleri 22
Çizelge 3.2	Detay türüne göre eşleme örnekleri (Li [22]) 24
Çizelge 4.1	Alt bölgelerde manuel olarak eşlenmiş çift sayıları 45
Çizelge 4.2	Bölge 1’de MatchingPlugin eklentisi ile elde edilen 1:N ve 1:1 eşleme sayıları 46
Çizelge 4.3	Bölge 2’de MatchingPlugin eklentisi ile elde edilen 1:N ve 1:1 eşleme sayıları 46
Çizelge 4.4	Bölge 3’de MatchingPlugin eklentisi ile elde edilen 1:N ve 1:1 eşleme sayıları 47
Çizelge 4.5	MatchingPlugin otomatik eşleme yüzdeleri ($(\text{OtomatikManuel}) \times 100$) 48
Çizelge 4.6	MatchingPlugin otomatik doğru eşleme yüzdeleri ($(\text{DoğruManuel}) \times 100$) 48
Çizelge 4.7	MatchingPlugin otomatik yanlış eşleme yüzdeleri ($(\text{YanlışOtomatik}) \times 100$) 49
Çizelge 4.8	RoadMatcher eklentisi ile elde edilen 1:N ve 1:1 eşleme sayıları 50
Çizelge 4.9	RoadMatcher 1:N ve 1:1 eşleme yüzdeleri ($\times 100$) 51
Çizelge 4.10	Bölge 1’de optimizasyon modeli ile elde edilen eşleme sayıları 51
Çizelge 4.10	Bölge 1’de optimizasyon modeli ile elde edilen eşleme sayıları (devamı) 52
Çizelge 4.10	Bölge 1’de optimizasyon modeli ile elde edilen eşleme sayıları (devamı) 52
Çizelge 4.11	Bölge 1’de α eşik mesafe ve β katsayı değerlerine göre hesaplanmış eşleme yüzdeleri ($\times 100$) 53
Çizelge A.1	Manuel eşleme (Bölge 1) 64
Çizelge A.2	Ağırlık merkezine dayalı eşleme (Bölge 1) 65
Çizelge A.3	Hausdorff mesafesine dayalı eşleme (Bölge 1) 66
Çizelge A.4	Kısmi-Hausdorff mesafesine dayalı eşleme (Bölge 1) 67
Çizelge A.5	RoadMatcher (Bölge 1) 68
Çizelge A.6	Optimizasyon modeli (Bölge 1) 69

MEKÂNSAL VERİ ALTYAPILARINDA GEOMETRİK ENTEGRASYON

Müslüm HACAR

Harita Mühendisliği Anabilim Dalı

Yüksek Lisans Tezi

Tez Danışmanı: Doç.Dr. Türkay GÖKGÖZ

Mekânsal verilerin üretimi ve paylaşımı, içinde bulunduğumuz bilgi ve yüksek teknoloji çağında hızla artmaktadır. Bir yandan üretim niceliği büyürken, diğer yandan; erişim, dönüşüm, entegrasyon, bütünleştirme vb. süreçlerde şekilsel farklılıklar, belirsizlikler ve uyum temelli sorunlar oluşabilmektedir.

Ulusal Mekânsal Veri Altyapısı, verilerin üretim ve kullanım döngüsündeki kurallarını önceden belirlenmiş standartlara göre koymakta ve süreçteki tüm politikaları ihtiyaçlara göre yönlendirmektedir. Entegrasyon süreçlerindeki harcanan çabaları en aza indirmek için verilerin standartlara göre üretilmesi gerekmektedir.

Karar alıcılar ve araştırmacılar mekânsal verisetlerinin analiz edilmesi, birleştirilmesi ya da yenilerinin üretilmesi için, farklı kaynaklardan verisetlerine ihtiyaç duymaktadır. Farklı projeksiyon, ölçek, doğruluk, amaç ve tarihlere bağlı olarak aynı varlık farklı verisetlerinde farklı özellikte nesneler ile temsil ediliyor olabilir. Bu sebeple verisetleri birleştirilirken geometrik, topolojik ve öznitelik farklılıklardan doğan bazı sorunlarla karşılaşmaktadır. Bu sorunlar entegrasyon sürecinin en önemli aşamalarından olan eşleme işleminin otomasyonunu olumsuz etkilemektedir.

Eşleme işlemi, genel olarak, aynı varlıkları temsil eden farklı verisetlerinin nesneleri arasında bağlantılar kurmaktadır. Bağlantılar nesneleri tanımlayarak ve bağlayarak aralarında köprüler inşa etmekte ve bu sayede verisetlerinin birlikte çalışabilir olmalarını sağlamaktadır. Kullanıcılar bu bağlantıları veri transferi, güncelleme ve bütünleştirme

gibi entegrasyon süreçlerinde kullanabilmektedir. Farklı kaynaklardan gelen mekânsal verilerin otomatik eşlenmesi için bilim insanları çok sayıda algoritma geliştirmiştir. Ancak bu algoritmaların çoğu veri-bağımlı olduğu için, çeşitli kaynaklardan gelen verilerin eşlenmesinde başarısız olmaktadır.

Bu çalışmada, mevcut eşleme algoritmalarından bazıları (MatchingPlugin, RoadMatcher ve optimizasyon modeli) ile Türkiye'deki biri kamu ve diğeri özel kurumda üretilen aynı bölgeye ait iki farklı yol ağı veriseti otomatik eşlenmiştir. Otomatik eşleme sonuçları manuel eşleme sonuçları ile karşılaştırılarak algoritmaların yerli verisetlerindeki başarısı test edilmiştir. Sonuçlar, test edilen algoritmaların yeterli ölçüde eşleme sayısına ulaşamadıklarını göstermiştir.

Anahtar Kelimeler: Mekânsal veri altyapısı, geometrik entegrasyon, bütünleştirme, eşleme algoritmaları, Hausdorff mesafesi

ABSTRACT

GEOMETRICAL INTEGRATION WITHIN SPATIAL DATA INFRASTRUCTURES

Müslüm HACAR

Department of Geomatic Engineering

MSc. Thesis

Adviser: Doç.Dr. Trkay GKGZ

Producing and sharing spatial data is rapidly increasing in information and high-tech age. On one hand, quantity of producing is growing, and on the other hand, shape differences, ambiguities and compactibility-based problems may occur during the processes of access, transformation, integration, conflation, etc.

National Spatial Data Infrastructure determines data rules within producing and consuming cycle according to previously determined standards and leads all the policies on this period according to needs. In order to minimize the efforts on integration process, data is required being produced with standards.

To analyze, combine or re-produce spatial datasets, decision makers and researchers need datasets from different sources. The same entity may be represented by different featured objects in different datasets depending on different projection, scale, accuracy, target and time. Therefore, some problems depending on geometric, topologic and semantic differences during combining datasets may be encountered. These problems affect matching process, one of the most challenging process of integration, negatively.

A matching process, in general, establishes links between objects from different datasets that represent the same entities. Links build bridges among objects by identifying and linking them and so make datasets more interoperable. Users may use these links during integration processes such as data transferring, updating, conflation.

Scientists developed many algorithms to match spatial data from different sources automatically. However, since most of these algorithms are data-dependent, they may not succeed in matching process of the data from varied sources.

In this study, two different road network datasets in an area produced by public and private organisations in Turkey were matched automatically by some of the existing matching algorithms (MatchingPlugin, RoadMatcher ve optimization model). The achievements of the algorithms in native datasets were tested by comparing manuel matching results with the automated matching results. Results show that the tested algorithms could not reach enough matching number.

Keywords: Spatial data infrastructure, geometrical integration, conflation, matching algorithms, Hausdorff distance

GİRİŞ

Haritalar, mekâna ait veri ve bilgilerin insanlar arasında iletilmesini sağlar. İcat edilen en eski iletişim araçları (diğerleri dil ve müzik) arasında yer almaktadır [1]. Zamanla taş, toprak, tahta vb. üzerine, kâğıdın tarih sahnesine inişiyile önce papirüs ve parşömen ardından selüloz kâğıdı üzerine çizilmiştir. Coğrafi keşifler ile haritaların üretim ve kullanımı belli bir zümrede artarken, sanayi devrimi ile birlikte modern toplum inşası için mülkiyet temelli mekânsal veriye olan gereksinim artmış ve harita üretimi çoğu toplum için bir devlet politikası haline gelmiştir. Bilgi ve teknoloji çağına girişimiz ile birlikte hem kâğıt haritalar sayısal ortama aktarılmış hem de hava fotoğrafları ve modern ölçme aletleri ile sürekli güncel haritalar üretilmeye başlanmıştır.

Modern toplumlar için; savunma, inşaat, altyapı, çevre kirliliği gibi alanlarda harita kullanımı zaruri hale gelmiştir. Kullanıcılar için harita gereksinimi arttıkça farklı çalışma alanlarından farklı kuruluşlar yetkilendirilmiştir. Bir yandan üretim niceliği ve çeşitliliği artarken, diğer yandan; erişim, dönüşüm, uygulanabilirlik, birlikte kullanım, entegrasyon vb. süreçlerde uyum temelli sorunlarla karşılaşmıştır. Ayrıca veri üretimindeki kontrolsüzlük, aynı verinin farklı kurumlar tarafından üretilmesine neden olmuştur. Mekânsal verilerin gerek birbirileri ile gerekse diğer Coğrafi Bilgi Sistemi bileşenleri; donanım, yazılım, yöntem ve insan ile uyumlu olması için belli kurallar çerçevesinde tanımlanması gerekir. Ulusal Mekânsal Veri Altyapısı (UMVA), verilerin üretim ve tüketim döngüsündeki kurallarını önceden belirlenmiş standartlara göre koyar ve süreçteki tüm politikaları ihtiyaçlara göre yönlendirir. Devletlerin sürdürülebilir gelişiminin sağlanmasında UMVA önemli rol oynar.

Yönetimler tarafından beşeri ve doğal değişimlerin kontrolünün sağlanması, afet gibi acil müdahale gerektiren olayların yönetilmesi için stratejilerin belirlenmesi, mekânsal verilerin ortak sistemlere entegrasyonu ile mümkün olmaktadır. Entegrasyon, verisetlerinin daha güncel, doğru ve ihtiyaç duyulan veri bakımından zengin olması için yapılır. Ancak üst üste veya bitişik şekilde birebir örtüşen geometriye veya ortak özniteliklere sahip verisetleri çoğu Coğrafi Bilgi Sistemi (CBS) yazılımı kullanılarak otomatik eşlenebilirken tam olarak örtüşmeyen ve birebir aynı öznitelik sütununa sahip olmayan verisetlerinin entegrasyonu için daha detaylı eşleme algoritmaları kullanmak gerekmektedir. Entegrasyon süreçlerinde başarının en kritik ölçütü olan eşleme işlemlerinin otomatik yapılamaması zaman ve ekonomik kayıplara neden olmaktadır.

Bu tez çalışmasında, eşleme yöntemleri üzerine araştırmalar yapılmış (bk. Bölüm 1), mekânsal veri altyapısı ihtiyacının ortaya çıkışı ve dünyadaki farklı örnekleri ile ülkemizin bu konudaki mevcut durumu değerlendirilmiş (bk. Bölüm 2), mekânsal veri entegrasyonu ve eşleme çeşitleri hakkında bilgiler verilmiş (bk. Bölüm 3), gerçekleştirilen uygulama ile seçilen bazı otomatik eşleme yöntemleri test edilmiş (bk. Bölüm 4) ve yapılan tüm araştırma ve uygulamalara göre bazı sonuç ve öneriler elde edilmiştir (bk. Bölüm 5).

1.1 Literatür Özeti

Bütünleştirme (map conflation), Lynch ve Saalfeld [2] tarafından “iki sayısal harita dosyasının birleştirilmesiyle, bunlardan daha iyi üçüncül bir harita dosyasının üretilmesi” olarak tanımlanmıştır. İlk uygulama olarak Amerika Birleşik Devletleri’nin (ABD) iki harita üreticisi olan Birleşik Devletler Jeolojik Araştırma ve Nüfus Sayımı İdaresi (Census Bureau) kurumlarının sayısal yol orta eksen vektör haritaları bütünleştirilmiştir [2],[3]. Bu çalışmalarda bütünleştirme işleminin en önemli adımları eşleme işlemi olarak anılmıştır. Daha sonraki araştırmaların çoğunda eşleme işlemlerinin otomasyonuna yönelik çalışmalar yapılmıştır.

Eşleme, özellikle harita bütünleştirme süreçlerinde sıklıkla kullanılan bir işlemdir. Çeşitleri hakkında detaylı bilgiler Bölüm 3.4’te verilmiştir. Eşleme işleminin ilk çözüm denemelerinde araştırmacılar verisetleri arasındaki geometrik uyumsuzluğa yönelmişlerdir [2],[3],[4]. Saalfeld [5] eşleme ve birleştirme işlemlerinin başarısını

verilerin geometrik olarak hizalanmasına (alignment) bağlamıştır. Saalfeld [5] eşleme işleminin otomasyonu için geliştirdiği yöntemini; (1) kriter (örneğin, eşlenecek iki düğüm noktası arasındaki maksimum mesafe ya da düğüm noktalarına bağlanan çizgi sayısı) belirleme aşamasında seçilen ilk (en güçlü) kriter ile en az üç noktada eşlemenin yapılması, (2) bu noktalara göre üçgenleme yapılması, (3) her bir üçgen içindeki noktaların üçgen köşelerine göre belirlenmiş rubber-sheeting parametreleriyle hizalanması, (4) üçgen içindeki noktalarla birinci adıma geçilerek sıradaki kriter ile iterasyona devam edilmesi, şeklinde tanımlamıştır. Bu yaklaşım kaynak verisetlerinin eş biçimli olduğunu varsaydığı için 1-1 (birebir) eşleme işlemleri için geliştirilmiştir.

Deretsky ve Rodny [6] geliştirdiği yöntemde iki sayısal haritaya ait geometrik ve öznitelik verileri birleştirmek için; yolların kesişim noktalarını, yol çizgilerini ve bunlara ait öznitelikleri, ilişkisel veritabanı yönetimi sistemi içerisindeki standart operatörleri kullanmıştır.

Gabay ve Doytscher [7] tarafından geliştirilen yöntemde ilk olarak sonlu yol çizgileri eşlenir. İkinci aşamada hem haritalar arasındaki ortak elementler tanımlanır hem de tek bir haritaya özgü özellikler ortaya çıkarılır. Bu sayede geometrik belirsizlikler ve topolojik farklılıklar tanımlanır. Bu yöntemde 1:m ve n:m (bir-çok ve çok-çok) eşleme gerçekleştirilebilir.

Cobb vd. [8] ABD'deki Ulusal Görüntü ve Harita Ajansı tarafından üretilen Vektör Ürün Formatı (VPF) gibi zengin öznitelik içeren verisetlerini eşlemek için yeni bir yöntem geliştirmiştir. Geleneksel yakınlık, geometri ve topoloji kriterlerine ek olarak, veri kalitesi ve detay kodu gibi semantik ve öznitelik bilgilerini de içeren veri ile ilgili tüm bilgileri kullanan bir yöntemdir. Öznitelik verilerinin semantik ve çizgi detayların şekilsel benzerliklerini temel alan yöntem, Saalfeld [5] tarafından geliştirilen yöntemdeki gibi, güçlü kriterden zayıf kritere doğru seçilen kriterlerle yeni eşlenen detay çifti kalmayana kadar iteratif işleme devam eder.

Sester vd. [9] aynı ölçekli verisetleri arasında eşleme işlemini gerçekleştirmiştir. Klasik eşik değerler yerine, haberleşme sistemlerinin alıcı ile vericisi arasındaki sinyalde meydana gelen değişikliğe dikkat çekmiştir. Verisetlerini, bilgi teorisi çerçevesinde, mevcut sinyal düzeltme algoritmaları ile eşlemiştir.

Walter ve Fritsch [10] birbirine yakın ölçekli yol ağlarını istatistiksel yaklaşım ile eşlemiştir. Yönteme göre ilk olarak eşleme öncesi afin dönüşümü ile yol ağları hizalanır. Yol çizgilerinde tampon bölgeler oluşturulur ve bu bölgeler içindeki diğer verisetine ait detaylar potansiyel detay eşi olarak belirlenir. Aday detaylar; konum, büyüklük, yön (direction) ve topolojik bilgilerine göre karşılaştırılır. Eşlenecek detaylar önceden eşlenen verilerin istatistiksel analizi ile eşik değerlere göre belirlenir.

Yuan ve Tao [11] tarafından ilk kez bütünleştirme problemi ayrıntılı olarak sınıflandırılmıştır. Bütünleştirme, “yatay” (kenarlaştırma) ve “dikey” (üst üste çakıştırma) olmak üzere iki genel sınıfa ayrılmış, en zor olarak da eşleme işleminin gerçekleştirildiği adıma dikkat çekilmiştir. Kang [12] ise eşleme işlemini “mekânsal” ve “mekânsal olmayan” eşleme olmak üzere iki türe ayırmıştır. Mekânsal eşlemede; nokta, çizgi ve alan nesneler için, geometrik ölçülerin ve topolojik özelliklerin, mekânsal olmayan eşlemede ise, istatistiksel veya kurala dayalı yöntemlerin kullanılabileceğini belirtmiştir.

Yuan ve Tao [11], alanlardan meydana gelen iki veriseti ile yapılan eşleme çalışmasında raster yontma (raster shrink) yöntemi ile alanların birinci verisetinde belirlenen ağırlık merkezlerinden yararlanarak ikinci verisetindeki eşlerini bulmuş ve böylece eşlenen alanlar arasında öznitelik transferini gerçekleştirmiştir.

Chen vd. [13] vektör yol ağını uydu görüntüsüne entegre ederek görüntüdeki yolları klasik yöntemin dörtte biri kadar az hata ile tanımlamıştır. Bunu iki aşamada gerçekleştirmiştir: (1) farklı çevrimiçi kaynaklardan gelen kontrol noktası çiftlerinin tespiti ve (2) uydu görüntüsünde görüntü işlemek için vektör verisinin bilgilerinden faydalanılması. Bu teknik aşamalar ile aynı bölgeyi temsil eden uydu görüntüleri otomatik olarak hizalanabilmekte veya vektör verisi uydu görüntüsüne entegre edilebilmektedir.

Xiong ve Sperling [14] tarafından geliştirilen yarı otomatik yöntem ile hava fotoğraflarından ve vektör formatındaki yol ağından elde edilen yol detayları eşlenebilmektedir. Nokta, segment, kenar eşlemesi otomatik olarak yapılmakta fakat bunlara ek olarak operatör yardımıyla işlem sonucunda oluşan eşleme çiftleri manuel olarak kontrol edilmektedir. Bu kontrol ile hataların tekrarlanması engellenmekte ve gerek performans gerekse güvenilirlik artmaktadır.

Samal vd. [15], ortak öznitelikleri karşılaştırmış, detaylar arasındaki öznitelik benzerlikleri saptamış, öznitelik değerleri önemli ölçüde farklı olan detayların eşlenmeyeceğini belirlemiş, şekilsel olarak detayların eşlenebilirliğini ölçmek için tampon bölge ile analiz yapmış, alan detayları için önemli yer işaretlerini (landmarks) belirleyerek verisetleri arasındaki benzerlikleri hesaplamış ve son olarak tüm bulguları her bir detay için benzerlik puanı oluşturacak şekilde toplamıştır.

Zhang ve Meng [16] tarafından Basis DLM (Digital Land Cover Model) ve TeleAtlas'ın yol tabakaları, topolojik özellikler ve tek tarafta gelişen tampon bölge (unsymmetrical buffer growing) yöntemi yardımıyla eşleştirilmiş ve TeleAtlas'taki bina numaralarını da içeren posta verileri DLM'e entegre edilmiştir.

Mustière ve Devogele [17] farklı ölçekli yol ağları üzerinde çalışmıştır. NetMatcher adı verilen eşleme yöntemine göre öncelikle eşleme aday noktaları basit bir şekilde veri bağımlı eşik mesafe değerine göre belirlenmiştir. Eşleme aday çizgileri ise Hausdorff yöntemi kullanılarak tespit edilmiştir. Ancak klasik Hausdorff mesafesi yerine, iki yol çizgisi arasında çok detaylı olandan az detaylı olana doğru hesaplanan yarı Hausdorff mesafesi kullanılmıştır. Kesin nokta eşlerinin belirlenmesi için ağ topolojisi ve noktalara bağlanan aday çizgilerin saat yönündeki sıralamaları gibi ölçütler uygulanmıştır. En yakın yol (the closest path) yaklaşımıyla da kesin yol eşleri tespit edilmiştir. Mustière ve Devogele [17] çalışmalarında IGN-Fransa'nın ürettiği iki veritabanından alınan verisetlerini test etmiştir. Ancak her ne kadar ölçekleri farklı olsa da, verisetlerinin üreticileri aynı olduğu için eşleme işleminin başarısı yeterli ölçüde test edilememiştir [18]. İleriki çalışmalarda [19], [20] yöntem Belief Teorisi ile geliştirilerek farklı kaynaktan gelen mekânsal verilere uygulanmıştır.

Kim vd. [21] Voronoi diyagramı, üçgenleme ve geometrik ölçüleri kullanarak binaları içeren büyük ölçekli verisetlerinin alan detaylarını eşlemiştir. Öznitelik verilerinden yer işaretlerini tespit ederek önce Voronoi hücreleri oluşturmuş, ardından üçgenleme yapmıştır. Üçgenlere ait geometrik ölçüleri (alan ve çevre) karşılaştırarak benzerlikleri hesaplamış ve eşleme işlemini gerçekleştirmiştir.

Li [22] ile Li ve Goodchild [23] eşleme işlemi için bir optimizasyon modeli geliştirmiştir. Bu modelde eşlenecek detayların toplam benzerliklerinin en büyük olmasını sağlayan

afin dönüşümüne göre ilgili detaylar eşlenir. Yöntemle ilgili ayrıntılı bilgiler Bölüm 4.2.2’de verilmiştir.

Song vd. [24] eşleme işlemini güvenilirlik matrisi oluşturarak gerçekleştirmiştir. Veri setleri arasındaki maksimum ötelenme miktarına göre bir eşik mesafesi belirlenmiştir. Birbirine eşik mesafeden daha yakın olan ikinci veri setindeki her bir nokta eşleme adayı olarak kabul edilmiştir. Birinci verisetindeki noktaya bağlanan çizgi sayısı ile her bir eşleme aday noktasına bağlanan çizgi sayısı karşılaştırılmıştır. Çizgi sayısı birinci veri kümesindeki noktanın çizgi sayısına eşit olan eşleme adayı noktalara, oluşturulan güvenilirlik matrisinde 1 değeri atanmıştır. Belirlenen uygunluk denklemiyle iki iterasyon yapıldıktan sonra maksimum güvenilirliğe sahip olanlar eşleştirilmiştir.

Pourabdollah vd. [25] açık verilerin (open data) kalitesini arttırmak için eşleme işlemini yapmıştır. Devamlı güncellenen fakat öznitelik bakımından eksik olan Open Street Map verileri ile öznitelik bakımından zengin Ordnance Survey (İngiliz Ulusal Harita Ajansı) verilerini bütünleştirerek hem güncel hem de öznitelik bakımından daha kaliteli bir veriseti elde etmiştir.

Yang vd. [26], yol ağını oluşturan çizgileri desenlerine (pattern) göre sınıflandırarak (çok parçalı veya tek parçalı) eşleme adaylarını belirlemiştir. Yönteme göre, klasik geometrik ve topolojik kriterleri sağlayan çizgilerden aynı desen sınıfına girenler eşlenmektedir.

Eşleme işlemi sadece entegrasyon ve bütünleştirme süreçlerinde kullanılmaz. Fotogrametride farklı hava fotoğraflarından otomatik elde edilen aynı yol ve bina çizgi nesnelerinin seçimi için eşleme yöntemleri kullanılır [27]. Ayrıca, Woo ve Dipanda [28] genetik biliminde kullanılan gen modeli çizgilerinin, krossover ve mutasyon gibi kriterlere göre etiketlendiği genetik algoritmalarını kullanarak, aktif stereo görüntü sisteminde çizgi ve nokta eşlemiştir. Zhang ve Shu [29] ise, kullanımı giderek artan avuç içi tarama ve tanımlama sistemlerinde, avuç içi çizgilerinin aynı kişiye ait olup olmadığının saptanması için bir eşleme yöntemi geliştirmiştir.

1.2 Tezin Amacı

Bu tez çalışması, mekânsal veri entegrasyonu sürecinin en kritik aşamalarından biri olan eşleme işleminin, farklı kurumlar tarafından üretilmiş, İstanbul ili Beykoz ilçesine ait yol orta eksen verisetlerinde test edilmesi amacı ile yapılmıştır.

1.3 Hipotez

Bu tez aşağıdaki hipotezlere dayanmaktadır:

- Mekânsal veri entegrasyonu süreçlerinin en kritik aşamalarından biri eşleme işlemidir.
- Mevcut eşleme algoritmaları veri-bağımlıdır.

BÖLÜM 2

MEKÂNSAL VERİ ALTYAPISI

2.1 Ortaya Çıkışı

1980'lerin başında, Canada'daki eyalet ve şehir harita ajansları arasındaki mekânsal veri transferi için üretim standartları ve değişim protokollerinde ilk kez mekânsal bilgi altyapısı (Geo Information Infrastructure) terimi kullanılmıştır [1].

1992'de Rio de Janeiro'da gerçekleştirilen Çevre ve Gelişim Üzerine Birleşmiş Milletler Konferansı'nda küresel çevre sorunlarıyla ilgili Agenda 21 raporu oluşturulmuştur. Agenda 21'de, çevreyi algılamayı geliştiren önemli bir araç olarak kaliteli mekânsal veriye olan ihtiyaç vurgulanmış ve küresel çevre sorunlarının tüm dünya milletlerinin koordinasyonlu bir şekilde çalışmasıyla azaltılabileceğinden söz edilmiştir [30].

1994'te ABD başkanı Bill Clinton coğrafi verilerin üretim ve erişimini yönetmek için Ulusal Mekânsal Veri Altyapısını (NSDI) kurduğunu duyurmuştur. Bu hareket UMVA kurmak için atılan ilk somut adım olarak kabul edilebilir [1], [31].

2.2 Bir İhtiyaç Olarak UMVA

Mekânsal veriler, dijital dünyaya girmesinin hemen ardından, günlük hayatın içerisinde her zaman ve her yerde kullanılabilir hale gelmiştir. Hükümetler tarafından tıpkı diğer kamu hizmetleri gibi vatandaşlara sunulması gereken bir hizmet olmuştur. Ancak üretim ve tüketim döngüsünün kurallarının belli olmaması; erişim, paylaşım, güncelleme, entegrasyon vb. süreçlerde eşgüdümlü çalışmayı engellemektedir. Örneğin, sunucular için paylaşım stratejisinin belirlenmemesi, kullanıcıların edindikleri verileri kullanırken

uyum sorunları yaşamalarına neden olmaktadır. Ya da mevcut mekânsal veri ürünlerinin katalog ve dizinlerine erişilememesi, kurumların aynı veriyi birden fazla üretmelerine (dublication) neden olmaktadır. Bunlarla birlikte, birden fazla üretilmiş uyum sorunu olan verilerin başka verilerle birleştirilmesi ya da güncelleme ve entegrasyon gibi işlemlere girdi olması, sonlandırılması zor süreçlere geçmek demektir.

CBS, mekânsal verilerin üretildiği, yönetildiği, entegre edildiği, analiz edildiği ve görüntülendiği bilgisayar sistemleridir [31]. Raster veya vektör veri yapıları ile çalışılabilmektedir.

İçme sularının kalitesinin arttırılmasından sel bölgelerinin belirlenmesine, kıyı bölgelerindeki yapılaşmadan turizm alanlarının yönetimine kadar birçok sistem CBS teknolojisine ihtiyaç duyar. Bu teknoloji birçok farklı kaynaktan gelen büyük miktarda mekânsal veriyi işleyebilmektedir [31]. Mekânsal verilerin gerek birbirileri ile gerekse diğer CBS bileşenleri; donanım, yazılım, yöntem ve insan ile uyumlu olması için önceden belirlenmiş kriterlere bağlı olarak tanımlanması şarttır. UMVA, verilerin üretim ve tüketim döngüsündeki kurallarını ISO (International Organization for Standardization) ve OGC (Open Geospatial Consortium) gibi önceden belirlenmiş standartlara göre koyar ve süreçteki tüm politikaları ihtiyaçlara göre yönlendirir. Birinci dereceden mekânsal verilerin ortak kullanım ve değerlendirilmesine bağlı olan; taşkın, yangın, deprem gibi afetlere müdahalelerde UMVA'ların belirlediği risk ve afet yönetim planları zararı en aza indirme amacıyla hazırlanmaktadır. Örneğin, Edirne ili içerisinde geçen Meriç nehri her yıl taşkınlara neden olmaktadır. Nehir, Bulgaristan'da doğarak Türkiye'ye girer ve Edirne üzerinden Ege Denizi'ne dökülür. Ayrıca, Yunanistan ile Türkiye sınırının bir kısmını oluşturur. Taşkınların büyük çoğunluğunun nedeni Meriç'in Bulgaristan'daki bölümüne kurulmuş olan barajların kapaklarını açmasından kaynaklandığı bilinmektedir. 480 km uzunluğa sahip bu nehrin herhangi bir bölgede taşkınlara sebep olmaması ya da oluşan taşkınlardaki zararın en aza indirilmesi için ihtiyaç duyulan, üç ülkenin UMVA'larının kurulmuş ve birbirileri ile iletişim içinde çalışabiliyor olmasıdır.

Devletlerin sürdürülebilir gelişiminin sağlanmasında UMVA önemli rol oynar. Masser [31] UMVA'nın faydalarını ekonomik, sosyal ve çevresel olmak üzere üç ana başlıkta toplamıştır.

Ekonomik faydaları:

- Mekânsal bilgilerin ürün ve hizmetlerinde genişleyen bir iç pazar,
- Mekânsal bilgilerin ürün ve hizmetlerinin ihracatında daha fazla fırsat ve rekabet,
- Hem kamu hem özel sektör kurumlarında artan verim,
- Gelişmiş altyapı ve nakil yönetim sistemi,
- İş dünyasında mekânsal bilgi uygulama ve hizmetleri için yeni fırsatlar oluşturmaktadır.

Sosyal faydaları:

- Gelişmiş ulusal ve yerel hükümet,
- Daha etkili devlet güvenliği,
- Demokratik süreci yaşatmak için daha fazla olanak,
- Sağlık, eğitim gibi bölge ve toplulukların ihtiyaçlarını karşılamadaki olanaklar,
- Daha hızlı acil durum müdahalenin oluşturulmasıdır.

Çevresel faydaları:

- Gelişmiş kıyı bölgesi yönetimi,
- Sürdürülebilir gelişimi desteklemek,
- Doğal kaynakların daha iyi görüntülenme ve yönetilmesidir.

Teknolojik ilerlemeler, konum bilgisinin önemi konusunda hükümetler ve toplumlar üzerinde farkındalık oluşturmıştır [31]. Ekonomik, sosyal ve çevresel konularda sürdürülebilir ilerlemenin gerçekleştirilmesi için UMVA ya ihtiyaç duyduklarını anlamışlardır. Bugün ABD’de National Spatial Data Infrastructure (NSDI), Canada’da Canadian Geospatial Data Infrastructure (CGDI), Avusturalya’da Australian Spatial Data Infrastructure (ASDI), Almanya’da Geodateninfrastruktur Deutschland (GDI-DE), Çin’de Digital China Geospatial Framework (DCGF) adıyla UMVA’lar kurulmuştur.

Farklı UMVA örnekleri, sıradaki başlık altında incelenmiştir.

2.3 UMVA Örnekleri

2.3.1 Dünyada UMVA

2.3.1.1 Amerika Birleşik Devletleri (NSDI)

“Mekânsal verilerin edinimi ve erişiminin koordinasyonu: UMVA” başlığıyla 1994 yılında bir Başkan emri yayınlandı. Buna göre UMVA, mekânsal verilerin kamu ve özel kurumlar, kâr amacı gütmeyen kuruluşlar ve akademik toplum arasında paylaşımına imkân veren teknolojiler, politikalar ve insanlar olarak tanımlanır. Bu altyapının amacı, kurumlar arasındaki ikili veri üretimini azaltmak, kaliteyi arttırmak, coğrafi veri üretiminin maliyetini düşürmek, verilerin kamuya daha açık olmasını sağlamak, mevcut verilerin kullanımındaki faydaları arttırmak ve veri kullanımını arttırmak için eyaletler, iller, ilçeler, akademi ve özel sektörle kilit ortaklıkları kurmaktır [32]. Federal Coğrafi Veri Komitesi UMVA’yı üç önemli eylem üzerinden yönetiyor; ulusal boyutta mekânsal veriler için paylaşım ortamları (clearinghouse ve geoportal) kurulması, verilerin dökümantasyonu, toplanması ve değişimi için standartların geliştirilmesi ve ulusal sayısal mekânsal veri üretmek için ortaklıkların, yöntemlerin ve politikaların geliştirilmesi [33].

FGDC, UMVA’nın altı adet yapı taşını uygulamakla uğraşır. Bunlar; metaveri, paylaşım ortamları, standartlar, framework (çerçeve), mekânsal veri, ortaklıklardır. Tüm bunlar mekânsal verinin dökümantasyonunda yapı ve uyumu oluşturmayı sağlar. Ayrıca mekânsal veri üretici ve kullanıcılarının dağıtık ağlar kurarak verileri paylaşabilmelerine olanak tanır. Temel veri temaları olarak; ortofoto, jeodezik kontrol, yükseklik ve batimetri, ulaşım, hidrografiya, kadastro ve idari birimleri kullanılır [34].

FGDC mekânsal verilerin geoplatform.gov portalından paylaşılmasını koordine etmektedir [35]. Geospatial One-Stop İnisiyatifi tarafından desteklenen portal, ABD içindeki her türlü kayıtlı mekânsal bilgiye ve çevrimiçi hizmetlere erişilebilmesini sağlamaya çalışır.

Standartlar, FGDC tarafından uluslararası eğilimlere (ISO TC 211, OGC, W3C) bağlı kalınarak eyalet, il yönetimleri, özel sektör ve akademi dünyası ile işbirliği içinde belirlenmiştir.

2.3.1.2 Avustralya (ASDI)

ASDI mekânsal verinin sağlayıcı ve kullanıcılarını birbirine bağlayan ulusal bir yapıdır. ASDI, mekânsal verinin tüm kamu kurumlarında, özel sektörde, kâr amacı gütmeyen organizasyonlarda ve akademide kullanımına imkân vermek için insanı, teknolojiyi ve politikaları kapsayan bir yapıdır.

Temel veri temaları olarak; idari sınırlar, kadastro, yükseklik, arazi kullanımı, yer adları, yollar, cadde adresleri, bitki örtüsü ve suyu kullanırlar.

1995'te kurulan Metaveri Çalışma Grubu'nun görevi ulusal boyutta mekânsal veri dizini oluşturmaktır. Avustralya Mekânsal Veri Dizini bütün Avustralya'daki mevcut mekânsal verileri keşfetmeye olanak tanıyan çevrimiçi bir dizindir [36].

ISO19115 ve OGC katalog servisi gibi uluslararası standartları temel almıştır. ASDI, dönüşümünü giderek web servisi modeline doğru sürdürmektedir.

2.3.1.3 Güney Afrika Cumhuriyeti (NSIF)

Ulusal Mekânsal Bilgi Çerçevesi (NSIF) karar verme mekanizmasındaki süreçlerde, mekânsal bilginin kullanımını desteklemek için ihtiyaç duyulan altyapının geliştirilmesini sağlayan ulusal bir inisiyatiftir. Dünyadaki diğer örneklerinde olduğu gibi; politikaları, kurumsal düzenlemeleri, gelişen insan kaynakları ve coğrafi bilgi standartlarını içerir. Mekânsal Veri Keşif Hizmeti, verilerin değişimine olanak tanımak için mekânsal veri kaynakları ile kullanıcıları birbirine bağlayan bir sistemdir. Böylece tekrarlı (dublication) veri üretimine engel olunmaya çalışılır [34].

Temel veri temaları; kadastral sınırlar, arazi örtüsü, hidrografiya, yükseklik (DEM), idari sınırlar ve sayısal ortofotolardır. Ancak bunların hangi standartlarda üretildiği gerek web sayfalarından gerekse dökümanlardan tespit edilememiştir.

2.3.2 Avrupa'da UMVA ve INSPIRE Direktifi

2001'de Avrupa Birliği Komisyonu çevresel ve mekânsal verisetleri ile ilgili çalışmaları başlatmıştır. 15 Mayıs 2007 tarihinde kısaca INSPIRE (Infrastructure for Spatial Information in Europe) ismiyle anılan mekânsal bilgi altyapısı direktifi yürürlüğe girmiştir [37]. INSPIRE direktifi, Avrupa içinde mekânsal veri altyapısını oluşturmayı

amaçlamaktadır. Oluşturulacak altyapı Avrupa içinde çevreyle ilgili mekânsal verilerin kamu kurumları arasında daha hızlı paylaşılmasına ve mekânsal bilgilere hem kamusal hem de özel erişimin daha kolay hale gelmesine olanak tanıyacaktır. Ülkemizde bu amacı gerçekleştirmek için kurulan Çevre ve Şehircilik Bakanlığı'na bağlı Coğrafi Bilgi Sistemleri Genel Müdürlüğü'nün web sitesinde de yer aldığı gibi "INSPIRE, sınırları aşan bir politikanın oluşturulmasına yardımcı olacaktır" [38].

INSPIRE bir dizi ortak ilkeye dayanmaktadır [38]:

- Veri yalnız bir kez toplanmalı ve en etkin şekilde saklanacağı yerde tutulmalıdır.
- Avrupa içindeki farklı kaynaklardan gelen muntazam verinin bir araya getirilmesine, çok sayıda kullanıcı tarafından uygulamalarda paylaşılmasına olanak sağlamalıdır.
- Belirli bir katman ve ölçekte toplanan veri, diğer tüm katman ve ölçeklerde paylaşılabilir, araştırmalar için ayrıntılı olmalı, stratejik amaçlar için genel özellikler içermelidir.
- İyi bir yönetim için her düzeyde ihtiyaç duyulan coğrafi bilgiye kolaylıkla ve şeffaf bir şekilde erişilebilmelidir.
- Hangi coğrafi bilgilere ulaşılabilir, istenilen amacı karşılamak için nasıl kullanılacağı ve hangi şartlar altında elde edilebileceği ve istifade edilebileceği bilgisine kolaylıkla ulaşılabilir.

INSPIRE'yi benimseyen ve veri altyapılarını buna göre hazırlayan ülkelerin, UMVA çalışmalarında INSPIRE'nin standartlarına uyması gerekir. Örneğin, detay sınıfı, öznitelik ve tanımlamalarının ISO 191XX serisi standartlarına uygun yapılması gerekmektedir [39].

INSPIRE direktifinde veri temalarından kurumların koordinasyonuna kadar çok sayıda protokol ve standart mevcuttur. Ek-1 (Annex 1), Ek-2 (Annex 2) ve Ek-3 (Annex 3) gibi ana başlıklarda çeşitli veri temaları bulunmaktadır. Ek-1'deki veri temaları; adres, idari birimler, kadastral parseller, koordinat referans sistemleri, coğrafi grid sistemleri, coğrafi isimler, hidrografiya, koruma bölgeleri, ulaşım ağlarıdır. Ek-2; yükseklik, jeoloji, arazi örtüsü ve ortofotodan oluşmaktadır. Ek-3 ise; toprak, canlı türleri dağılımı, deniz bölgeleri, arazi kullanımı gibi 21 tematik veri temasından oluşmaktadır (Şekil 2.1) [40].

ANNEX 1

 Addresses	 Geographical names
 Administrative units	 Hydrography
 Cadastral parcels	 Protected sites
 Coordinate reference systems	 Transport networks
 Geographical grid systems	

ANNEX 2

 Elevation
 Geology
 Land cover
 Orthoimagery

ANNEX 3

 Agricultural and aquaculture facilities	 Habitats and biotopes	 Population distribution and demography
 Area management / restriction / regulation zones & reporting units	 Human health and safety	 Production and industrial facilities
 Atmospheric conditions	 Land use	 Sea regions
 Bio-geographical regions	 Meteorological geographical features	 Soil
 Buildings	 Mineral Resources	 Species distribution
 Energy Resources	 Natural risk zones	 Statistical units
 Environmental monitoring Facilities	 Oceanographic geographical features	 Utility and governmental services

Şekil 2.1 INSPIRE veri temaları [40]

2.3.3 Ülkemizde UMVA

Ulusal düzeyde, UMVA çalışmaları, Çevre ve Şehircilik Bakanlığı'na bağlı CBS Genel Müdürlüğü tarafından yürütülmektedir. CBS Genel Müdürlüğü'nün INSPIRE Direktifine uygun olarak üzerinde çalıştığı Türkiye Ulusal Coğrafi Bilgi Sistemleri (TUCBS), UMVA oluşturmada birincil öneme sahiptir.

TUCBS'nin birlikte çalışabilirlik için standartların belirlenmesi, standartlara ve veri temalarına göre veri ve metaveri üretimi, sunumu ve üçüncü kişilerle paylaşımı gibi aşamaları bulunmaktadır. Tüm bu aşamaların INSPIRE Direktifi gözetilerek gerçekleştirilmesi gerekmektedir. "TUCBS Standartlarının Belirlenmesi Projesinin" son raporları, yüklenici TÜRKSAT A.Ş. ile alt yüklenici İTÜ Arı Teknokent Proje Geliştirme Planlama A.Ş. tarafından Aralık 2012'de teslim edilmiştir [41]. Temel veri temaları Ek-1'de; Adres, Bina, Tapu-Kadastro, İdari Birim, Ulaşım, Hidrografiya, Arazi Örtüsü, Ortofoto, Topografiya ve Jeodezik Altyapı olarak belirlenmiştir. Tematik veri temaları ise Ek-2'de; Yasak/Koruma Bölgeleri, Plan Bölgeleri, Sosyal/Kültür, Altyapı, Doğal Kaynaklar, Biyoçeşitlilik, Hava/İklim ve Jeoloji/Çevre olarak belirlenmiştir.

Büyük ölçekli haritaların üretilmesi ve paylaşılmasında Çevre ve Şehircilik Bakanlığı Tapu ve Kadastro Genel Müdürlüğü ve İller Bankası A.Ş., orta ve küçük ölçekli haritaların üretilmesi ve paylaşılmasında ise Milli Savunma Bakanlığı Harita Genel Komutanlığı (HGK) önemli rol oynamaktadır. Şu anda Çevre ve Şehircilik Bakanlığı, Orman ve Su İşleri Bakanlığı ve TÜRKSAT A.Ş. çeşitli faaliyet alanlarında geoportallar (TAKBIS, ATLAS, GEODATA ve TURKSAT GLOBE) sunmaktadır. Ancak bu portallar genel olarak lokal çalışmaktadır. Henüz aktif olarak kullanıma açık değildir.

2010 yılı INSPIRE Türkiye raporunda; “Mevcut durumda, HGK, veriyi oluşturan başlıca kurum olup, sınırlı sayıda metaveriyi sunan en iyi görüntülenen internet ortamına sahiptir. Oluşturulan coğrafi bilgi kullanımında izin gerekliliği, kurumun askeri özelliğini ortaya koymaktadır.” ifadesine yer verilip sivilleşmenin gereğine değinilmiştir [42]. HGK’nın sivilleşme süreci henüz başlamamıştır.

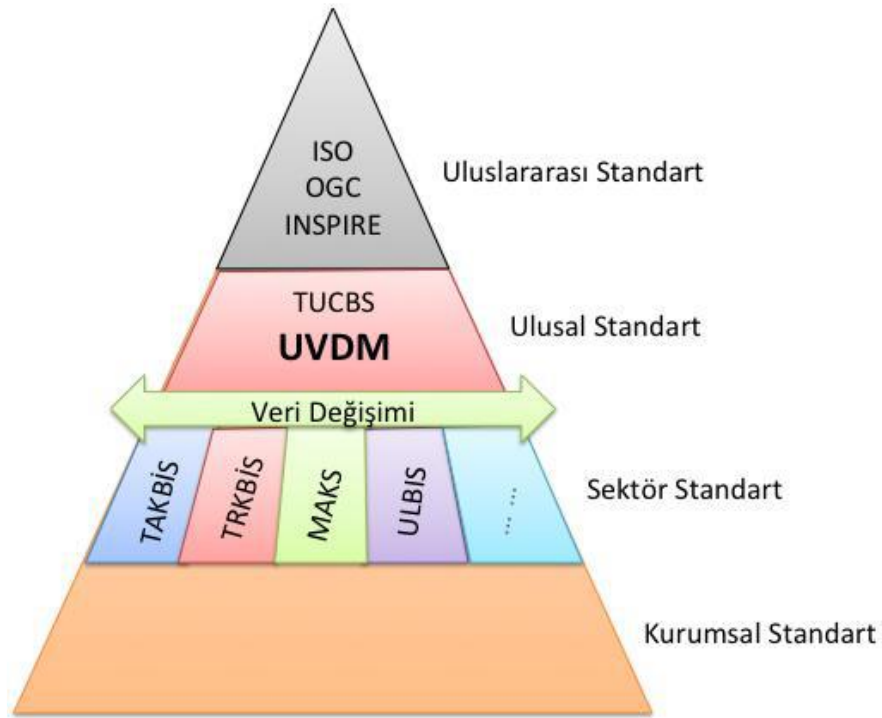
Ülkemizde TUCBS’nin gerçekleştirilme sürecindeki yavaşlığını Güney vd. [41] aşağıdaki nedenlere bağlamıştır.

- Bugüne kadar yapılan çalışmaların çoğunda üstlenilen rol önderlik değil diğer ülkelerdeki mekânsal bilgi yönetimini takip etmek olmuştur.
- Çağın gerektirdiği şekilde, mekânsal bilgi konusunda yeterli “farkındalık” oluşturulmuş ancak “değişim” ve “sonuca en hızlı ulaşma” konusunda adımlar atılmamıştır.
- Son 15 yılda TAKBIS projesinin bitmemesi ve mekânsal bilgi altyapısı oluşturmadaki süreç göz önünde bulundurulursa, sürdürülebilir biçimde bir mekânsal birlikte çalışabilirlik inşasının yakın gelecekte tamamlanmasının olası olmadığı görülmektedir.
- TUCBS projesi genel olarak farklı kaynaklardan elde edilen verilerin, metaveriler sayesinde paylaşılmasına dayalı bir portal yaratmak ve tekrarlı veri üretimini engellemek üzerine kurulmuştur. Bunlar gerekli ancak yeterli değildir.

Standartlar

Mekânsal verilerin farklı disiplinler, bilgisayarlar ve kullanıcılar tarafından anlaşılmasını ve doğru yorumlanmasını sağlamak için veri sınıflandırmasının ve veri yapılarına ait detayların belirli ilkeler doğrultusunda standartlaştırılması gerekmektedir [43]. TUCBS

için temel bir detay modelinin tasarlanması ve gerçekleştirilecek tüm coğrafi veri modellerinin bu detay modeline uygun olması sağlanmalıdır. Aydınöğlu vd. [43]'te "TUCBS Mekânsal Veri Altyapısının Kurulması Projesi"'nde veri modellerinin özelliklerinin tanımlanmasında ve tutarlılığın sağlanmasında ISO 19109'da tanımlanan Genel Detay Modeli (General Feature Model) standardının temel alınacağını belirtmişlerdir. Şekil 2.2'deki piramitte TUCBS standart hiyerarşisi gösterilmektedir. Bu piramit, en alt bölümde yer alan kurumsal standartlardan Tapu ve Kadastro Bilgi Sistemi (TAKBİS) ve Türkiye Kent Bilgi Sistemi (TRKBİS) gibi sektörel standartlara, orta bölümde yer alan Ulusal Veri Değişim Modeli ve TUCBS gibi ulusal standartlardan ISO, OGC ve INSPIRE gibi uluslararası standartlara kadar hiyerarşik yapıyı göstermektedir. Düşey ilişki ile veri modelleri arasında veri değişimi mümkün olmalı ve farklı hiyerarşideki veri grupları ve modelleri arasında da ilişki olmalıdır [43].



Şekil 2.2 Standart hiyerarşisi [43]

2.4 Değerlendirme

UMVA; uyumsuz mekânsal bilgi, tekrarlı veri üretimi, veri sağlama ve paylaşım zorlukları ile karşılaşılması için önceden belirlenmiş kurallara göre tüm mekânsal veri sürecini yönetir. UMVA'nın başarısı mekânsal verilerin uyumlu, metaverilerin tanımlanmış şekilde standartlara, mevzuata, politikaya vb. uygun olmasıyla yakından ilgilidir. Birlikte

alışabilme, uyumlu mekânsal bilgilerin hızlı bir şekilde yorumlandığı bir sistem becerisidir. Birbirini tanıyan sistemler farklı kaynaklardan gelen verilerin kullanımı sırasında, kullanıcıların uyum nedenli karşılaştıkları sorunları giderir ve verileri uyumlu hale getirirler.

Bu bölümde UMVA kavramının tarihsel olarak ortaya çıkışı, ihtiyaç duyulma nedenleri ve dünyadaki mevcut durumu hakkında genel bilgiler verilmiştir. Dört kıtadan, ülkemiz dâhil beş UMVA örneğı; genel yapısı, kapsamı, veri temaları ve standartları yönünden incelenmiştir. Tüm ülkeler tarafından ISO temel standart olarak kabul edilmiştir. Bununla birlikte, sektörün daha içinde olduğu için, OGC standartları da kabul edilmiştir. Ayrıca, genel olarak UMVA’larda veri temalarının; idari sınırlar, kadastro, yükseklik, hidrografya, ulaşım ve topografya olduğu görülmüştür.

Sonuç olarak bulgular, gelişen ülkelerden Güney Afrika Cumhuriyeti ile Türkiye’nin UMVA oluşturmada hemen hemen aynı seviyede olduğunu ancak ABD’den hem yasal hem teknolojik hem de yetişmiş insan bileşenleri bakımından geri olduğunu ortaya koymaktadır. Türkiye’nin “UMVA reçetesi” INSPIRE direktifi uygulamaya hazır beklemektedir. İş, ülkemize uyarlanması ve eksik olan en büyük kısmın yani yasal altlığın oluşturulmasıdır. Ayrıca gerek bu genel çerçevenin yasal altlığında gerekse kurumların kendi yönetmeliklerinde birbirine uyumlu organizasyon hiyerarşisi yaratılmalıdır. Son olarak, harita üretiminde önemli role sahip olan HGK’nın sivilleşerek, tıpkı Avrupa ve dünyadaki diğer örneklerde olduğu gibi, mekânsal veri pazarında daha etkin rol alması gerekmektedir.

MEKÂNSAL VERİLERİN ENTEGRASYONU

3.1 Kapsam

Bilgi teknolojilerinin kullanımının artmasıyla birlikte, birbirinden farklı ölçekli, amaçlı, zamanlı, izdüşüm yüzeyli, doğruluklu, öznelikli verisetlerinin çeşitli amaçlar doğrultusunda birlikte kullanımını gerektiren durumlar oluşmuştur. Sürekli güncellenmesi gereken navigasyon verileri, çevresel değişimlerin izlenmesi ve analiz edilmesi, numarataj verilerinin farklı geometrik yapıdaki yol veya parsel verisetlerine aktarımı gibi konular ve sel, deprem, yangın, terör saldırısı gibi acil müdahale gerektiren durumlar, mekânsal verilerin birlikte kullanımını ve entegrasyonunu gerektirmiştir.

Mekânsal veri entegrasyonu; güncelleme (update), bütünleştirme (vertical conflation), kenarlaştırma (horizontal conflation) gibi mevcut verisetleri arası veri aktarımının/değişiminin gerçekleştirilerek, belirlenmiş amaçlar çerçevesinde, yeni veri veya ürünlerin elde edildiği süreçler olarak tanımlanabilir. Oluşturulacak entegrasyon stratejilerinin başarısı birinci dereceden verilerin birbirileri ile uyumlu olmasına ve entegrasyon sürecinin otomasyonuna bağlıdır.

Veri kaynaklarını birlikte kullanmak, birbirine entegre olabilen sistemleri kurmak ve verisetleri arasında veri aktarımını gerçekleştirmek için aynı varlığı temsil eden detayların eşlenmesi gereklidir. Eşleme, mekânsal veri entegrasyonunda bir işlemdir ve geometri, topoloji ve öznelilik kriterlerle belirlenen yöntemler ile çözülmektedir. Eşleme işlemi entegrasyon süreçlerindeki önemli adımlardandır.

Bu bölümde etkili entegrasyon için UMVA gerekliliği, entegrasyon çeşitleri ve bu süreçlerde en çok karşılaşılan sorunlardan biri olan eşleme işlemi üzerinde durulmuştur.

3.2 UMVA Kapsamında Veri Entegrasyonu

Sürdürülebilir modern toplum inşası için karar alıcıların bilimsel olarak üretilen verilere göre politikalar üretmesi gerekir. Örneğin, doğal ve beşeri nedenlerle oluşan çevresel değişimleri gösteren birden fazla kaynaktan toplanan mekânsal veriler, yönetimler tarafından; çevre sorunları ve hızlı kentleşme gibi sorunların çözümünde kullanılabilir. Sel, deprem, yangın, terör saldırısı gibi acil müdahale gerektiren durumlarda da ihtiyaç duyulan veriler genellikle birden fazla kaynaktan gelmektedir. Ancak bu gibi durumlarda mekânsal verilere normalde olduğundan daha hızlı ihtiyaç duyulur. Bir yangına müdahale sırasında, müdahale araçlarının geçebileceği genişlikte yolların bilgisinden binaya ait kat planlarına kadar birçok farklı mekânsal veriye saniyeler içinde ihtiyaç duyulabilir. Sel ve taşkın gibi afetlerde ise ilk etapta hidrografiya ve yerleşim planları gibi veriler, afet sonrası da bu verilere ek olarak kadastro verisi kullanılabilir.

Yönetimler tarafından beşeri ve doğal değişimlerin kontrolünün sağlanması, afet gibi acil müdahale gerektiren olayların yönetilmesi için stratejilerin belirlenmesi, mekânsal verilerin ortak sistemlere entegrasyonu ile mümkün olmaktadır. Yukarıdaki durumlarda olduğu gibi birçok hizmet sürecinde farklı kaynaktan gelen farklı veriler eş zamanlı olarak kullanılır. Ortak platformlarda değerlendirilecek verilerin birbirileri ile uyumsuz olması yapılacak işin süresini uzatır. Bazı durumlarda ise yeterli zaman ayrılmasına rağmen verilerin birbirileri ile çalışması mümkün olmamaktadır. Böyle durumlarda verilerin bir kısmı ya da tamamı ya tekrar düzenlenmekte ya da üretilmektedir. Bu da söz konusu iş için hem ekonomik hem de zamansal yük getirmektedir. Doğru ve standartlara uygun sistemler sayesinde hızlı, ekonomik, kolay ve kaliteli hizmetin önü açılır.

Verilerin bir bilgisayarı oluşturan donanım parçaları gibi birbiri ile çalışmayı desteklemesi gerekir. Standartlaştırılmış veritabanları INSPIRE verilerinin tüm masaüstü ve çevrimiçi platformlarda uyumlu olmasını sağlar. Ancak şu an UMVA'da karşılaşılan büyük sorunlardan biri; mevcut verilerin nasıl UMVA standartlarındaki verilere dönüştürüleceğidir. FME ve ArcGIS yazılımları mevcut verilerden INSPIRE uyumlu veri modellerine geçiş için şema dönüşümünü gerçekleştirebilmektedir [44].

Verilerin metaverilerle zenginleştirilmesi; kapsamı, zamanı, nasıl üretildiği ve veri yapısı gibi bilgiler entegrasyonun hızını, kalitesini ve uygulanabilirliğini etkileyen başlıca etkenlerdir. Etkili entegrasyon için UMVA'lar kurumsal, sosyal, resmi yapılarla ve yönetimlerin politikalarıyla desteklenmelidir [45].

3.3 Mekânsal Veri Entegrasyonunun Çeşitleri

Mekânsal veriler üretildikleri kaynağa göre; zaman, yapı, hassasiyet, öznitelik, çözünürlük, vb. bakımından farklı olabilirler. Fakat birlikte çalışabildikleri takdirde sözdizimsel, öznitelik ve geometrik entegrasyon süreçlerinde kullanılabilirler [46].

Geometri ve öznitelik entegrasyonundan tam olarak ayrılmayan sözdizimsel entegrasyon; XML (Extensible Markup Language), GML (The Geography Markup Language) gibi işaretleme dillerinde dosya yapısı ile oynar. Bazı durumlarda içerisinde hem öznitelik hem geometri entegrasyonu gerçekleştirilir.

Metaveriler orijinal veya değiştirilmiş veri hakkında detaylı bilgiler (nerede, ne zaman, ne, detay türü, vb.) sakladığı için çoğu entegrasyon sürecinde ihtiyaç duyulur. Güncelleme ve bütünleştirme işlemlerinde verisetlerinin şematik bilgilerinin tamlığı, gerçekleştirilecek eşleme veya entegrasyon yöntemlerinin belirlenmesine yardım eder.

Geometri/öznitelik entegrasyonu, sırasıyla geometrik/öznitelik bakımdan verisetlerinin daha güncel, doğru ve ihtiyaç duyulan veri bakımından zengin olması için yapılır. Üst üste veya bitişik şekilde birebir çakışan geometriye veya ortak özniteliklere sahip verisetleri çoğu CBS yazılımı kullanılarak detay-detay eşlenebilir ve bu bağlantılar yardımıyla birinden diğerine geometri/öznitelik aktarımı otomatik olarak gerçekleştirilebilir. Ancak tam olarak çakışmayan ve birebir aynı öznitelik sütununa sahip olmayan verisetlerinin entegrasyonu için daha detaylı eşleme algoritmaları kullanmak gerekir.

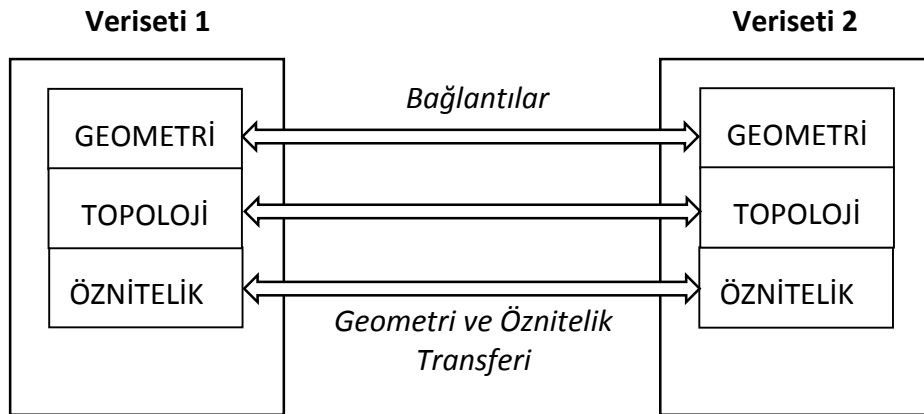
Sözdizimsel entegrasyon bu tezin kapsamı dışındadır. Tezin kalan kısmında eşleme süreçleri araştırılmıştır.

3.4 Mekânsal Verilerin Eşlenmesi

Veri üretim mekanizmasının önceden belirlenmiş bir plana veya kabul edilmiş standartlara göre gerçekleştirilmemesinden dolayı, veriler arası öznitelik veya geometri

transferi, güncelleme, bütünleştirme, veri entegrasyonu, eşleme gibi uygulamalar otomatik olarak yapılamamaktadır. Entegrasyon süreçlerinde başarının en kritik ölçütü olan eşleme işlemlerinin otomatik yapılamaması zaman ve ekonomik açıdan büyük kayıplara neden olmaktadır.

Eşleme işlemi genel olarak iki verisetinin detaylarının birbirlerine bağlantılar (linkler) ile bağlanması olayıdır. Amaç, geometrik, öznitelik ve/veya topolojik olarak eşlenen detaylar arasında geometri ve/veya öznitelik aktarımının gerçekleştirilmesidir. Şekil 3.1’de eşleme çeşitleri ve entegrasyon süreci, verisetleri çerçevesinde gösterilmektedir. Veriseti 1’de yer alan detaylar, verisetlerinin karmaşık yapısına göre, geometrik, öznitelik veya topolojik kriterlerden biri, ikisi veya üçü birden kullanılarak Veriseti 2’deki detaylarla eşlenebilir. Eşleme algoritmalarında kullanılan bu kriterlerden biri diğerlerinden daha belirleyici rol oynar. Geometri, topoloji veya öznitelik eşlemeleri karmaşık eşleme süreçlerinde aynı benzerlik eşitliklerinin bileşenleri olabilmektedirler [17], [18].



Şekil 3.1 Mekânsal veri entegrasyonu: eşleme ve veri transferi

3.4.1 Öznitelik Eşlemesi

Farklı verisetlerinde aynı gerçek varlıkları temsil eden nesnelerin çoğu zaman öznitelikleri birbirine benzer olmaktadır. Bu sebeple iki veriseti arasında eşleme yapılırken öznitelik verileri de benzerlik etkenlerinden biri olarak kullanılabilir.

Nesne ismi eşleme işlemlerinde sık kullanılan bir özniteliktir. İki isim arasındaki benzerliğin belirlenmesi için çeşitli yöntemler mevcuttur. “Hamming Mesafesi” ve “Levenshtein Mesafesi” iki nesne isminin birbirine ne kadar benzemediğini

hesaplayabilmektedir. Hamming mesafesine göre nesneye ait ismi oluşturan harfler sırasıyla diğer verisetindeki aynı uzunluktaki nesne isminin aynı sıradaki harfleri ile karşılaştırılır [22]. Bu yöntem, karşılaştırılan harflerin farklı olması durumunda her farklılık için +1 değerlerinin toplamını Hamming mesafesi olarak kabul eder. Çizelge 3.1’de harf ve sayı örnekleri verilmiştir. “Harbiye” ile “Raptiye” isimleri arasındaki Hamming mesafesi hesaplanırken “Harbiye” kelimesindeki sırasıyla “h, r, t” harfleri yerine “Raptiye” kelimesinde “r, p, t” harfleri bulunmaktadır. Bu sebeple iki kelime arasındaki Hamming mesafesi 3’tür. Geriye kalan harfler iki kelimedede de aynıdır. Sayılar için ise aynı şekilde “5, 7” yerine “0, 6” bulunmaktadır. Hamming mesafesi 2’dir.

Çizelge 3.1 Hamming mesafesi örnekleri

	Veriseti 1	Veriseti 2	Hamming Mesafesi
İsim	Harbiye	Raptiye	3
Sayı	110575	110065	2

Levenshtein mesafesi ise başta arama motorları olmak üzere bilgi teknolojilerinde sık kullanılan bir yöntemdir. Bir kelime veya cümlemin başka bir kelime veya cümleye dönüşmesi için yapılması gereken en az sayıdaki işlem adımlarının sayısını verir [18]. Levenshtein mesafesinin hesaplanması için kullanılan örnek matrisler aşağıda Şekil 3.2’de verilmiştir. Buna göre, A matrisinde “Kitap” isminin “Kalem” ismine dönüşmesi için 4 işlemin, B matrisinde ise “Tatil” ve “Sahil” isimlerinin birbirine dönüşmesi için ise 2 işlemin yapılması gerektiği anlaşılmaktadır.

A matrisi

	K	İ	T	A	P
K	0	1	2	3	4
A	1	1	2	3	4
L	2	2	2	3	4
E	3	3	3	3	4
M	4	4	4	4	4

B matrisi

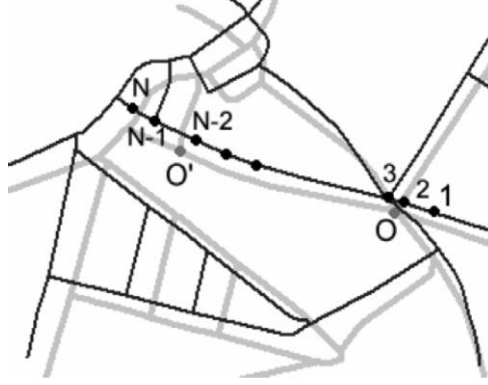
	T	A	T	İ	L
S	1	2	3	4	5
A	2	1	2	3	4
H	3	2	2	3	4
İ	4	3	3	2	3
L	5	4	4	3	2

Şekil 3.2 Levenshtein mesafesi için örnek matrisler

Hamming veya Levenshtein mesafeleri eşleme süreçlerinde genellikle tek başlarına kullanılmazlar. Daha çok geometri ve topoloji eşlemelerinde detaylar arasında kurulmak istenen benzerlik bağıntılarının iyileştirilmesi için kullanılırlar.

3.4.2 Topoloji Eşlemesi

Topoloji, bağlantılı ve komşu nesneler arasında ilişki kurmaya yaradığı için aynı mekâna ait farklı verisetlerinin nesneleri arasında eşleme için kullanılır. Aşağıda Şekil 3.3’de topoloji dikkate alınmadığı zaman O ve O’ noktaları arasında kalan çizgi diğer verisetindeki 1 ve N-2 noktaları arasında kalan çizgi ile eşlenebilir. Fakat topoloji dikkate alındığında 3 ve N-1 noktaları arasındaki çizgi ile eşlenecektir.

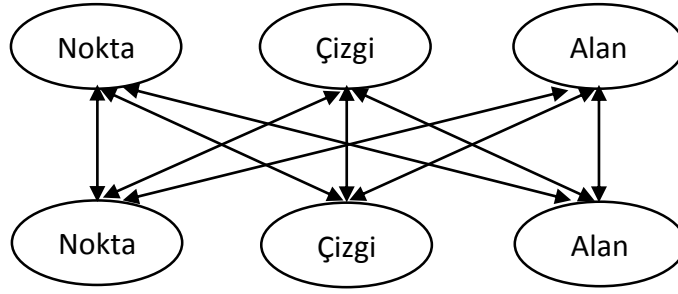


Şekil 3.3 Topolojiyi kullanarak eşleme (Zhang ve Meng [16])

Zhang ve Meng [16] eşleme için yaptığı çalışmada kavşaklara bağlanan yol sayılarını topolojik bir kriter olarak kullanmıştır. Örneğin Şekil 3.3’deki O noktasına bağlanan çizgi sayısı 5 iken O’ noktasına bağlanan çizgi sayısı 3’tür. Böylece, O’ ve O noktaları arasındaki çizgi diğer verisetinde bu kriterlere uyan noktalar (3 ve N-1) arasındaki çizgi ile eşlenir.

3.4.3 Geometri Eşlemesi

Gerçek dünyadaki varlıklar farklı verisetlerinde farklı geometriler ile temsil edilebilir. Geometri eşleme sürecinde ilk olarak detay türü dikkate alınır. Geometri türüne göre eşleme işlemi nokta, çizgi ve alan şeklinde yapılır (Şekil 3.4) [11]. Li [22], eşleme sürecinde detay türlerinin birbirileri ile olan ilişkisini Çizelge 3.2’deki gibi örneklendirmiştir.



Şekil 3.4 Detay türüne göre eşleme işlemi (Yuan ve Tao [11])

Çizelge 3.2 Detay türüne göre eşleme örnekleri (Li [22])

Detay Türü	Nokta	Çizgi	Alan
Nokta	İşaretler, kavşaklar, vb.	Yol orta eksenleri ile üzerinde ölçülen GPS noktaları, vb.	Bir verisetinde nokta olarak temsil edilirken diğer verisetinde alan olarak temsil edilen restoranlar, vb.
Çizgi		Yol ağları, dereler, vb.	Bir verisetinde çizgi olarak temsil edilirken diğer verisetinde alan olarak temsil edilen nehirler, vb.
Alan			İki kadastral parsel, vb.

Detay türü, eşleme yönteminin seçiminde birinci dereceden etkilidir. Nokta-nokta eşlemesinde kullanılacak eşik mesafe değeri için noktalar arasındaki en kısa mesafe olan Öklid mesafesi (3.1) kullanılabilirken, çizgi-çizgi eşlemesi için çizgiler arasındaki en kısa mesafelerin en uzununu olan Hausdorff mesafesinin (3.2) kullanılması daha uygundur [11], [18]. Hausdorff mesafesi otomatikleştirilmiş kartografya analizleri için nesneler arasındaki mekânsal ilişkiyi nicel veriler olarak sunar. Bütünleştirme süreçlerindeki eşleme işlemlerinde sıkça kullanılmasının yanı sıra genelleştirilmiş çizgilerin orijinal konumundan sapma miktarlarını belirlemek için de kullanılır. Hausdorff mesafesi ve diğer geometrik eşleme mesafeleri ile ilgili detaylı bilgiler Bölüm 4'te verilmiştir.

$$D_{\text{Ö}} = \sqrt{(X_2 - X_1)^2 + (Y_2 - Y_1)^2} \quad (3.1)$$

$$D_H = \max(\min(d_1, d_2)) \quad (3.2)$$

Verisetlerinin yapısal özellikleri eşleme işlemi için geometrik benzerlik kurmada önemlidir. Yol orta eksenleri eşlenirken ağ desen türleri dikkate alınabilir. Grid yapıda yol ağları eşlenirken yeterli olabilecek basit bir yöntem, dairesel veya karmaşık desenli yol ağlarında yeterli olamayabilir. Ayrıca kaynak verisetlerinin ölçekleri eşleme yönteminin belirlenmesinde önemli bir etkidir. İki verisetinde de nokta olarak temsil edilen bir bina için eşleme işlemi yapılırken yeterli olabilecek Öklid mesafesi (3.1), daha büyük ölçekli bir verisetinde alan olarak temsil edilen aynı bina ile eşleme yapıldığında yeterli olamayabilir. Nesneler arası ortalama veya maksimum uzaklık, doğrultular veya yön bilgileri iki veriseti arasında benzerlik kurmak için kullanılabilecek ölçülerdir.

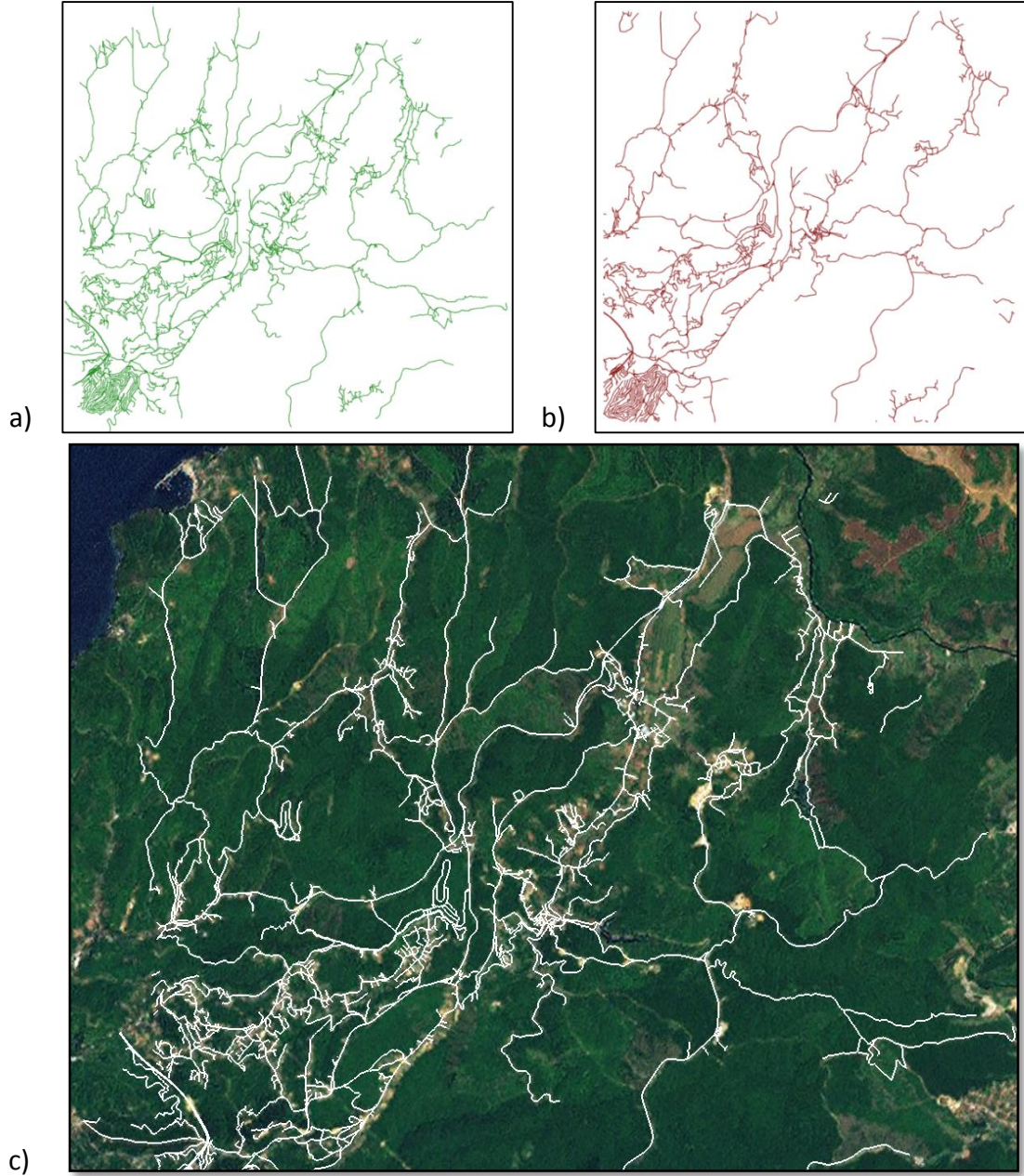
UYGULAMA: YOL ORTA EKSENLERİNİN EŞLENMESİ

4.1 Giriş

Mekânsal verilerin el ile tek tek eşlenmesi (manuel eşleme) zaman ve ekonomik bakımdan kayıplara neden olmaktadır. Bilim insanları eşleme işleminin otomasyonu için birçok algoritma geliştirmiştir. Ancak geliştirilen yöntemlerin çoğu veri bağımlıdır. Bir eşleme işleminde başarılı olan algoritma çoğu kez başka eşleme işleminde başarılı olamamaktadır.

Bu çalışmada, mekânsal veri entegrasyonunun önemli aşamalarından biri olan eşleme sürecinin otomatikleştirilmesi için geliştirilmiş yöntemler test edilmiştir. Uygulamada öncelikle temel (basit) eşleme yöntemlerine sahip MatchingPlugin (Eşleme Eklentisi) [47], ardından Vividsolution firması tarafından geliştirilen RoadMatcher (Yol Eşleyici) [48] ve son olarak Li [22] ile Li ve Goodchild [23] tarafından geliştirilen optimizasyon modeli denenmiştir.

Eşleme uygulaması için, biri İstanbul Büyükşehir Belediyesi Coğrafi Bilgi Sistemi Müdürlüğü (kamu) ve diğeri Başarsoft (özel) şirketinin ürettiği aynı bölgeye ait yol orta eksenlerini içeren iki veriseti kullanılmıştır (Şekil 4.1). Kamu kurumundan elde edilen verilerde bazı yol isimlerinin güncel olmadığı, özel firmadan alınan navigasyon verisinin ise öznitelik bakımından zengin olmasına rağmen geometrik bakımdan güncel olmadığı (bazı yolların bulunmadığı) ve bazı yolların daha basit geometriyle oluşturulduğu görülmüştür. Bu gibi farklılıklar veri entegrasyonunu (geometrik eşleme ve öznitelik transferi) gerekli kılmaktadır.



Şekil 4.1 (a) Kamu ve (b) özel şirketin ürettiği yol orta eksen verileri ile (c) uydu görüntüsü

Bu bölümde, öncelikle, uygulama bölgesi, kullanılan veriler ve yöntemler hakkında ayrıntılı bilgiler verilmiştir. Ardından uygulama sonuçları, tablolar ve grafiklerle verilmiş ve son bölümde sonuçlar değerlendirilmiştir.

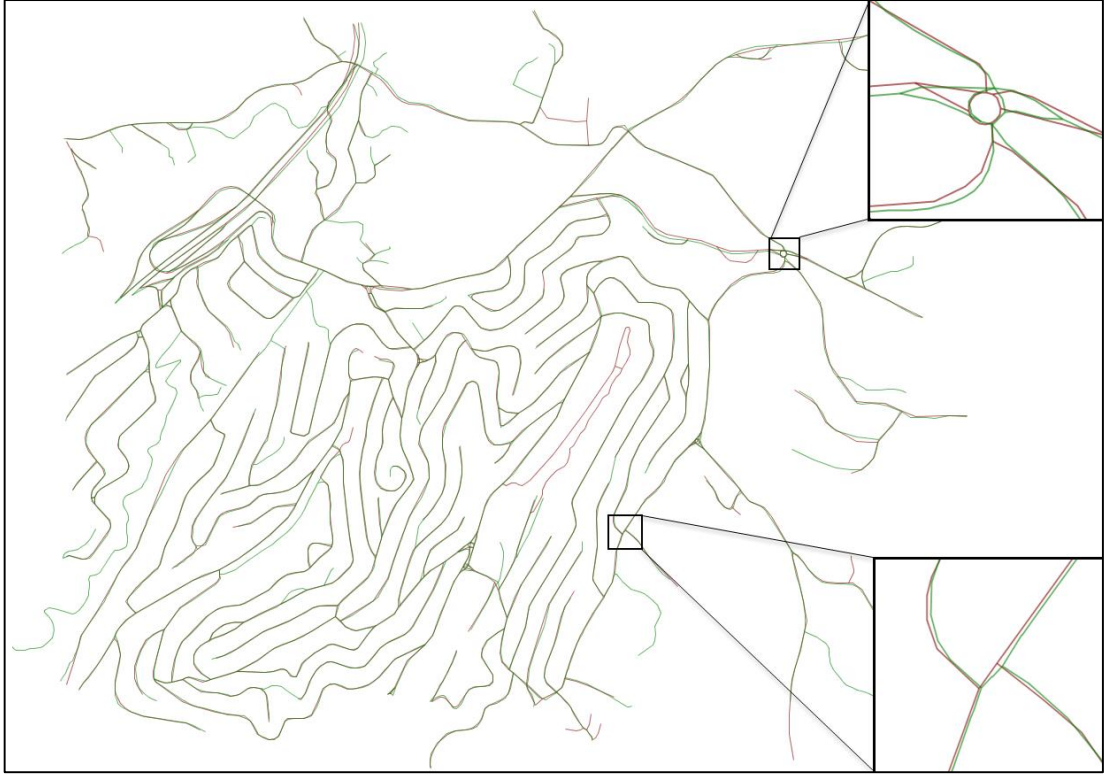
4.1.1 Uygulama Bölgesi ve Kullanılan Veriler

Uygulama, İstanbul ili Beykoz ilçesinin bir bölümüne (10 km x 10 km) ait yol ağları ile gerçekleştirilmiştir. Uygulama bölgesi otomatik eşlemenin doğruluğunun kontrolünü kolaylaştırmak amacıyla üç alt bölgeye ayrılmıştır (Şekil 4.2).



Şekil 4.2 Uygulama alanının alt bölgelere ayrılması

Bu bölümde kamu kurumundan elde edilen veriseti “Veriseti 1”, özel firmadan elde edilen veriseti “Veriseti 2”, alt bölgeler ise Bölge 1, Bölge 2 ve Bölge 3 olarak anılacaktır. Veriseti 1 1:1,000 ölçekli olup kurum tarafından CBS altyapısında kullanılmaktadır. Veriseti 2 1:5000 ölçekli olup firma tarafından navigasyon uygulamalarında kullanılmaktadır. Veriseti 1 ITRF2005.0, Veriseti 2 ise WGS84 datumundadır. Her iki verisetinde geometrik veriler yol orta eksenlerinden meydana gelmektedir. Geometrik verilerde üretim stratejilerinin (üretim yöntemi, zamanı, kaynak, ölçek gibi) farklılığından kaynaklanan konumsal ve topolojik farklar bulunmaktadır (Şekil 4.3). Veriseti 1’de bazı yol isimlerinin güncel olmadığı, Veriseti 2’de ise verilerin öznitelik bakımından zengin olmasına rağmen geometrik bakımdan güncel olmadığı (bazı yolların bulunmadığı) ve bazı yolların daha basit geometriyle oluşturulduğu görülmüştür. Veriseti 1 ve 2’deki öznitelik verileri Şekil 4.4’te görülmektedir.



Şekil 4.3 Bölge 1’de Veriseti 1 (yeşil) ve Veriseti 2 (kırmızı)’nin üst üste görünümü ve yolların birebir çakışmadığını gösteren örnekler

View Schema: Veriseti 1		View Schema: Veriseti 2	
Field Name	Data Type	Field Name	Data Type
GEOMETRY	Geometry	GEOMETRY	Geometry
OBJECTID	Integer	ID	Integer
YOL_ISMI	String	ADI	String
YOL_ID	Integer	TURU	String
YOL_UAVT	Integer	TIPI	String
YOL_TURU	Integer	DF	String
YOL_TIPI	Integer	DR	String
SERIT_SAYI	String	NL	String
TEKYON	String	Z_LEVEL	String
KAPLAMA_CI	String	FZLEVEL	String
BOLUNMUS_Y	Integer	TZLEVEL	String
HIZ	Integer	DS	String
DAKİKA	Double	CS	String
YAYAHIZ	Integer	BP	String
YAYADAKİKA	Double	WD	String
KULLANILAN	String	ANNAME	String
GENISLIK	Double	FC	String
ESKI_ISMI	String	FW	String
PARCA_ID	Integer	YAPI	String
ILCE_UAVT	Integer	TR	String
MAHALLE_KO	Integer	VTR	String
MAHALLE__1	String	HIZLIMIT	String
MAHALLE__2	Integer	ORTALAMAHI	String
ACIKLAMA	String	ILKOD	String
SOZEL_ALAN	String	SOLMAH	String
GLOBALID	String	SAGMAH	String
SHAPE_LEN	Double		

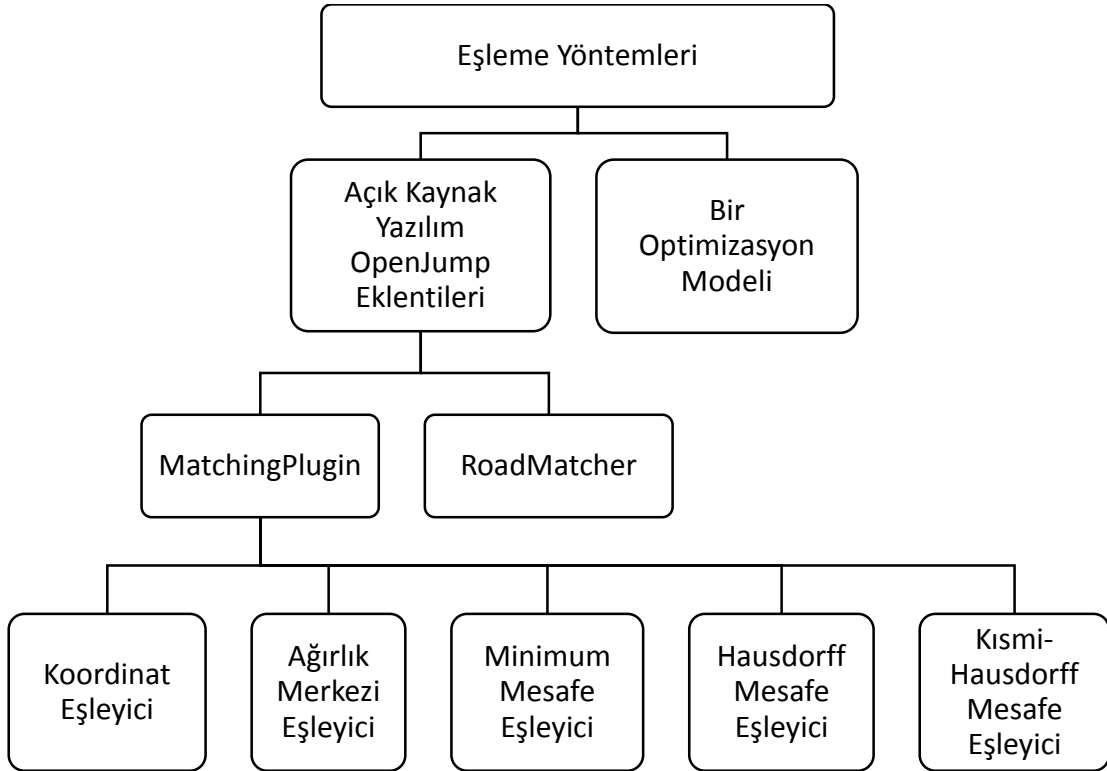
Şekil 4.4 Veriseti 1 (solda) ve Veriseti 2 (sağda) verisetlerinde öznitelik verileri

Bu çalışmada, Veriseti 1'deki çizgi nesnelerin Veriseti 2'deki çizgi nesnelerle eşlenmesi ve verisetlerinin olası bir geometri veya öznitelik entegrasyonuna hazır hale getirilmesi hedeflenmiştir.

4.2 Yöntemler

Uygulamada öncelikle temel (basit) eşleme yöntemlerine sahip Michaël Michaud tarafından geliştirilen MatchingPlugin, ardından Vividsolution firması tarafından geliştirilen RoadMatcher ve son olarak Li [22] ile Li ve Goodchild [23] tarafından geliştirilen optimizasyon modeli kullanılmıştır (Şekil 4.5). MatchingPlugin ve RoadMatcher, Java programlama dilinde yazılmıştır ve her ikisi de OpenJump [49] açık kaynak yazılımında eklenti olarak kullanılır. Optimizasyon modeli ise açık kaynak kodlu

Python dilinde Arcpy modülü kullanılarak yazılmıştır. ArcGIS yazılımında toolbox olarak kullanılır.



Şekil 4.5 Uygulamada kullanılan eşleme yöntemleri

4.2.1 Açık Kaynak Yazılım (OpenJump) Eklentilerinde Eşleme Yöntemleri

OpenJUMP, Vividsolution şirketi tarafından ilk olarak JUMP ismiyle kullanıcılara sunuldu. İsim, Unified Mapping Platform kelimelerinin ve yazım dili olan Java'nın baş harfleri kullanılarak oluşturulmuştur. Open kelimesi ise herkesin erişebilme imkânını, yani açık kaynağı temsil etmektedir. OpenJump, GNU genel kamu lisansına (GPL: General Public Licence) sahiptir ve dünyanın her yerinden programcılar tarafından geliştirilmektedir.

Haritaların üretilmesi, düzenlenmesi, görselleştirilmesi ve coğrafi verilerin analizi için çeşitli araçlara sahip OpenJump yazılımının genel özellikleri aşağıdaki gibidir:

- Vektör tabanlı CBS,
- GNU genel kamu lisansına [50] sahip,
- OGC standartları [51] ile uyumlu,
- Java dili ile yazılmış ve açık kaynaklı,
- Düzenleme ve analiz fonksiyonlarının kullanımı kolay,

- Yeni eklentiler (Plugin) ile yeni özellikler/işlevler kazandırılabilir ve
- İngilizce olan arayüzleri istenen dile (Türkçe, Almanca, vd.) çevrilebilir.

Eklentisiz halinde OpenJump yazılımı yalnızca genel CBS araçlarını (buffer, join, vb.) bünyesinde barındırmaktadır. Bununla birlikte zaman içinde geliştirilen çeşitli eklentilerle yeni işlevler (yazdır, genelleştirme, eşleme vb.) kazandırılmış ve kazandırılmaya devam edilmektedir. Bu bakımdan sürekli gelişen bir yazılımdır. Eklentiler [52] gibi çeşitli internet sitelerinde, dizin şeklinde kullanıcıların hizmetine sunulmaktadır.

Bu çalışmada OpenJump yazılımının X.Y.Z sürümü kullanılmıştır. MatchingPlugin ve RoadMatcher eklentileri ile yol orta eksenleri eşlenmiş ve yeterlilikleri test edilmiştir.

4.2.1.1 MatchingPlugin ile Eşleme

MatchingPlugin (MatchingPlugin0.7.2), birbirinden farklı yöntemler (Koordinata dayalı eşleme, Ağırlık merkezine dayalı eşleme, En kısa mesafeye dayalı eşleme, Hausdorff mesafesine dayalı eşleme, Kısmi-Hausdorff mesafesine dayalı eşleme vd.) ile geometrik ve/veya öznitelik verileri kullanarak nokta, çizgi ve alan nesnelerini eşleme imkânı vermektedir. Eşleme işlemi kaynak veri setinden hedef (referans) verisetine öznitelik transferi ile son bulmaktadır.

MatchingPlugin'in genel özellikleri aşağıdaki gibidir [47]:

- Mekânsal ve/veya öznitelik kriterlere göre kaynak katmandaki nesneler ile hedef (referans) katmandaki nesneleri eşleme
- Kesişim, içermeye gibi katmana dayalı kaplaşım analiz fonksiyonları yanında, minimum mesafe, Hausdorff mesafesi gibi kriterlere de sahip olma
- Ayrıca Levenshtein mesafesi gibi özniteliksel kriterlere sahip olma
- 1:1, 1:m ve n:1 eşlemelerini gerçekleştirebilme
- Öznitelik verilerinin kaynak veri setinden hedef veri setine transferi
- Özniteliklerin eşlenmesinde sadece harf karakterli verilerin kullanımı

Ayrıca eşleme işleminin çıktıları kullanıcının tercihiyle üç ayrı katmanda (eşlenen ve eşlenmeyen nesneler ile bunlar arasındaki bağlantıları (links) gösteren

katmanlar) toplanabilir. Hedef çıktı katmanında, eşlenen nesnelere ait özniteliklerin dışında aşağıdaki öznitelikler de yer alır.

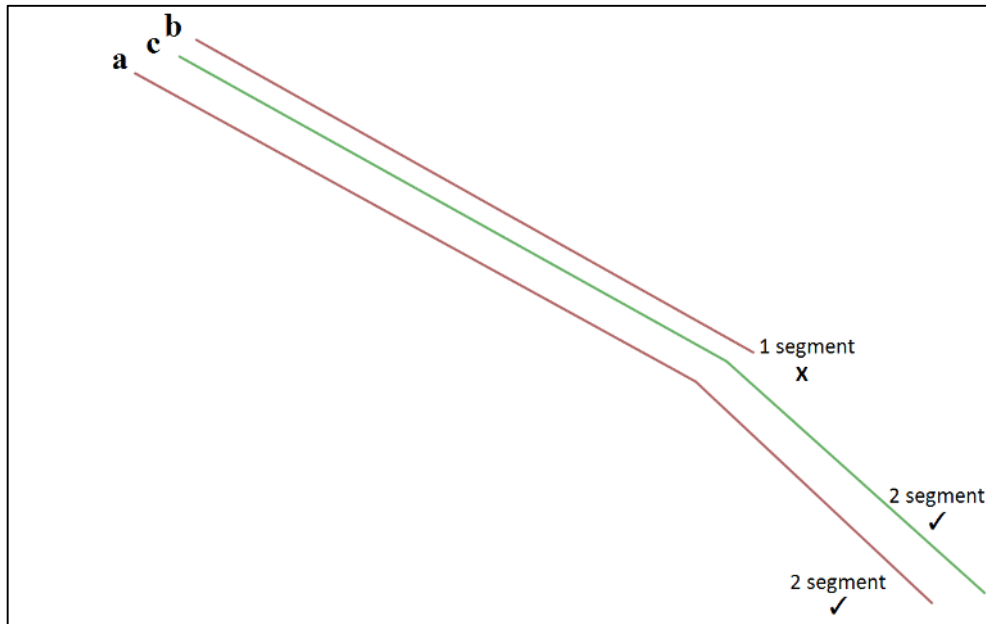
- X_COUNT: Hedef veri setindeki ilgili nesne ile eşlenen kaynak veri setindeki nesne sayısı
- X_MAX_SCORE: İlgili hedef nesne için hesaplanmış en yüksek puan
- X_MIN_DISTANCE: İlgili hedef nesne ile eşlenen kaynak nesneler arasındaki minimum mesafeler.

MatchingPlugin 19 adet eşleme yöntemine sahiptir. Bunlardan bazıları sadece nokta, çizgi veya alan eşlemelerinde kullanılırken, bazıları ise ikisi veya üçü için de kullanılabilir. Bu çalışmada aşağıda ayrıntılı olarak açıklanan beş farklı çizgi eşleme yöntemi kullanılmıştır.

Koordinata dayalı eşleme (Equals With Coordinate Tolerance Matcher)

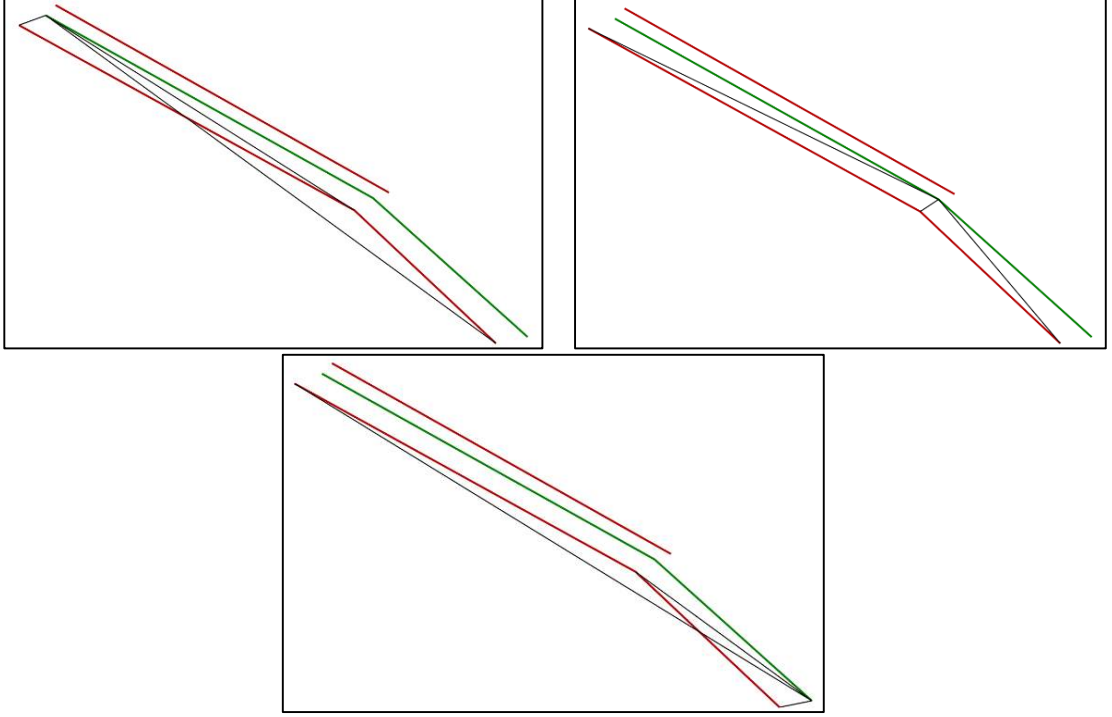
Bu yöntem, ilgili katmanların (hedef ve kaynak) ve eşik mesafe değerinin (T_e) kullanıcı tarafından belirtilmesinin ardından, sırasıyla aşağıdaki adımları uygular:

- Farklı katmanda yer alan, aynı sayıda segmente sahip çizgiler eşleme adayı olarak belirlenir (Şekil 4.6).



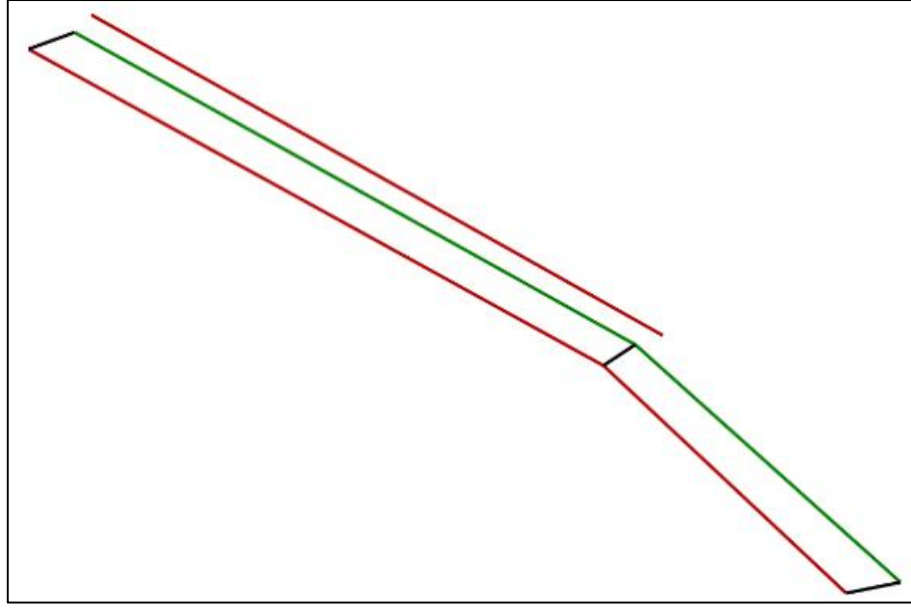
Şekil 4.6 Aynı (a ve c çizgileri eşleme adayları) ve farklı sayıda segmente sahip çizgiler

- Çizgilere ait tüm noktalar $N_n=\{1, 2, \dots, k\}$ ile bunların eşleme adayı çizgilerinin tüm noktaları $N_m=\{1, 2, \dots, l\}$ arasındaki D_{nk-m} Öklid mesafeleri hesaplanır (3.1). Burada N_n ; n numaralı çizginin k eleman sayılı nokta kümesini, N_m ise; m numaralı çizginin l eleman sayılı nokta kümesini göstermektedir. Aşağıda Şekil 4.7’de gösterilen, bir noktadan diğer katmandaki eş adayı çizginin bütün noktalarına uzanan mesafeler hesaplanır.



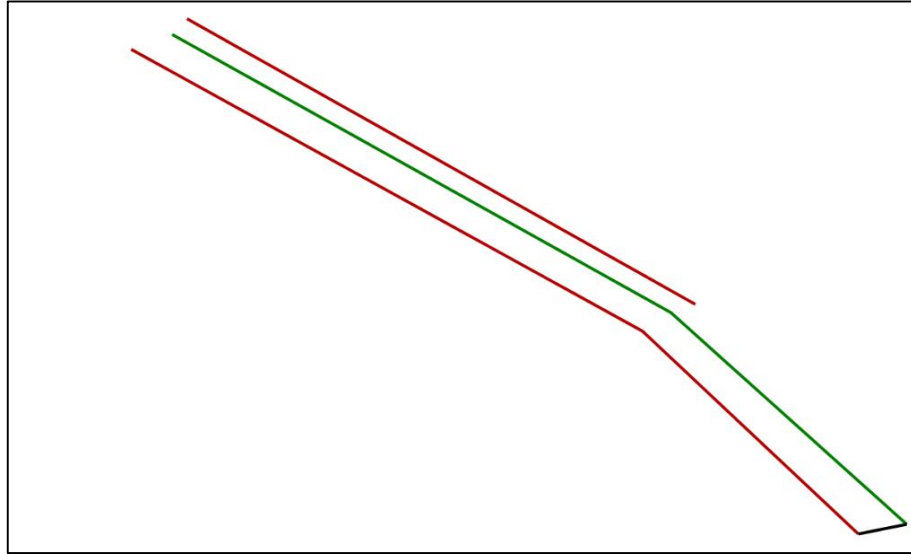
Şekil 4.7 Eşleme adayı çizgilerin noktaları arasındaki Öklid mesafeleri

- Her bir noktadan diğerlerine uzanan mesafelerden en kısa olanı seçilir. Örneğin aşağıda Şekil 4.8’de gösterildiği gibi a ve c çizgileri için D_{11-21} , D_{12-22} ve D_{13-23} mesafeleri seçilir.



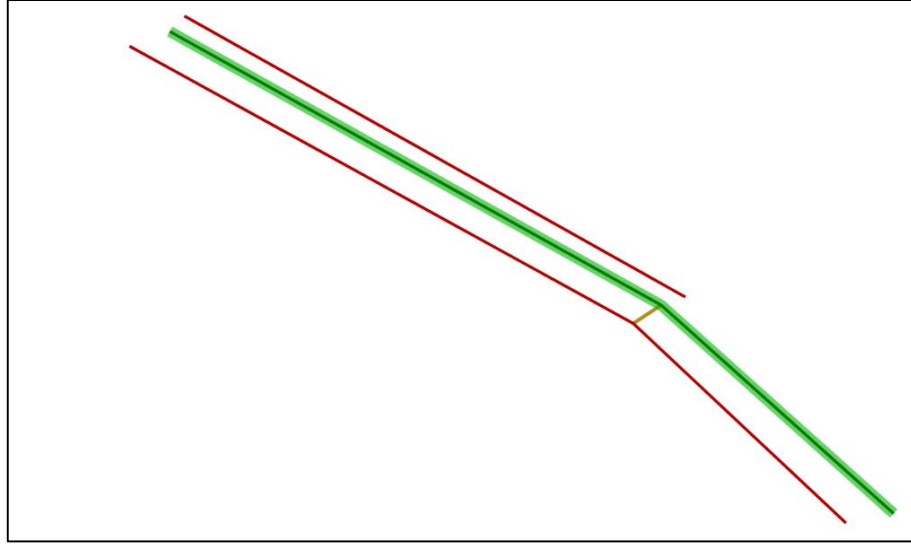
Şekil 4.8 Noktalar arası en kısa mesafeler

- İki eşleme adayı çizgi (a ve b) arasındaki mesafelerin (D_{11-21} , D_{12-22} ve D_{13-23}) en büyüğü ($D_{maks} = D_{13-23}$) seçilir (Şekil 4.9).



Şekil 4.9 Eşleme adayı çizgiler arasındaki seçilmiş mesafe D_{maks}

- $D_{maks} < T_e$ koşulunu sağlayan i ve j çizgileri eşlenir (Şekil 4.10).



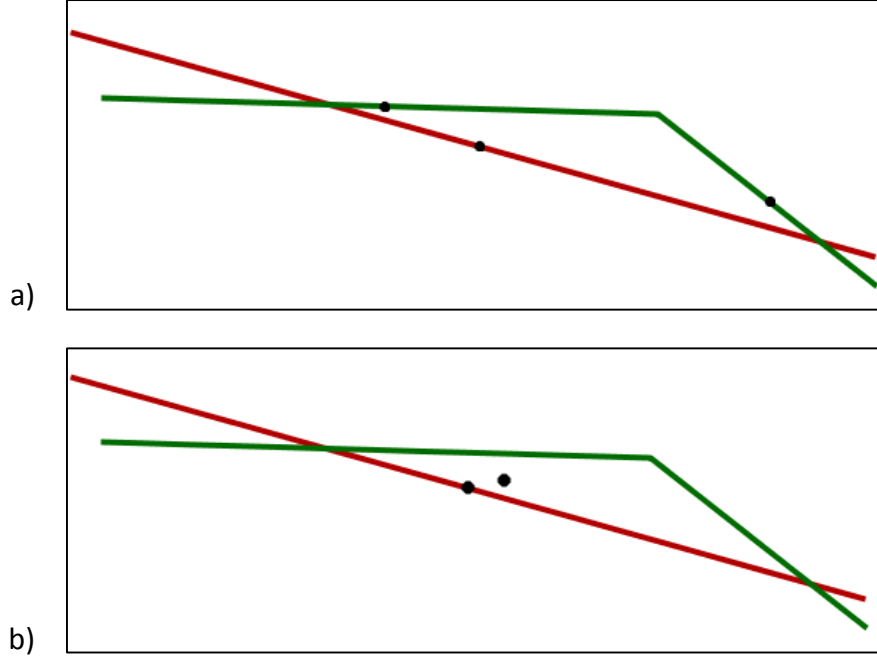
Şekil 4.10 Koordinata dayalı eşleme yöntemi ile eşlenmiş a ve b çizgileri

Ağırlık merkezine dayalı eşleme (Centroid Distance Matcher)

Ölçek değıştikçe nesnelerin geometrileri de değışime uğrar. Bu değışimden en az etkilenen geometri parametresi nesnelerin ağırlık merkezleridir. Bu sebeple farklı ölçekli verilerin eşlenmesinde iyi bir seçim olabilir. Bu yöntem, ilgili katmanların (hedef ve kaynak) ve eşik mesafe değeri (T_e) kullanıcı tarafından belirtilmesinin ardından, sırasıyla aşağıdaki adımları uygular:

- Her çizginin ağırlık merkezinin koordinatları X_g ve Y_g eşitlik (4.1)'de verilen formüllerle hesaplanır. Burada S_i i numaralı segmentin uzunluğunu, S toplam çizgi uzunluğunu, x_i ve y_i ise segmentlerin orta noktalarının koordinatlarını temsil etmektedir. Şekil 4.11'de segment ve çizgilerin ağırlık merkezleri görülmektedir.

$$X_g = \sum_{i=1}^n \frac{S_i x_i}{\sum S} \quad Y_g = \sum_{i=1}^n \frac{S_i y_i}{\sum S} \quad (4.1)$$



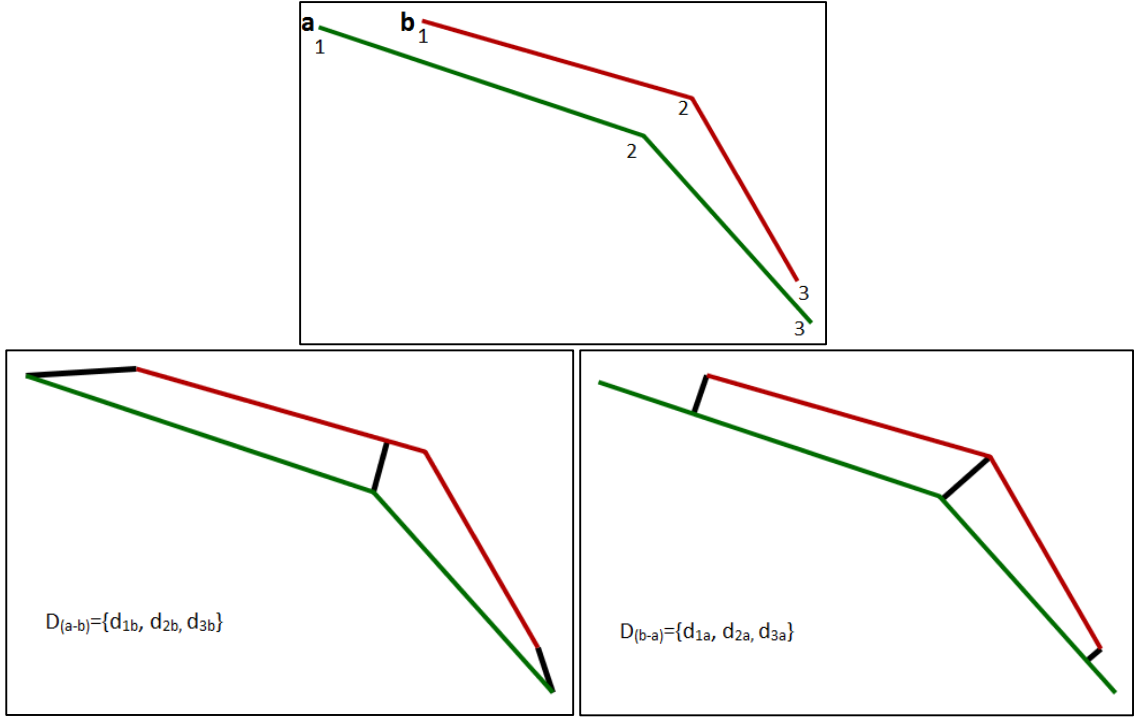
Şekil 4.11 (a) Segmentlerin ve (b) çizgilerin ağırlık merkezleri (siyah nokta)

- İlgili katmanlarda yer alan i ve j numaralı çizgilerin ağırlık merkezleri arasındaki d_{ij} Öklid mesafeleri ölçülür (3.1).
- $d_{ij} < T_e$ koşulunu sağlayan i ve j çizgileri eşlenir.

En kısa mesafeye dayalı eşleme (Minimum Distance Matcher)

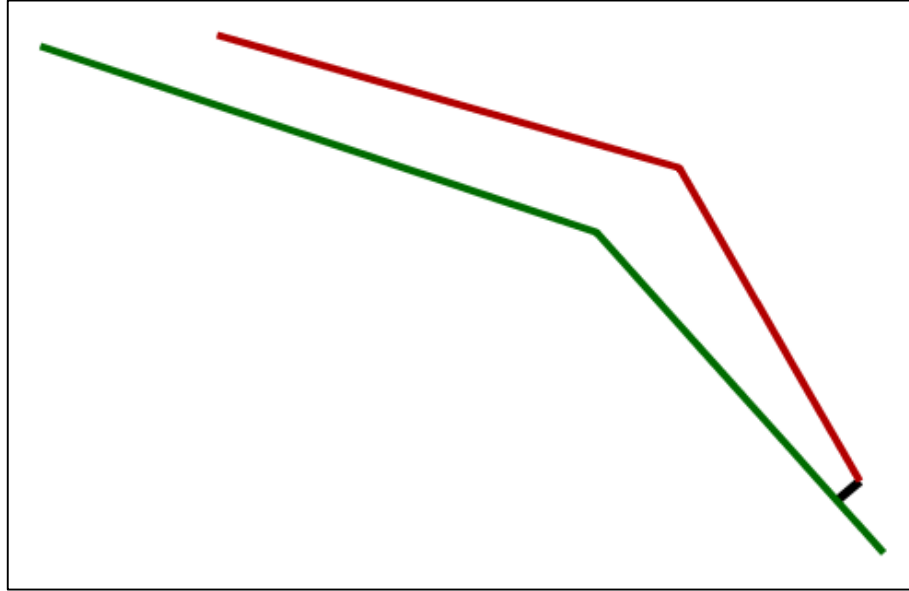
Bu yöntem nokta eşlemeleri için daha uygundur. Nokta çiftleri arası Öklid mesafelerinin eşik mesafeden küçük bir değere sahip olması durumunda eşleme işlemi gerçekleşir. Yöntem, çizgilere uygulanması durumunda, ilgili katmanların (hedef ve kaynak) ve eşik mesafe değerinin (T_e) kullanıcı tarafından belirtilmesinin ardından, sırasıyla aşağıdaki adımları uygular:

- Her bir çizgiyi oluşturan noktalar ile diğer katmandaki çizgiler arasındaki en kısa mesafe değerleri $D_{(n-m)} = \{d_{11}, d_{12}, \dots, d_{km}\}$ ve $D_{(m-n)} = \{d_{11}, d_{12}, \dots, d_{ln}\}$ hesaplanır. Burada d_{km} , n numaralı kaynak çizginin k numaralı noktası ile m numaralı hedef çizgi arasındaki en kısa mesafe değerini göstermektedir. Ayrıca, d_{ln} , m numaralı hedef çizginin l numaralı noktası ile n numaralı kaynak çizgi arasındaki en kısa mesafe değerini göstermektedir. Örneğin Şekil 4.12’de gösterildiği gibi hedef katmanında yer alan a ve kaynak katmanında yer alan b çizgileri arasındaki en kısa mesafe değerleri $D_{(a-b)} = \{d_{1b}, d_{2b}, d_{3b}\}$ ve $D_{(b-a)} = \{d_{1a}, d_{2a}, d_{3a}\}$ şeklinde gösterilebilir.



Şekil 4.12 Noktalar ile çizgiler arasındaki en kısa mesafeler

- Farklı katmanda yer alan çizgiler arasındaki bu mesafelerden $\min(D_{(n-m)}, D_{(m-n)})$ olanları seçilir. Örneğin aşağıda Şekil 13’de gösterildiği gibi $\min(D_{(a-b)}, D_{(b-a)}) = d_{3a}$ seçilir.



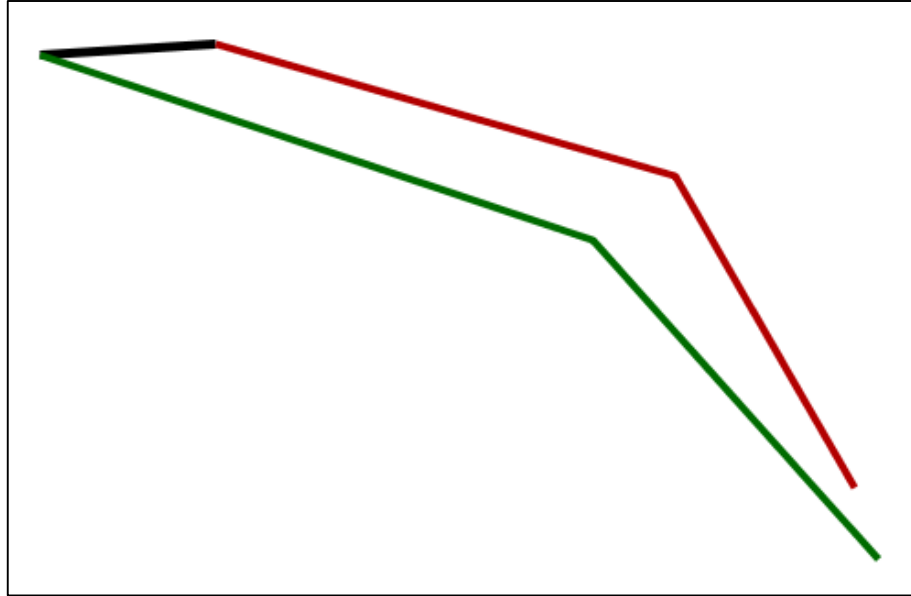
Şekil 4.13 Seçilen en kısa mesafe

- $\min(D_{(i-j)}, D_{(j-i)}) < T_e$ koşulunu sağlayan i ve j çizgileri eşlenir.

Hausdorff mesafesine dayalı eşleme (Hausdorff Distance Matcher)

Hausdorff mesafesi, iki çizgi arasındaki en kısa mesafelerin en büyüğü olarak kısaca tanımlanabilir. Bu yöntem, ilgili katmanların (hedef ve kaynak) ve eşik mesafe değerinin (T_e) kullanıcı tarafından belirtilmesinin ardından, sırasıyla aşağıdaki adımları uygular:

- Her bir çizgiyi oluşturan noktalar ile diğer katmandaki çizgiler arasındaki en kısa mesafe değerleri $D_{(n-m)} = \{d_{11}, d_{12}, \dots, d_{km}\}$ ve $D_{(m-n)} = \{d_{11}, d_{12}, \dots, d_{ln}\}$ hesaplanır. Burada; d_{km} , n numaralı kaynak çizginin k numaralı noktası ile m numaralı hedef çizgi arasındaki en kısa mesafe değerini göstermektedir. Ayrıca, d_{ln} , m numaralı hedef çizginin l numaralı noktası ile n numaralı kaynak çizgi arasındaki en kısa mesafe değerini göstermektedir. Örneğin Şekil 4.12’de gösterildiği gibi hedef katmanında yer alan a ve kaynak katmanında yer alan b çizgileri arasındaki en kısa mesafe değerleri $D_{(a-b)} = \{d_{1b}, d_{2b}, d_{3b}\}$ ve $D_{(b-a)} = \{d_{1a}, d_{2a}, d_{3a}\}$ şeklinde gösterilebilir.
- Farklı katmanda yer alan çizgiler arasındaki bu mesafelerden $\max(D_{(n-m)}, D_{(m-n)})$ olanları seçilir. Örneğin aşağıda Şekil 4.14’de gösterildiği gibi $\max(D_{(a-b)}, D_{(b-a)}) = d_{1b}$ seçilir.



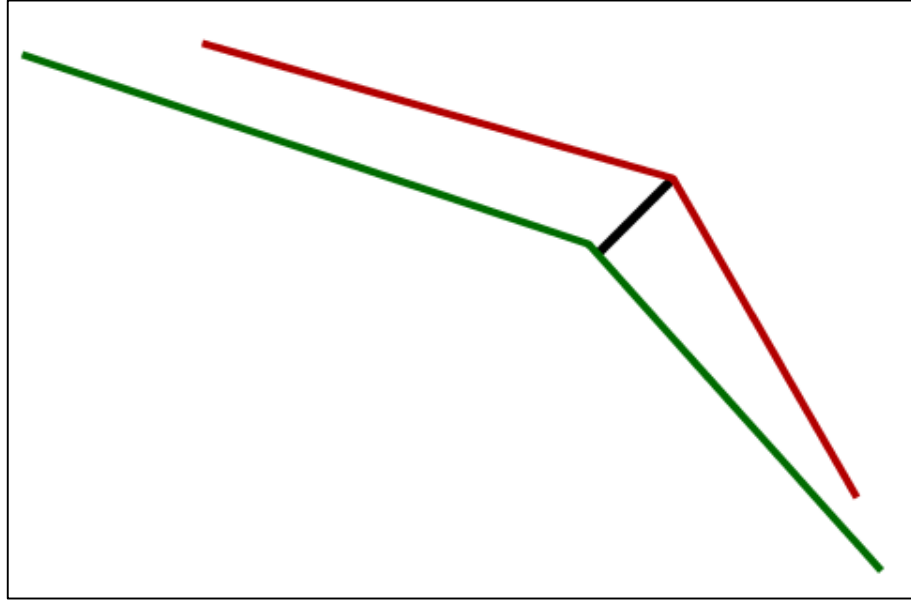
Şekil 4.14 Seçilen en uzun mesafe

- $\max(D_{(i-j)}, D_{(j-i)}) < T_e$ koşulunu sağlayan i ve j çizgileri eşlenir.

Kısmi-Hausdorff mesafesine dayalı eşleme (Semi-Hausdorff Distance Matcher)

Kaynak katmandaki çizgiden, hedef (referans) katmandaki çizgiye olan tek yönlü Hausdorff mesafesine dayalı eşleme işlemidir. Bu yöntem, ilgili katmanların (hedef ve kaynak) ve eşik mesafe değerinin (T_e) kullanıcı tarafından belirtilmesinin ardından, sırasıyla aşağıdaki adımları uygular:

- Kaynak katmandaki her bir çizgiyi oluşturan noktalar ile diğer katmandaki çizgiler arasındaki en kısa mesafe değerleri $D_{(n-m)} = \{d_{11}, d_{12}, \dots, d_{km}\}$ hesaplanır. Burada d_{km} , n numaralı kaynak çizginin k numaralı noktası ile m numaralı hedef çizgi arasındaki en kısa mesafe değerini göstermektedir. Örneğin Şekil 4.12’de gösterildiği gibi a ve b çizgileri arasındaki en kısa mesafe değerleri $D_{(a-b)} = \{d_{1b}, d_{2b}, d_{3b}\}$ dir.
- Bu mesafelerden $\max(D_{(n-m)})$ olanları seçilir. Örneğin a çizgisinden b çizgisine Kısmi-Hausdorff mesafesi $\max(D_{(a-b)}) = d_{1b}$ (Şekil 4.14), fakat b çizgisinden a çizgisine Kısmi-Hausdorff mesafesi $\max(D_{(b-a)}) = d_{2a}$ seçilir (Şekil 4.15).



Şekil 4.15 b çizgisinden a çizgisine Kısmi-Hausdorff mesafesi

- $\max(D_{(i-j)}) < T_e$ koşulunu sağlayan i ve j çizgileri eşlenir.

4.2.1.2 RoadMatcher ile Eşleme

RoadMatcher (RoadMatcher15) eklentisi, aynı yere ait yol orta eksen ağlarını segment (iki düğüm noktası arasında kalan çizgi parçası) bazında eşleme imkânı sunar. Eşleme

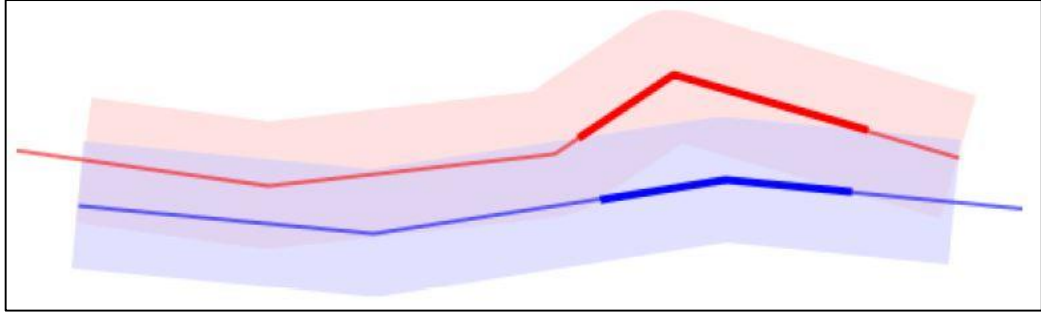
işleminin gerçekleşmesinin ardından istenildiği takdirde bir sonuç katmanı çıkartılabilmektedir. Eklenti, sonuç katmanı çıkartırken, kullanıcının transfer edilecek geometri ve özniteliklere müdahale etmesine olanak tanır. Kullanıcı eşleme sonrası oluşturulacak ürüne ait geometrik verilerin, hangi katmandan alınacağını belirleyebilmektedir. Ayrıca sonuç üründe öznitelik tablolarının tümünün birleştirilmesi yerine, sadece kullanıcının belirlediği sütunların transferi gerçekleştirilebilmektedir.

RoadMatcher eklentisinde otomatik eşlemenin başlatılabilmesi için verilerin aşağıdaki koşulları sağlaması gereklidir [48]:

- Bütün segmentler çizgi (linestring) geometrili olmalıdır.
- Çizgilerin kesiştiği yerlerde çizgiler sonlanmış olmalıdır. Yani verisetleri çizgi-düğüm noktası yapısında olmalıdır.

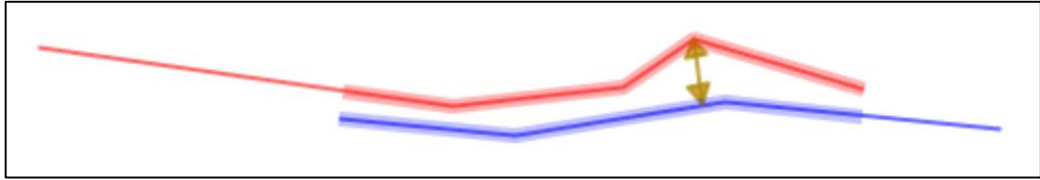
Kullanıcı, eklenti arayüzündeki seçenek menüsünden, değerlerini önceden belirlemek zorunda olduğu aşağıdaki dört eşik değerden ikisini zorunlu diğer ikisini isteğe bağlı kullanabilir.

- **En kısa segment uzunluğu (Zorunlu):** Bu kriterdeki segment tanımı, düğüm noktaları arasındaki uzunluk değildir. Burada segment iki kırılma noktası (verteks) arasındaki uzaklıktır. Eşleme sürecinde 1-1 ve 1-m eşleme başarısını arttırmak için algoritma bazı segmentleri parçalar. Segmenti parçalarken segmentin uzunluğunun eşik değerden uzun olması gerekir.
- **Yakınlık (Zorunlu):** Otomatik eşleme sürecinde kullanılmaz fakat oluşacak sonuç üründe analiz yapılması açısından önemlidir. Sonuç üründe eşlenmiş çizgiler arasındaki yakınlık değeri öznitelik tablosuna yazılır. Çizgilerin ne kadarının, eşik mesafesi içinde yer aldığını gösteren kesirli bir sayıdır. Eşleme adayı çizginin tamamı eşik mesafesi içinde yer alıyorsa “Yakınlık” değeri 1 olur (Şekil 4.16).



Şekil 4.16 Eşlenen çiftler arası yakınlığın belirlenmesi [48]

Eşik mesafesi seçilen birkaç çizginin ortak kısmı için belirlenen Hausdorff mesafesi (kırılmış Hausdorff mesafesi) ile belirlenebilir (Şekil 4.17). Çizgilerin ortak kısımları, başlangıç ve bitiş noktalarından diğer aday çizgiyi dik kestiği noktalar arasında kalan çizgiler şeklinde belirlenir.



Şekil 4.17 Kırılmış Hausdorff mesafesi [48]

- **En kısa mesafe:** Verisetlerindeki yalnız (başka yollar ile bağlantısı olmayan) yol nesnelerinin tespiti için kullanılır. Segmentler arası mesafe değerinin eşik değerinden büyük olduğu durumlarda yalnız nesneler belirlenir.
- **Hausdorff mesafesi:** bk. Bölüm 4.2.1.1

Eşleme işleminin başarısı, örnek verilerle yapılacak bir önışlem (preprocessing) sürecinde eşik değerlerin belirlenmesi ile doğrudan ilgilidir. Bu sebeple kullanıcının yazılımda varsayılan (default) olarak gelen eşik değerler yerine önışlem sürecinde belirlediği eşik değerleri kullanması gereklidir.

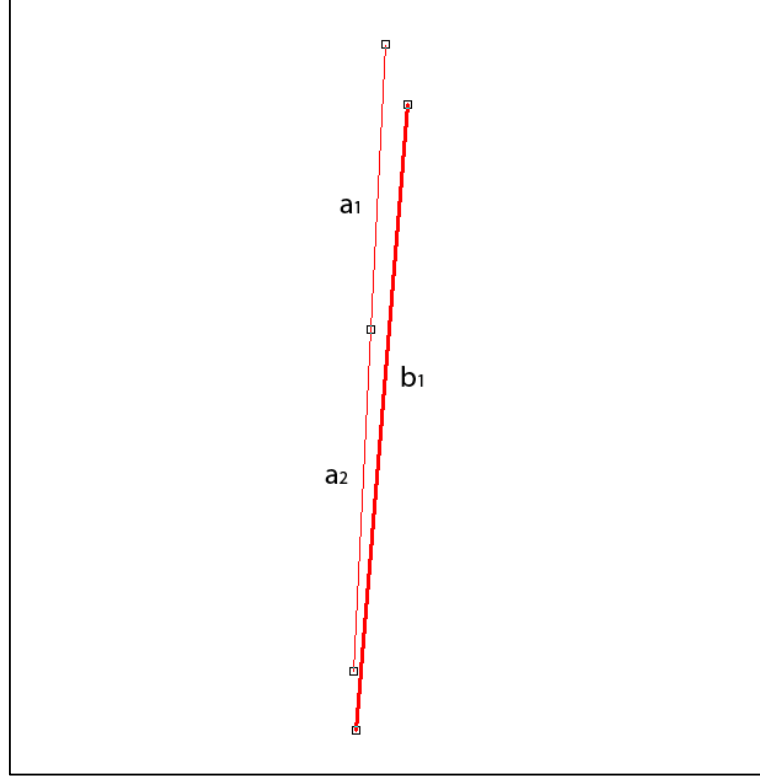
4.2.2 Bir Optimizasyon Modeli ile Eşleme

Bu çalışmada yukarıda tanıtılan yöntemler dışında, Li [22] ile Li ve Goodchild [23] tarafından geliştirilen optimizasyon modeli de test edilmiştir. Yönteme göre,

- Şekil 4.18’de görüldüğü gibi a_1 ve a_2 çizgileri ile b_1 çizgisinin eşlenmesi gerektiği açıktır. Ancak a_1 ile b_1 veya a_2 ile b_1 arasında Hausdorff mesafesi b_1 ’in diğerlerinden uzun

olması sebebiyle yüksek çıkacağından benzerlik kriteri sağlanmayacaktır. Bu sebeple farklı verisetlerinde yer alan çizgiler arasında yönlendirilmiş Hausdorff mesafeleri belirlenir (4.2).

$$d_{i-j} = maks(\min(d_1, d_2)) \quad (4.2)$$



Şekil 4.18 Yönlendirilmiş Hausdorff mesafesi (Li ve Goodchild [23])

- Yönlendirilmiş mesafelere göre, benzerlik eşitliği S_{i-j} ve a eşik mesafe değeri ile aday eşleme sayısı azaltılır (4.3).

$$S_{i-j} = \begin{cases} 0 & d_{i-j} > a \\ a - d_{i-j} & d_{i-j} < a \end{cases} \quad (4.3)$$

- Eşlemenin daha başarılı sonuç vermesi için eşleme adayı çizgi uzunluğu koşul olarak belirlenmiştir. Buna göre, ikinci verisetindeki her bir çizgi ile eşlenecek birinci verisetindeki çizgilerin toplam uzunluğunun ikinci verisetindeki çizginin uzunluğunun β katını geçmemesi gerekir. Diğer bir deyişle β kat sayısı öyle bir değer ki, ölçeği veya detay seviyesi küçük olan verisetinin detay uzunluğu ile çarpılarak diğer verisetindeki eşi ile yaklaşık olarak aynı uzunluğa gelmesini sağlamaktadır. Ancak bu katsayı tüm

veriseti için önceden belirlenmiş tek değer olduğu için tüm uzunlukları eşitleyebilmesi beklenmez.

- Tüm adımlar tersinir şekilde ($j \rightarrow i$) de tekrarlanır ve toplam benzerlikler hesaplanır.

Ayrıca afin dönüşümü yaparak iki veriseti arasındaki geometrik uyumsuzlukların azaltılması için aşağıdaki adımlarla yöntemine devam edilmektedir [22], [23]:

- (1) Birinci verisetindeki çizgilerin başlangıç ve sonlarında nokta oluşturulur.
- (2) Bölge, kontrol noktası sayısı k kadar alt bölgelere ayrılır. Her bir bölgeden bir kontrol noktası belirlenir. Böylece seçilen kontrol noktalarının olabildiğince dağılması sağlanır.
- (3) Birinci verisetindeki her bir p_r kontrol noktası için, olabilecek en uzun ötelemeye göre belirlenen, tampon eşik mesafesi içindeki diğer verisetinin $Q_r = \{q_{r1}, q_{r2}, \dots, q_{rs}\}$ noktaları eşleme adayı olarak belirlenir. Burada s mümkün olan eşleme sayısını temsil etmektedir.
- (4) Birinci verisetindeki her bir p_r noktası için Q_r kümesindeki noktalardan biri seçilir ve bu noktalara dayalı afin dönüşüm parametreleri (a, b, c, d, e ve f) hesaplanır.
- (5) Dönüşüm sonrası eşleme işlemi tekrarlanır ve toplam benzerlik değeri hesaplanır.
- (6) 4 ve 5'deki adımlar bütün kontrol noktası kombinasyonları için tekrarlanır. Bu model toplam benzerlik değerinin en yüksek çıktığı afin dönüşüm parametrelerini belirler.

Böylelikle toplam benzerliğin en yüksek olduğu dönüşüm sonrasında eşlenmiş çizgiler belirlenmiş olur.

Li [22] ile Li ve Goodchild [23], yol isimleri arasındaki benzerlikleri belirlediği Hamming mesafesini de bir değişken olarak eşitlik (4.3)'te kullanmıştır. Bu çalışmada öznitelik verileriyle eşleme işlemi hakkında herhangi bir uygulama yapılmamıştır. Bu sebeple uygulamanın kapsamında olmadığı için Hamming mesafesi değişkeni eşitlikten çıkarılmıştır.

4.3 Uygulama Sonuçları

Uygulama bölgesi otomatik eşlemenin doğruluğunun kontrolünü kolaylaştırmak amacıyla Bölge 1, Bölge 2 ve Bölge 3 olmak üzere üç alt bölgeye ayrılmıştır (Şekil 4.2). Her alt bölgede, kullanılan otomatik yöntemlerin başarısını ölçmek amacıyla, manuel eşleme yapılmıştır (Çizelge A.1). Aynı yola ait detaylar (düğüm noktaları arasında kalan çizgiler) operatör kontrolüyle teker teker eşlenmiştir.

Bazı nesneler Veriseti 1’de iki veya daha fazla detayla temsil edilirken, Veriseti 2’de tek detay ile temsil edilmektedir (ya da tam tersi). Bu sebeple manuel eşlemede eşlenen detay çiftlerinin sayısı 1:1, 1:N ve N:1 eşleme durumları dikkate alınarak belirlenmiştir (Çizelge 4.1).

Çizelge 4.1 Alt bölgelerde manuel olarak eşlenmiş çift sayıları

	Bölge 1		Bölge 2		Bölge 3	
	Veriseti 1	Veriseti 2	Veriseti 1	Veriseti 2	Veriseti 1	Veriseti 2
Detay Sayısı	365	367	293	272	306	190
Eşlenen Detay Sayısı	332	332	248	232	215	176
Eşlenmeyen Detay Sayısı	33	35	45	40	91	14
Eşlenen Çift Sayısı	332		248		215	

Manuel eşleme ile karşılaştırılan otomatik yöntemlerin başarısını ölçmek için; “Eşleme”, “Doğru Eşleme”, “Yanlış Eşleme”, “Eksik Eşleme” ve “Eşlenmeme” değerleri belirlenmiştir. Bu çalışmada kullanılan yöntemler ayrı başlıklar altında aşağıda analiz edilmiştir.

4.3.1 MatchingPlugin Eklentisi

Bölüm 4.2.1.1’de açıklanan “Koordinata dayalı eşleme”, “Ağırlık merkezine dayalı eşleme”, “En kısa mesafeye dayalı eşleme”, “Hausdorff mesafesine dayalı eşleme” ve “Kısmi-Hausdorff mesafesine dayalı eşleme” yöntemleri üç alt bölgede uygulanmıştır

(Çizelge A.2, Çizelge A.3 ve Çizelge A.4). Eklentinin seçenek olarak sunduğu 1:1 ve 1:N eşleme işlevleri tüm yöntemlerde S eşik mesafe değeri 10 metre alınarak test edilmiştir. Eşik değer, verisetlerindeki çizgiler arası mesafelerin manuel ölçülmesi ile bulunmuştur ve çizgiler arasında olabilecek maksimum mesafeyi temsil etmektedir.

Bölge 1, Bölge 2 ve Bölge 3'e ait test sonuçları sırasıyla Çizelge 4.2, Çizelge 4.3 ve Çizelge 4.4'te manuel eşleme sayıları, otomatik eşleme sonuçlarına göre doğru, yanlış ve eksik eşleme sayıları ve eşlenmeyen detay çiftlerinin sayıları şeklinde verilmiştir.

Çizelge 4.2 Bölge 1'de MatchingPlugin eklentisi ile elde edilen 1:N ve 1:1 eşleme sayıları

Bölge 1	Koordinata dayalı eşleme		Ağırlık merkezine dayalı eşleme		En kısa mesafeye dayalı eşleme		Hausdorff mesafesine dayalı eşleme		Kısmi-Hausdorff mesafesine dayalı eşleme	
	1:N	1:1	1:N	1:1	1:N	1:1	1:N	1:1	1:N	1:1
Manuel	332		332		332		332		332	
Otomatik	35	33	238	222	332	292	185	185	326	247
Doğru	24	24	157	205	17	150	148	171	44	203
Yanlış	9	7	78	7	314	121	32	7	281	13
Eksik	2	2	3	10	1	21	5	7	1	31
Eşlenmeyen	297	299	94	110	0	40	147	147	6	85

Çizelge 4.3 Bölge 2'de MatchingPlugin eklentisi ile elde edilen 1:N ve 1:1 eşleme sayıları

Bölge 2	Koordinata dayalı eşleme		Ağırlık merkezine dayalı eşleme		En kısa mesafeye dayalı eşleme		Hausdorff mesafesine dayalı eşleme		Kısmi-Hausdorff mesafesine dayalı eşleme	
	1:N	1:1	1:N	1:1	1:N	1:1	1:N	1:1	1:N	1:1
Manuel	248		248		248		248		248	
Otomatik	11	10	163	151	248	208	140	139	246	176
Doğru	11	10	145	149	22	126	139	138	45	154
Yanlış	0	0	17	1	226	67	0	0	201	2
Eksik	0	0	1	1	0	15	1	1	0	20
Eşlenmeyen	237	238	85	97	0	40	108	109	2	72

Çizelge 4.4 Bölge 3’de MatchingPlugin eklentisi ile elde edilen 1:N ve 1:1 eşleme sayıları

Bölge 3	Koordinata dayalı eşleme		Ağırlık merkezine dayalı eşleme		En kısa mesafeye dayalı eşleme		Hausdorff mesafesine dayalı eşleme		Kısmi-Hausdorff mesafesine dayalı eşleme	
	1:N	1:1	1:N	1:1	1:N	1:1	1:N	1:1	1:N	1:1
Manuel	215		215		215		215		215	
Otomatik	12	12	115	112	214	143	96	96	205	125
Doğru	12	12	98	107	36	66	93	93	58	109
Yanlış	0	0	14	1	178	69	0	0	145	2
Eksik	0	0	3	4	0	8	3	3	2	14
Eşlenmeyen	203	203	100	103	1	72	119	119	10	90

Eşlenen detay çiftlerinin doğru olup olmadığını ölçmek için manuel eşleme referans alınarak yüzdelikler hesaplanmıştır. Yöntemlerin çoğunun eşleme işlemini yüksek oranda gerçekleştirdiği fakat eşlenen çiftlerin büyük kısmının yanlış olduğu görülmektedir (Çizelge 4.5, Çizelge 4.6 ve Çizelge 4.7). Örneğin 1:N eşlemede en fazla eşleme sayısına sahip En kısa mesafeye dayalı eşleme yöntemi ile detayların ortalama %99,8’i eşlenmesine (Çizelge 4.5) rağmen, eşlenen çiftlerin büyük kısmının yanlış olduğu anlaşılmaktadır (Çizelge 4.7). Bununla birlikte, 1:1 eşlemede en iyi sonucu veren Kısmi-Hausdorff yöntemi ile detayların ortalama %58’si doğru eşlenebilmiştir (Çizelge 4.6).

Çizelge 4.5 MatchingPlugin otomatik eşleme yüzdeleri $((Otomatik/Manuel) \times 100)$

	Bölge 1		Bölge 2		Bölge 3		Ortalama	
	1:N	1:1	1:N	1:1	1:N	1:1	1:N	1:1
Koordinata dayalı eşleme	10.5	9.9	4.4	4.0	5.58	5.58	6.85	6.52
Ağırlık merkezine dayalı eşleme	71.7	66.9	65.7	60.9	53.5	52.1	63.6	60.0
En kısa mesafeye dayalı eşleme	100	88.0	100	83.9	99.5	66.5	99.8	79.4
Hausdorff mesafesine dayalı eşleme	55.7	55.7	56.5	56.1	44.7	44.7	52.2	52.1
Kısmi-Hausdorff mesafesine dayalı eşleme	98.2	74.4	99.2	71.0	95.4	58.1	97.6	67.8

Çizelge 4.6 MatchingPlugin otomatik doğru eşleme yüzdeleri $((Doğru/Manuel) \times 100)$

	Bölge 1		Bölge 2		Bölge 3		Ortalama	
	1:N	1:1	1:N	1:1	1:N	1:1	1:N	1:1
Koordinata dayalı eşleme	7.2	7.2	4.4	4.0	5.6	5.6	5.8	5.6
Ağırlık merkezine dayalı eşleme	47.3	61.8	58.5	60.1	45.6	49.8	50.5	57.2
En kısa mesafeye dayalı eşleme	5.1	45.2	8.9	50.8	16.7	30.7	10.3	42.2
Hausdorff mesafesine dayalı eşleme	44.6	51.5	56.1	55.7	43.3	43.3	48.0	50.1
Kısmi-Hausdorff mesafesine dayalı eşleme	13.3	61.1	18.2	62.1	27.0	50.7	19.5	58.0

Ağırlık merkezine dayalı eşleme yöntemi neredeyse Kısmi-Hausdorff yöntemi ile aynı sonucu vermiştir. Hatta genel olarak yanlış eşleme sayısı Hausdorff yönteminden sonra en iyi ikinci sonucu vermiştir (Çizelge 4.7).

Çizelge 4.7 MatchingPlugin otomatik yanlış eşleme yüzdeleri
 $((Yanlış/Otomatik) \times 100)$

	Bölge 1		Bölge 2		Bölge 3		Ortalama	
	1:N	1:1	1:N	1:1	1:N	1:1	1:N	1:1
Koordinata dayalı eşleme	25.7	21.2	0.0	0.0	0.0	0.0	8.57	7.07
Ağırlık merkezine dayalı eşleme	32.8	3.2	10.4	0.7	12.2	0.9	18.5	1.6
En kısa mesafeye dayalı eşleme	94.6	41.4	91.1	32.2	83.2	48.3	89.6	40.6
Hausdorff mesafesine dayalı eşleme	17.3	3.8	0.0	0.0	0.0	0.0	5.8	1.3
Kısmi-Hausdorff mesafesine dayalı eşleme	86.2	5.3	81.7	1.1	70.7	1.6	79.6	2.7

Dikkat çeken bir diğer nokta ise Hausdorff yöntemi ile yapılan eşlemelerin çok az sayıda detayı yanlış eşlemesidir. Ancak Çizelge 4.5'te görüldüğü gibi Hausdorff yöntemi eşlenebilecek detayların yalnızca %52'sini eşleyebildiği için yetersizdir.

4.3.2 RoadMatcher Eklentisi

Bölüm 4.2.1.2'de açıklanan RoadMatcher ile üç alt bölgede uygulama yapılmıştır. Eklentinin seçenek olarak sunduğu 1:1 ve 1:N eşleme işlevleri S eşik mesafe değeri önceki uygulamada olduğu gibi 10 metre alınarak test edilmiştir. Eşleme durumu Çizelge A.5'te, eşleme sayıları ise Çizelge 4.8 ve 4.9'da verilmiştir.

Çizelge 4.8 RoadMatcher eklentisi ile elde edilen 1:N ve 1:1 eşleme sayıları

	Bölge 1		Bölge 2		Bölge 3	
	1:N	1:1	1:N	1:1	1:N	1:1
Manuel	332		248		215	
Otomatik	316	316	239	239	197	197
Doğru	287	266	229	209	187	176
Yanlış	21	21	10	10	9	9
Eksik	8	29	0	20	1	12
Eşlenmeyen	16	16	9	9	18	18

Yanlış eşleme tüm bölgelerde aynıken, eksik eşleme sayıları 1:N eşlemede daha azdır (Çizelge 4.8). Bu sebeple, RoadMatcher, 1:N ve 1:1 eşlemede eşit sayıda detay eşlemesinden dolayı, tüm alt bölgelerde 1:N eşlemede daha iyi (doğru sayısı fazla) sonuç vermiştir. RoadMatcher, MatchingPlugin’de yer alan en iyi yöntemlerden (Ağırlık merkezine dayalı ve Kısmi-Hausdorff eşleme yöntemi) yaklaşık 1.5-2 kat daha fazla detay çiftini doğru eşlemiştir (Çizelge 4.2, 4.3, 4.4 ve 4.8).

Yöntem en iyi sonucu Bölge 2’de yaklaşık %92 doğrulukla vermiştir (Çizelge 4.9). Demek oluyor ki, Bölge 2 için ortalama her 13 yol detayından 1’i ya eşlenmemiş, ya eksik ya da hatalıdır. Yol ağları ile eşleme yaparken, eşlemenin amacına göre bu doğruluk yeterli olmayabilir. Örneğin, ambulans, itfaiye, posta ya da okul servisi, taksi gibi araçlar, yol ağlarının görselleştirildiği navigasyon sistemlerini kullanırlar. Bu sistemler daima güncel ve güvenilir olmalıdır. Bu bağlamda, böyle bir sistemde örneğin RoadMatcher yöntemi ile veriler güncellenirse, en başarılı durumda bile ambulans ve itfaiye gibi acil müdahale araçları 1/13 ihtimalle yanlış yollarda vakit kaybedebilirler.

Çizelge 4.9 RoadMatcher 1:N ve 1:1 eşleme yüzdeleri ($\times 100$)

	Bölge 1		Bölge 2		Bölge 3		Ortalama	
	1:N	1:1	1:N	1:1	1:N	1:1	1:N	1:1
Otomatik/Manuel	95.2	95.2	96.4	96.4	91.6	91.6	94.4	94.4
Doğru/Manuel	86.5	80.1	92.3	84.3	87.0	81.9	88.6	82.1
Yanlış/Otomatik	6.7	6.7	4.2	4.2	4.6	4.6	5.1	5.1

4.3.3 Optimizasyon Modeli

Bölüm 4.2.2’de açıklanan optimizasyon modeli Bölge 1’de uygulanmıştır (Çizelge A.6). Yöntem 1:N eşlemeyi seçenek olarak sunmadığı için 1:1 eşleme işlevi test edilmiştir. Çizelge 4.10’da sabit β değerleri ($\beta=\{1.4, 1.5, 1.6\}$) ile a eşik mesafelerine göre eşleme sayılarındaki değişim görülmektedir.

Çizelge 4.10 Bölge 1’de optimizasyon modeli ile elde edilen eşleme sayıları

Bölge 1 $\beta = 1.4$	a=10 m	a=15 m	a=20 m	a=25 m	a=30 m	a=35 m
Manuel	332	332	332	332	332	332
Otomatik	248	266	276	282	287	292
Doğru	207	219	221	226	225	225
Yanlış	11	17	25	26	32	35
Eksik	30	30	30	30	30	32
Eşlenmeyen	84	66	56	50	45	40

Çizelge 4.10 Bölge 1’de optimizasyon modeli ile elde edilen eşleme sayıları (devamı)

Bölge 1 $\beta = 1.5$	a=10 m	a=15 m	a=20 m	a=25 m	a=30 m	a=35 m
Manuel	332	332	332	332	332	332
Otomatik	248	266	276	282	287	292
Doğru	206	218	220	225	224	224
Yanlış	12	18	26	27	32	37
Eksik	30	30	30	30	31	31
Eşlenmeyen	84	66	56	50	45	40

Çizelge 4.10 Bölge 1’de optimizasyon modeli ile elde edilen eşleme sayıları (devamı)

Bölge 1 $\beta = 1.6$	a=10 m	a=15 m	a=20 m	a=25 m	a=30 m	a=35 m
Manuel	332	332	332	332	332	332
Otomatik	248	266	276	282	287	292
Doğru	206	221	224	228	227	228
Yanlış	12	15	22	24	29	32
Eksik	30	30	30	30	31	32
Eşlenmeyen	84	66	56	50	45	40

Yöntem eşleme işlemini yüksek oranda gerçekleştirmiş fakat eşlenen çiftlerin bir kısmının eksik olduğu görülmektedir (Çizelge 4.10). Bunun nedeni ise 1:N eşlemenin yapılamamış olmasıdır.

Yapılan deneysel çalışmada, a eşik mesafe ve β değerlerinin arttıkça yanlış eşleme sayısının arttığı, doğru eşleme sayısının ise belli bir aralıkta (%67-68) kaldığı görülmektedir (Çizelge 4.11).

Çizelge 4.11 Bölge 1’de a eşik mesafe ve β katsayı değerlerine göre hesaplanmış eşleme yüzdeleri ($\times 100$)

Bölge 1	Otomatik /Manuel	Doğru /Manuel	Yanlış /Otomatik
a=10 m; $\beta=1.4$	74,7	62,3	4,4
a=10 m; $\beta=1.5$	74,7	62,0	4,8
a=10 m; $\beta=1.6$	74,7	62,0	4,8
a=15 m; $\beta=1.4$	80,1	66,0	6,4
a=15 m; $\beta=1.5$	80,1	65,7	6,8
a=15 m; $\beta=1.6$	80,1	66,6	5,6
a=20 m; $\beta=1.4$	83,1	66,6	9,1
a=20 m; $\beta=1.5$	83,1	66,3	9,4
a=20 m; $\beta=1.6$	83,1	67,5	8,0
a=25 m; $\beta=1.4$	84,9	68,1	9,2
a=25 m; $\beta=1.5$	84,9	67,8	9,6
a=25 m; $\beta=1.6$	84,9	68,7	8,5
a=30 m; $\beta=1.4$	86,4	67,8	11,1
a=30 m; $\beta=1.5$	86,4	67,5	11,1
a=30 m; $\beta=1.6$	86,4	68,4	10,1
a=35 m; $\beta=1.4$	88,0	67,8	12,0
a=35 m; $\beta=1.5$	88,0	67,5	12,7
a=35 m; $\beta=1.6$	88,0	68,7	11,0

Bölge 1 için yöntem, a=25 metre ve $\beta=1.6$ değerlerinde maksimum doğru ve minimum yanlış yüzdelerinde sonuç vermiştir. Ancak %68 doğruluğa sahip bir eşleme ile güncelleme, veri transferi, bütünleştirme gibi veri kalitesini arttırmak için yapılan işler yapılamamaktadır.

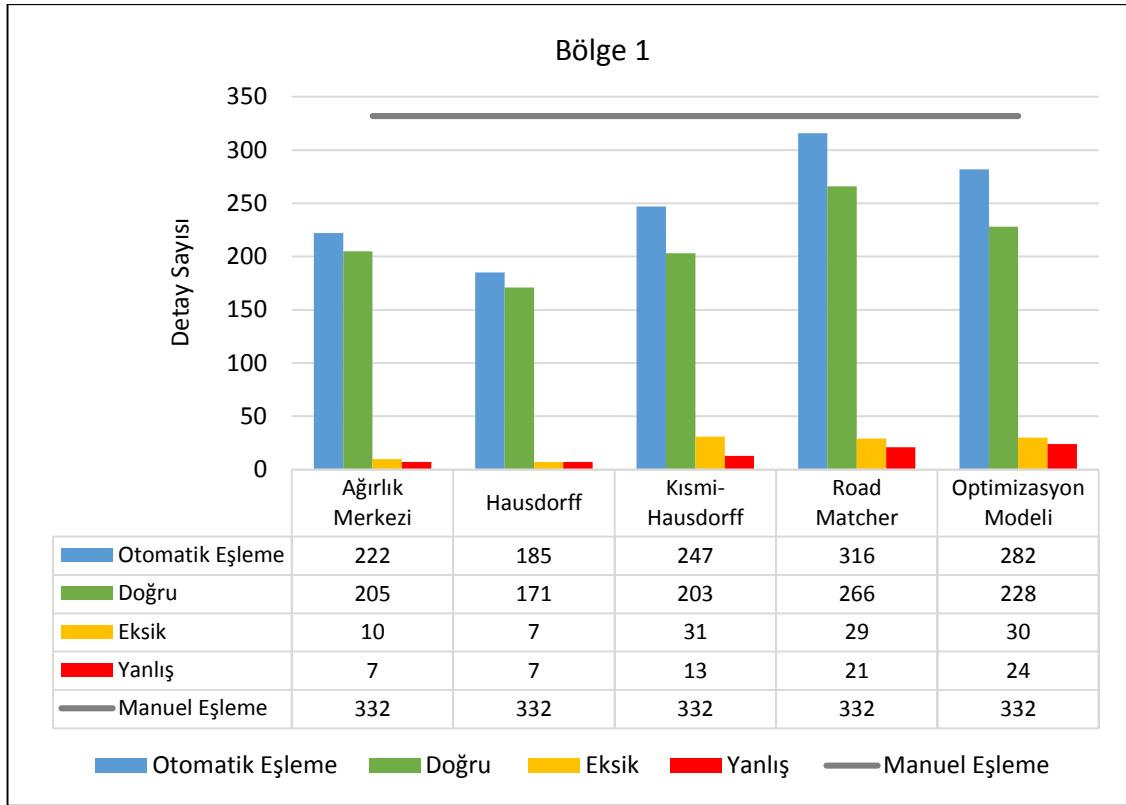
4.4 Sonuçların Değerlendirilmesi

Yol ağları ile eşleme yaparken, eşlemenin amacına göre elde edilen doğruluğun yeterliği değişkenlik gösterebilmektedir. Örneğin, ambulans, itfaiye, posta ya da okul servisi, taksi gibi araçlar, yol ağlarının görselleştirildiği navigasyon sistemlerini kullanırlar. Bu sistemler daima güncel ve güvenilir olmalıdır. Bu bağlamda, kullanılan eşleme yönteminin doğruluğunun navigasyon sisteminin gereksinimlerini karşılayacak seviyede olması gerekmektedir. Bu çalışmada yöntemlerin başarısı gerçekleştirdikleri doğru ve yanlış eşleme sayıları ile belirlenmiştir.

Birçok yol ağı eşleme yönteminde kullanılan Hausdorff ve Kısmi-Hausdorff mesafelerinin tüm alt bölgeler için verdiği sonuçlar tek başlarına kullanılamayacaklarını göstermektedir. Ancak, bu iki yöntemin (özellikle Hausdorff) yanlış sayısının az olması, birçok farklı yöntemde ana ölçütlerden biri olarak kullanılma nedenini açıklamaktadır (Çizelge 4.2, 4.3 ve 4.4).

Ağırlık merkezine dayalı eşleme yöntemi çizgi eşlemede çok kullanılan bir yöntem değildir. Ancak, yanlış sayısının Hausdorff yönteminden sonra en az olarak gerçekleştiği yöntem Ağırlık merkezine dayalı eşleme yöntemidir. Doğru sayısı bakımından da kısmi-Hausdorff yönteminden daha başarısız olmasına karşın Hausdorff yönteminden daha başarılıdır (Çizelge 4.2, 4.3 ve 4.4). Bu bağlamda, bir basit eşleme yöntemi için verdiği sonuçlar yeterli olmamakla birlikte, iki ağırlık merkezi arasındaki mesafe, farklı kriterler içeren algoritmalarda Hausdorff mesafesi gibi kullanılabilecek bir ölçüttür.

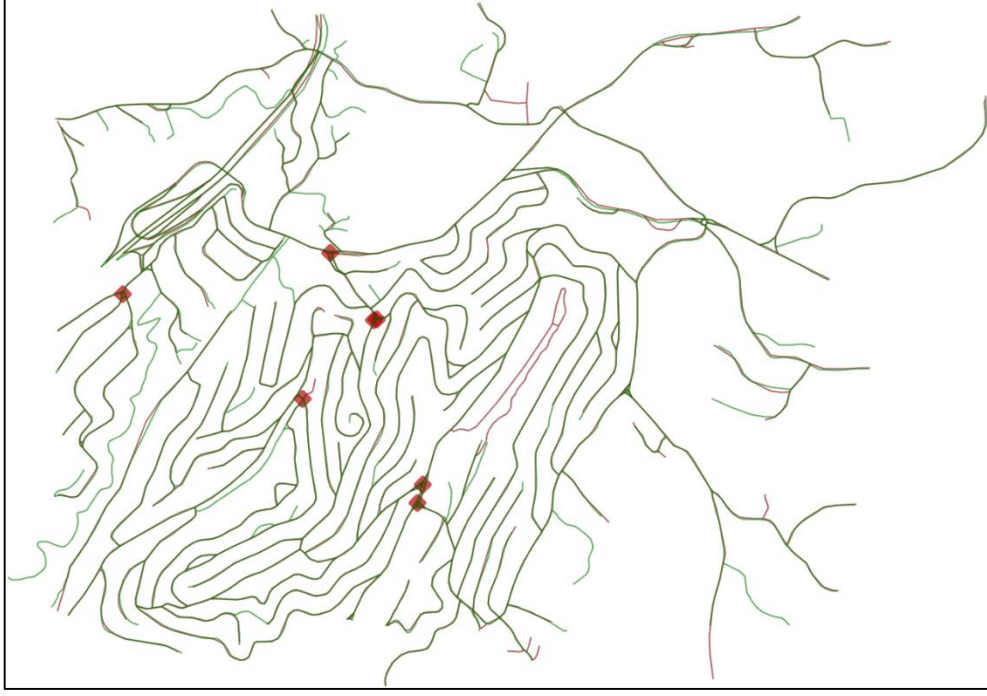
Ağırlık merkezine, Hausdorff ve Kısmi-Hausdorff mesafesine dayalı eşleme yöntemleri ile optimizasyon modeli ve RoadMatcher yöntemlerinin Bölge 1 için verdiği sonuçlar Şekil 4.19’da sütun ve çizgi grafik olarak verilmiştir. RoadMatcher ve optimizasyon modelinin diğer basit yöntemlere göre, doğru eşleme sayısı bakımından, daha başarılı olduğu; ancak, yanlış eşleme sayısı bakımından, daha başarısız olduğu görülmektedir. optimizasyon modeli ise RoadMatcher’e göre yaklaşık %12 daha az doğru eşleme yüzdesine sahiptir (Çizelge 4.8, 4.9, 4.10 ve 4.11).



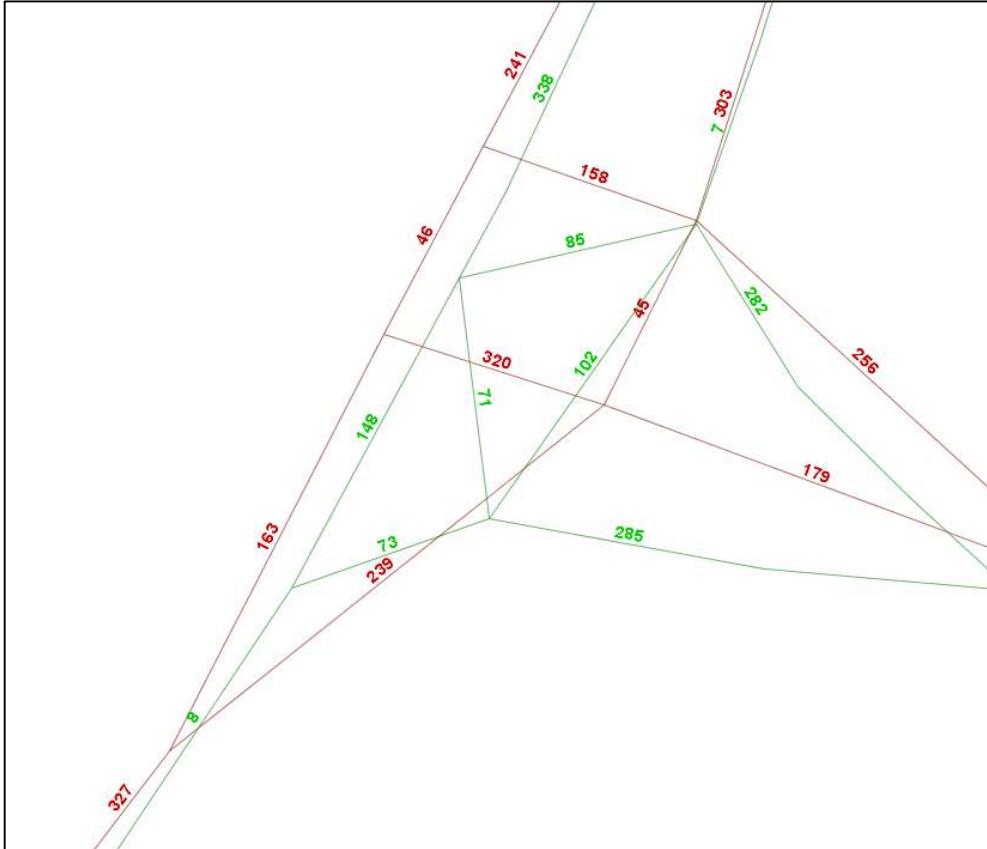
Şekil 4.19 Yöntemlere ait istatistiksel verilerin sütun ve çizgi grafik ile gösterimi

RoadMatcher doğru eşleme bakımından en başarılı yöntemdir. Ancak, yanlış ve eksik eşleme sayıları (toplam 50) ve eşlenmeyen sayısı (16) bakımından, kullanım amacına göre değişmekle birlikte, yeterli olmadığı açıkça görülmektedir (Çizelge 4.8 ve 4.9).

Bölge 1 için gerçekleştirilen tüm uygulamalarda yanlış eşlenen yollar aşağıda Şekil 4.20’de görülmektedir. Tüm yöntemler ile yanlış eşlenen bu yolların ortak özelliği her birinin kavşakları temsil eden çizgilerden biri olmasıdır. Şekil 4.21’de görüldüğü gibi Veriseti 1’de kavşağı oluşturan çizgiler bir üçgen iken, Veriseti 2’de dörtgendir. Gerek topolojisi gerekse geometrisi farklı olan bu yolların otomatik eşlenmesi oldukça zordur. Tüm yöntemlerin kavşakları oluşturan yolları eşlemedeki başarısızlığı, yeni bir yöntemin tasarımında buna dikkat edilmesi konusunda önem arz etmektedir.



Şekil 4.20 Bölge 1’de tüm yöntemlerin yanlış eşlediği yollar (kırmızı nokta işaretler)



Şekil 4.21 Veriseti 1 (yeşil) ve Veriseti 2’de (kırmızı) kavşağın farklı temsil edilmesi

SONUÇ VE ÖNERİLER

Bu tez çalışmasında farklı kaynaklardan gelen mekânsal verilerin entegrasyonunun gerekliliği, gerçekleştirme süreci ve en zor problemlerinden biri olan eşleme işleminin otomasyonu üzerine araştırmalar yapılmıştır.

Üretim hızı sürekli artan mekânsal verilerin kontrolünü sağlamak ve kısa sürede farklı amaçlar doğrultusunda analiz edebilmek için verilerin gerek birbirileri ile gerekse CBS'nin diğer bileşenleri ile uyumlu olması gerekmektedir. Uyumsuz, doğruluk bilgisi bulunmayan, kalitesiz, tekrarlı vb. veri üretimini engellemek için UMVA'nın tüm harita üreten kurumları kapsayacak şekilde veri üretim ve paylaşım politikalarını belirlemesi gerekmektedir. Birlikte çalışabilme, uyumlu mekânsal bilgilerin hızlı bir şekilde yorumlandığı bir sistem becerisidir.

Farklı kaynaklarda üretilmiş verilerin birlikte kullanımı veriler hakkındaki mevcut bilgilere, verinin kalitesine, uyumlu olmasına ve doğruluğuna bağlıdır. Bu nedenle, UMVA geliştirilirken entegrasyon süreçleri dikkate alınmalıdır. Örneğin, kurumların ürettikleri mekânsal verilerin geometrik ve öznitelik özelliklerinin (doğruluk, ölçek, projeksiyon, şema bilgileri vb.) belirgin olması, birçok entegrasyon sürecinin başarılı yönetilmesine olanak verecektir.

Mekânsal veri entegrasyonunun zorlu bir aşaması olan eşleme işleminin otomasyonu için UMVA'nın gerçekleştirilmesi önemli fakat yeterli değildir. Eşleme işleminin otomatik gerçekleştirilmesi için geliştirilen yöntemlerin çoğu, geliştirildikleri bölgenin karakteristik özelliklerine göre tasarlandıkları için veri bağımlıdır. Bu nedenle bir bölgedeki farklı

kaynaklı verisetlerini başarılı şekilde eşleyen bir algoritma başka bölgede daha az başarılı olabilmektedir.

Yol ağları ile eşleme yaparken doğru eşleme sayısı kadar yanlış eşleme sayısı da önemlidir. Bu bağlamda, uygulama sonunda bulgular yorumlanırken sadece doğru eşleme sayıları/yüzdeleri değil, bununla birlikte yanlış eşleme sayıları/yüzdeleri de dikkate alınmalıdır.

Uygulamada kullanılan yöntemler eşleme işlemini tamamıyla başarılı bir şekilde gerçekleştirememektedir. MatchingPlugin içindeki beş farklı yöntemin yol ağları gibi karmaşık yapıdaki nesnelerin eşlenmesi işleminde tek başlarına yeterli olmaları beklenmez. Mesafeye dayalı bu yöntemlerin verdikleri sonuçlar birbirleriyle karşılaştırılırsa, hangi mesafe kriterlerinin eşleme işlemi için kullanılabileceği hakkında yorum yapılabilir. Bölüm 4’de gerçekleştirilen uygulamada çizgiler arasındaki mesafelere dayalı eşleme yöntemlerinden, en iyi sonuç verenleri; Ağırlık merkezine dayalı, Hausdorff mesafesine dayalı ve Kısmi-Hausdorff mesafesine dayalı eşleme yöntemleridir. Bu yöntemlerle yapılan eşlemeler her ne kadar doğru sayısı bakımından yetersiz olsalar da, ortalama %1.9 yanlış eşleme oranı ile dikkat çekmektedir. Bu oran Hausdorff ve Kısmi-Hausdorff mesafelerinin niçin birçok eşleme algoritmasında kullanıldığının kanıtıdır. Ancak aynı durum çizgilerin ağırlık merkezleri arasındaki mesafeye göre gerçekleşen Ağırlık merkezine dayalı eşleme yöntemi için geçerli değildir. Hausdorff yöntemleri ile hemen hemen aynı sonucu vermesine karşın Ağırlık merkezine dayalı eşleme yöntemi literatürdeki algoritmalarda sıkça kullanılan bir kriter değildir. Elbette bunun önemli bir sebebi Hausdorff mesafesinin çizgiler arasındaki maksimum mesafeleri belirleyerek verisetleri arasındaki ölçek veya çözünürlük kaynaklı geometrik farklılığın ölçülmesini sağlamasıdır. Fakat mevcut bulgular Ağırlık merkezine dayalı yöntemin eşleme algoritmalarında kriter olabileceğini göstermektedir.

Tüm yöntemler kavşakları oluşturan yolları eşlerken başarısız olmuştur. Bunun temel nedeni, kavşaklardaki desen farklılıklarıdır. Örneğin, Şekil 4.21’de de görüldüğü gibi, bir verisetinde üçgen biçiminde gösterilen kavşak diğer verisetinde dörtgen biçiminde gösterilebilmektedir. Bu sebeple manuel eşleme yapılırken çoğu kavşakta 1-m eşleme yapılmıştır. Ancak otomatik eşleme için kullanılan RoadMatcher ve optimizasyon modeli

bu nesneleri başarıyla eşleyememiştir. Bulgular ileriki çalışmalarda geliştirilecek eşleme yöntemlerinin, kavşakların bu durumunun göz önünde bulundurularak tasarlanması gerektiğini göstermektedir.

Geometrik kriterlere dayanan yöntemlerle eşleme işlemini gerçekleştirirken yeterli eşleme sayısına ulaşamamıştır. Yöntemlerin daha iyi sonuçlar vermesi için geometrik kriterlerle birlikte topolojik ve öznitelik kriterlerin kullanılması gerekmektedir.

Sonuç olarak bu tez kapsamında yapılan çalışmalar, mekânsal veri entegrasyonunun daha profesyonel ve önceden belirlenmiş standartlara göre üretilen verisetleri ile yapılması gerektiğini göstermiştir. Ayrıca, uygulamada kullanılan yöntemlerin veri-bağımlı oldukları ve bu bağlamda yeni yöntemlerin geliştirilmesi gerektiği görülmüştür.

KAYNAKLAR

- [1] Al-Shahrani, A.M., (2002). Towards the Development of A Strategy For A National Spatial Data Infrastructure, Doktora Tezi, University of Nottingham, Nottingham.
- [2] Lynch, M.P. ve Saalfeld, A., (1985). "Conflation: Automated Map Compilation—A Video Game Approach", Auto-Carto VII, 11-14 Mart 1985, Washington, D.C.
- [3] Rosen, B. ve Saalfeld, A., (1985). "Match Criteria for Automatic Alignment" Auto-Carto VII, 11-14 Mart 1985, Washington, D.C.
- [4] Lupien, A. E. ve Moreland, W.H., (1987). "A General Approach to Map Conflation", Auto-Carto VIII, 29 Mart- 3 Nisan 1987, Baltimore.
- [5] Saalfeld, A., (1988). "Conflation: Automated Map Compilation", International Journal of Geographical Information System, 2(3): 217-228.
- [6] Deretsky, Z. ve Rdony, U., (1993). "Automatic Conflation of Digital Maps", In Vehicle Navigation and Information Systems Conference, Ekim 1993, Ottawa.
- [7] Gabay, Y. ve Doytsher, Y., (1995). "Automatic Feature Correction in Merging of Line Maps". In ACSM/ASPRS Annual Convention & Exposition Technical Papers, 404-411.
- [8] Cobb, M.A., Chung, M.J., Foley III, H., Petry, F.E., Shaw, K.B. ve Miller, H.V., (1998). "A Rule-Based Approach for the Conflation of Attributed Vector Data", GeoInformatica, 2(1): 7-35.
- [9] Sester, M., Anders, K.H. ve Walter, V., (1998). "Linking Objects of Different Spatial Data Sets by Integration and Aggregation", GeoInformatica, 2(4): 335-358.
- [10] Walter, V. ve Fritsch, D., (1999). "Matching Spatial Data Sets: A Statistical Approach", International Journal of Geographical Information Science, 13(5): 445-473.
- [11] Yuan, S. ve Tao, C., (1999). "Development of Conflation Components", The Proceedings of Geoinformatics'99 Conference, 19-21 Haziran 1999, Ann Arbor, 1-13.

- [12] Kang, H., (2002). Analytical Conflation of Spatial Data from Municipal and Federal Government Agencies, Doktora Tezi, the Ohio State University, Ohio.
- [13] Chen C.-C., Thakkar S., Knoblok C.A. ve Shahabi C., (2003) "Automatically Annotating and Integrating Spatial Datasets", in Proceedings of the International Symposium on Spatial and Temporal Databases, 2003, Santorini Adası.
- [14] Xiong, D. ve Sperling, J., (2004). "Semiautomated Matching for Network Database Integration", ISPRS Journal of Photogrammetry and Remote Sensing, 59(1): 35-46.
- [15] Samal, A., Seth, S. ve Cueto, K., (2004). "A Feature-Based Approach To Conflation of Geospatial Sources", International Journal of Geographical Information Science, 18(5): 459-489.
- [16] Zhang, M. ve Meng, L., (2007). "An Iterative Road-Matching Approach for The Integration of Postal Data", Computers, Environment and Urban Systems, 31(5): 597-615.
- [17] Mustière, S. ve Devogele, T., (2008). "Matching Networks with Different Levels of Detail", Geoinformatica, 12(4): 435-453.
- [18] Zhang, M., (2009). Methods and Implementations of Road-Network Matching. Doktora Tezi, Technical University of Munich, Münih.
- [19] Olteanu-Raimond, A.M. ve Mustière, S., (2008). "Data Matching—A Matter of Belief", Proceedings of the International Symposium on Spatial Data Handling, 23–25 Haziran 2008, Montpellier, 501-19.
- [20] Olteanu-Raimond, A.M., Mustière, S. ve Ruas, A., (2015). "Knowledge Formalisation for Vector Data Matching Using Belief Theory", Journal of Spatial Information Science, 10: 21-46.
- [21] Kim, J.O., Yu, K., Heo, J. ve Lee, W.H., (2010). "A New Method for Matching Objects in Two Different Geospatial Datasets Based on the Geographic Context", Computers & Geosciences, 36(9): 1115-1122.
- [22] Li, L., (2010). Design of A Conceptual Framework and Approaches for Geo-Object Data Conflation, Doktora Tezi, University of California, Santa Barbara.
- [23] Li, L. ve Goodchild, M.F., (2011). "An Optimisation Model for Linear Feature Matching in Geographical Data Conflation", International Journal of Image and Data Fusion, 2(4): 309-328.
- [24] Song, W., Keller, J.M., Haithcoat, T.L. ve Davis, C.H., (2011). "Relaxation-Based Point Feature Matching for Vector Map Conflation", Transactions in GIS, 15(1): 43-60.
- [25] Pourabdollah, A., Morley, J., Feldman, S. ve Jackson, M., (2013). "Towards An Authoritative Openstreetmap: Conflating OSM and OS Opendata National Maps' Road Network", ISPRS International Journal of Geo-Information, 2(3): 704-728.

- [26] Yang, W., Lee, D. ve Ahmed, N., (2014). "Pattern Based Feature Matching for Geospatial Data Conflation", GEOProcessing, Mart 2014, Barcelona.
- [27] Baillard C., Schmid C., Zisserman A. ve Fitzgibbon A. "Automatic Line Matching and 3D Reconstruction of Buildings from Multiple Views", In ISPRS Conference on Automatic Extraction of GIS Objects from Digital Imagery, Eylül 1999, 69-80.
- [28] Woo S., ve Dipanda A., (2000). "Matching Lines and Points in An Active Stereo Vision System Using Genetic Algorithms", Proc. IEEE-ICIP2000, Vancouver, 332-335.
- [29] Zhang, D. ve Shu, W., (1999). "Two Novel Characteristics in Palmprint Verification: Datum Point Invariance and Line Feature Matching", Pattern Recognition, 32(4): 691-702.
- [30] SDI Cookbook, 2004, Developing Spatial Data infrastructures: The SDI Cookbook, <http://www.gsdi.org/docs2004/Cookbook/cookbookV2.0.pdf>, 20.07.2015.
- [31] Masser, I., (2007). Building European Spatial Data Infrastructures, Birinci Baskı, Esri Press, Redlands.
- [32] FGDC, N. (2008). Annual Report Federal Geographic Data Committee, NSDI, Virginia.
- [33] FGDC, N. (2005). Annual Report Federal Geographic Data Committee, NSDI, Virginia.
- [34] Onah, C.C., (2009). Spatial Data Infrastructures Model for Developing Countries: A Case Study of Nigeria, Yüksek Lisans Tezi, Geoinformatics, University of Muenster, Münster.
- [35] GeoPlatform, Coğrafi veri portalı, <http://www.geoplatform.gov/>, 20.07.2015.
- [36] Avustralya Mekânsal Veri Dizini, <http://asdd.ga.gov.au/asdd/>, 22.07.2015.
- [37] INSPIRE Directive (2007). Directive 2007/2/EC of the EU Parliament and of the Council (14 March 2007) Establishing An Infrastructure for Spatial Information in the EU Community (INSPIRE), Official Journal of the European Union, 50.
- [38] T.C. Çevre ve Şehircilik Bakanlığı Coğrafi Bilgi Sistemleri Genel Müdürlüğü, INSPIRE Nedir?, www.csb.gov.tr/gm/cbs/index.php?Sayfa=sayfaicerikhtml&IcId=187&detId=188&ustId=187, 22.07.2015.
- [39] INSPIRE Documents, Commission Regulation 1089/2010 on Interoperability of Spatial Datasets and Services or New Version 1.2 of Technical Guidelines on INSPIRE Metadata, <http://inspire.ec.europa.eu/index.cfm/pageid/2>, 22.07.2015.
- [40] INSPIRE, View Information for the INSPIRE Themes, <http://inspire.ec.europa.eu/index.cfm/pageid/2/list/7>, 22.07.2015.
- [41] Güney, C., Doğru, A.Ö., Başaraner, M., Çelik, R.N. ve Uluğtekin, N., (2013). "TUCBS'ye Giden Yolum Neresindeyiz? Türkiye Ulusal Mekânsal Veri Altyapısı

İçin Nasıl Devam Etmeliyiz?”, TMMOB Coğrafi Bilgi Sistemleri Kongresi, 11-13 Kasım 2013, Ankara.

- [42] INSPIRE, (2010). 2010 Türkiye Genel Durumu, Yayın No: 8.1, Ankara.
- [43] CBS Gen. Müd., (2012). Türkiye Ulusal Coğrafi Bilgi Sistemi Standartlarının Belirlenmesi Projesi: TUCBS Kavramsal Model Bileşenleri, T.C. Çevre ve Şehircilik Bakanlığı; Aydınoglu A.Ç., Kara A. ve Yomralioglu T. tarafından düzenlendi, Ankara.
- [44] Tamash, N., (2012). INSPIRE Compliant Datasets Transformation & Conformance Testing, Yüksek Lisans Tezi, TU Delft, Delft.
- [45] Mohammadi, M., Binns, A., Rajabifard, A. ve Williamson, I., (2006). “Spatial Data Integration”. In Proceedings of the Seventeenth UNRCC-AP Conference and Twelfth Meeting of the PCGIAP, Bangkok.
- [46] Lunde, K., (2005). Geographical and Syntactical Integration of Geospatial Data, Yüksek Lisans Tezi, Østfold University College, Halden.
- [47] Michaud M., MatchingPlugin Tutorial for Version 0.7.2., http://sourceforge.net/projects/jump-pilot/files/OpenJUMP_plugins/More%20Plugins/Matching%20Plugin/MatchingPlugin0.7.2.pdf/download, 22.07.2015.
- [48] Vividsolutions, RoadMatcher Tutorial for Version 1.4, http://sourceforge.net/projects/jump-pilot/files/OpenJUMP_plugins/More%20Plugins/Roadmatcher%201.5%20for%20OJ1.7/RoadMatcher15forOJ17.zip/download, 22.07.2015.
- [49] OpenJUMP wiki, OpenJump Tutorial, http://ojwiki.soldin.de/index.php?title=Main_Page, 22.07.2015.
- [50] GNU GENERAL PUBLIC LICENSE, A Quick Guide, <http://www.gnu.org/licenses/quick-guide-gplv3.html>, 22.07.2015.
- [51] OGC, OGC Standards, <http://www.opengeospatial.org/standards/is>, 22.07.2015.
- [52] The JUMP Pilot Project, Plugin Download page, www.sourceforge.net/projects/jump-pilot/files/OpenJUMP_plugins/More%20Plugins/, 22.07.2015.

UYGULAMA SONUÇLARI

Çizelge A.1 Manuel eşleme (Bölge 1)

Veriseti1	Veriseti2
1	
2	
3	64
4	
5	335
6	242
7	303
8	327
9	137
10	215
11	62/95
12	302
13	297
14	263
15	314
16	294
17	206
18	34
19	124
20	194
21	50
22	233
23	27
24	336
25	35

Çizelge A.2 Ağırlık merkezine dayalı eşleme (Bölge 1)

Veriseti1	Veriseti2
1	
2	
3	
4	
5	335
6	242
7	303
8	327
9	137
10	215
11	
12	302
13	
14	263
15	314
16	294
17	206
18	
19	
20	
21	50
22	233
23	27
24	336
25	

Çizelge A.3 Hausdorff mesafesine dayalı eşleme (Bölge 1)

Veriseti1	Veriseti2
1	
2	
3	
4	
5	335
6	242
7	303
8	327
9	
10	215
11	
12	302
13	
14	263
15	314
16	
17	206
18	
19	
20	
21	50
22	233
23	27
24	336
25	

Çizelge A.4 Kısmi-Hausdorff mesafesine dayalı eşleme (Bölge 1)

Veriseti1	Veriseti2
1	
2	
3	
4	
5	335
6	242
7	303
8	327
9	137
10	215
11	95
12	302
13	
14	263
15	314
16	294
17	206
18	
19	
20	194
21	50
22	233
23	27
24	336
25	

Çizelge A.5 RoadMatcher (Bölge 1)

Veriseti1	Veriseti2
1	
2	
3	64
4	
5	335
6	242
7	303
8	327
9	421
10	215
11	
12	302
13	297
14	263
15	314
16	294
17	206
18	34
19	124
20	194
21	287
22	233
23	27
24	336
25	35

Çizelge A.6 Optimizasyon modeli (Bölge 1)

a=25 m; B=1.4		a=25 m; B=1.5		a=25 m; B=1.6	
Veriseti1	Veriseti2	Veriseti1	Veriseti2	Veriseti1	Veriseti2
1		1		1	
2		2		2	
3		3		3	
4		4		4	
5	335	5	335	5	335
6	242	6	242	6	242
7	303	7	303	7	303
8	327	8	327	8	327
9	137	9	137	9	137
10	215	10	215	10	215
11	95	11	95	11	95
12	302	12	302	12	302
13		13		13	
14	264	14	264	14	263
15	314	15	314	15	314
16	294	16	294	16	294
17	206	17	206	17	206
18		18		18	
19		19		19	
20	194	20	194	20	194
21	50	21	50	21	50
22	233	22	233	22	233
23	27	23	27	23	27
24	336	24	336	24	336
25		25		25	

ÖZGEÇMİŞ

KİŞİSEL BİLGİLER

Adı Soyadı : Müslüm HACAR
Doğum Tarihi ve Yeri : 1987, Bakırköy
Yabancı Dili : İngilizce
E-posta : mhacar@yildiz.edu.tr

ÖĞRENİM DURUMU

Derece	Alan	Okul/Üniversite	Mezuniyet Yılı
Lisans	Harita Mühendisliği	Yıldız Teknik Üniversitesi	2012
Lise	Fen Bilimleri	İstanbul Kemal Hasoğlu Lisesi (YDA)	2007

İŞ TECRÜBESİ

Yıl	Firma/Kurum	Görevi
2012-	Yıldız Teknik Üniversitesi/ İnşaat Fakültesi	Araştırma Görevlisi