



TÜBİTAK-2209-B SANAYİYE YÖNELİK LİSANS ARAŞTIRMA PROJELERİ DESTEĞİ PROGRAMI
ARAŞTIRMA ÖNERİSİ FORMU

2020

2209-B
1. Dönem Başvurusu

2209/B SANAYİYE YÖNELİK LİSANS ARAŞTIRMA PROJELERİ DESTEĞİ PROGRAMI
ARAŞTIRMA ÖNERİSİ FORMU

A. GENEL BİLGİLER

Araştırma Önerisinin Başlığı: Gerçek Zamanlı Akan Lojistik Verileri Kullanılarak Güzergah Takibi ve Sürücü Davranış Kalıpları Kapsamında Yasal Olmayan Durum ve Sürücülerin Tespiti
Başvuru Sahibinin Adı Soyadı: Ali EKEN Proje Ortakları: Furkan YILDIZ, Melike OĞUZ, Şeyma Nur MUTLU
Akademik Danışmanın Adı Soyadı: Doç. Dr. Ahmet SAYAR
Sanayi Danışmanının Adı Soyadı: Kaan Bayram, CEVA Lojistik ARGE ve Entegrasyon Grup Müdürü
Araştırmanın Yürütüleceği Kurum/Kuruluşlar: Kocaeli Üniversitesi, CEVA Lojistik

ÖZET

<p>Bu projenin amacı, lojistik verileri özelinde (spatio-temporal, uzaysal ve zamansal veriler), gerçek ve/veya yakın gerçek zamanlı olarak <u>akan veri</u> üzerinde, güzergah ve sürücü davranış kalıplarına bakılarak, yasal olmayan durum ve sürücülerin tespitidir. CEVA Lojistik Firmasının sağlayacağı, firma tarafından kullanılan Araç Takip Sistemlerinden toplanan veriler üzerinde, CEVA Lojistik firması tarafından sağlanacak bir fiziksel altyapı ve veri kullanılarak bir Akan Büyük Veri Analitik Platformu oluşturulacak ve bu altyapı üzerinde gerçek zamanlı olarak anormal güzergahlar ve anormal sürücü davranışları tespit edilmeye çalışılacaktır. CEVA Lojistik tarafından sağlanacak veride hukuki bir sorun oluşturacak hiçbir özel bilgi bulunmayacak, tüm kişisel bilgiler anonim hale getirilerek kullanılacaktır. Oluşturulacak çözümün, araç filosu işleten tüm kurum ve kuruluşlar için önemli bir ihtiyaç olduğu, servis tabanlı mimaride de bu tür bir hizmeti sağlayan girişime destek olabileceği değerlendirilmektedir.</p>
Anahtar Kelimeler: Lojistik Veri, Araç Takip Sistemi Verileri, OBD Veri Seti, Akan Büyük Veri, Gerçek Zamanlı Büyük Veri Analizi.

1. AMACI, YENİLİKÇİ YÖNÜ ve TEKNOLOJİK DEĞERİ

1.1. Projenin amacı

Bu projenin amacı;

Lojistik verileri özelinde (spatio-temporal, uzaysal ve zamansal veriler) gerçek ve/veya yakın gerçek zamanlı olarak akan veri üzerinde

- güzergah
- sürücü

davranış kalıplarına bakılarak yasal olmayan

- durumların
- sürücülerin

tespitidir.

Müteakip alt maddelerde proje amacına yönelik gerekli detaylar sunulmuştur.

1.1.1. Lojistik Verileri

Günümüzde perakende satışların büyük bir bölümü çevrimiçi olarak yapılmaktadır (1). Dünya çapındaki COVID-19 salgını çevrimiçi alışveriş miktarını ciddi miktarda arttırmıştır (2). Çevrimiçi ticaretin artması, lojistik şirketlerinin iş imkanlarını ve taşıma sektöründeki rekabeti arttırmaktadır.

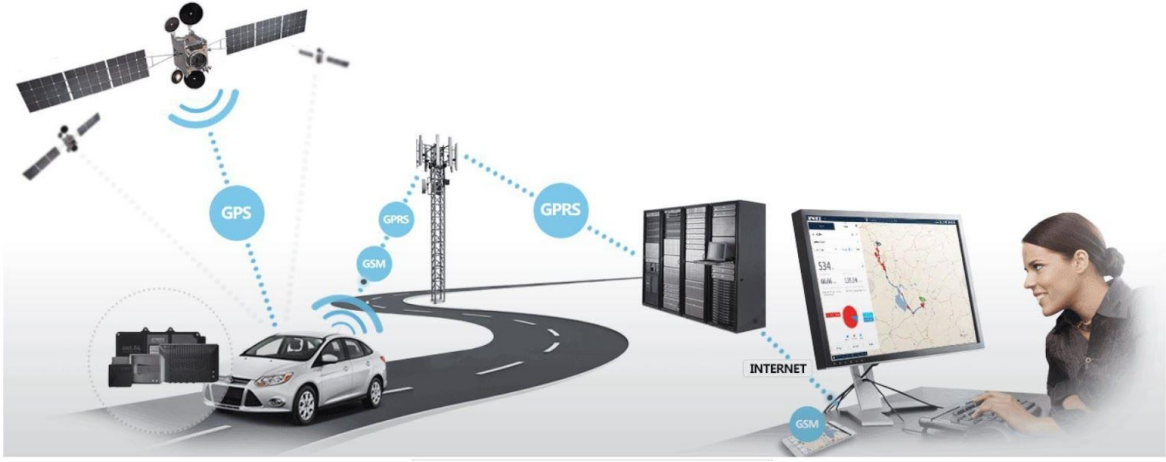
2209/B SANAYİYE YÖNELİK LİSANS ARAŞTIRMA PROJELERİ DESTEĞİ PROGRAMI ARAŞTIRMA ÖNERİSİ FORMU

Lojistik şirketleri, malların sevkiyatı için değişik büyüklüklerde yüzlerce kamyon/tır kullanmaktadır. Her türlü değer ve büyüklükte malların taşınması sağlanmaktadır. Söz konusu araç filolarındaki araçların yerlerinin ve şoför hareketlerinin takip edilebilmesi, lojistik şirketlerinin faaliyetlerini maliyet etkin yürütebilmeleri ve rekabet edebilmeleri için hayati öneme sahiptir. **Bu maksatla lojistik şirketleri araç takip sistemlerinden faydalanmaktadır. Bu çalışmada müteakip bölümlerde ifade edilecek araç takip sistemleri üzerinden elde edilen uzaysal ve zamansal (spatio-temporal veriler), lojistik verileri olarak ifade edilecektir.**

1.1.2. Araç Takip Sistemleri Üzerinden Toplanan Veriler

Araç takip sistemleri,

- Araç takip sistemi hizmetini veren bir şirket, (araç takip sistemi servis sağlayıcısı olarak isimlendirilecektir)
- Araç üzerine yerleştirilen bir cihaz, (araç takip cihazı olarak isimlendirilecektir)
- Araç takip cihazından bilgileri toplayan genellikle bulut ortamında çalışan bir altyapı,
- Araç takip cihazı ile söz konusu altyapının iletişimini sağlayacak bir ağ ortamından oluşmaktadır.



Şekil 1 Araç Takip Sistemi (3)

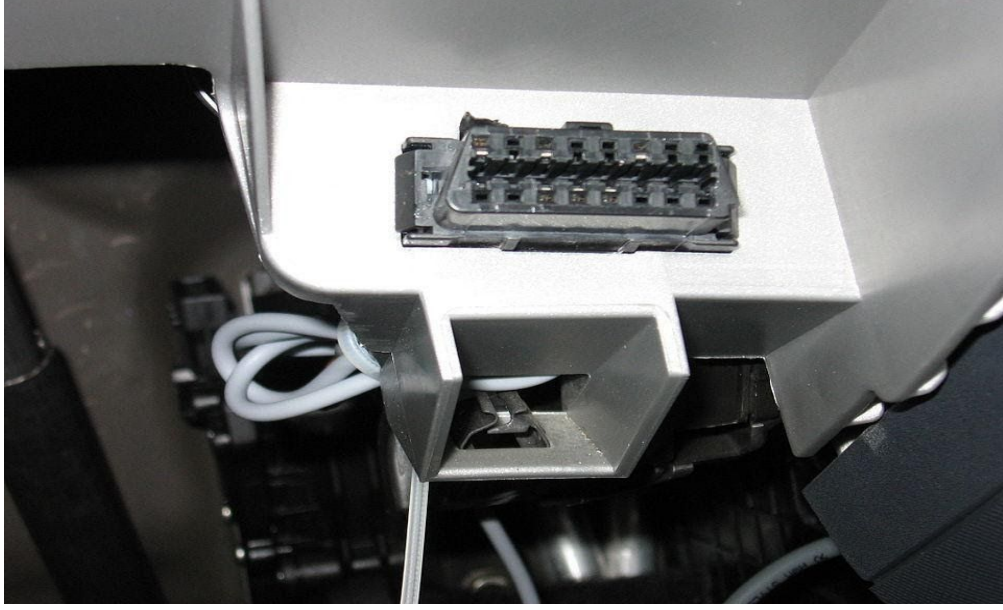
Ağ ortamı genellikle cep telefonu şebekeleridir, ancak farklı ihtiyaçlara göre uydu tabanlı global özel ağ sistemlerinin de kullanılması olasıdır.

Araç takip sistemi servis sağlayıcısı, araç takip cihazlarını, bu cihazlar vasıtası ile topladıkları verileri ve veri dağıtım sistemini pazarlamaktadır. Müşteriler, araç takip cihazlarını satın alabilmekte veya kiralayabilmektedir. Araç takip cihazından, verinin servis sağlayıcısına aktarılması için cep telefonu şebekesine dahil olması gerekmektedir. Bu kapsamda, araç takip sistemi cihazlarından GSM ağlarına dahil olabilmek için, aboneli olan GSM operatörlerinin SIM kartları takılmaktadır. GSM aboneliği yine müşterinin tercihinin göre kiralamalarda olduğu gibi servis sağlayıcısı tarafından sağlanabilmekte veya müşteri kendi kartlarını ve planını ilgili operatörden satın alarak kullanabilmektedir. Araç takip sistemi servis sağlayıcıları, araçlardan topladıkları veriyi ilgili firmaya ücret mukabilinde müşterinin tercihlerine ve fiyata göre değişik yoğunlukta ve detayda dağıtmaktadır.

Araç takip sistemleri üzerinden bulunduğu araca yönelik üç tür veri sağlamak mümkündür.

- Araç takip sistemi cihazlarında gömülü olarak GPS, GLONASS gibi Global Uydu Seyrüsefer Sistemlerinin çipleri/alıcıları bulunmaktadır. Böylelikle araç takip sistemleri aracın mevkisini anlık olarak takip ve rapor edebilmektedir.
- **Araç takip sistemleri genellikle araçlar üzerinden bulunan Gömülü Tespit Sistemi diyebileceğimiz "Onboard Diagnostic" OBD (4) sistemi ile iletişim halinde bulunmaktadır. OBD sistemi günümüz tüm güncel araçlarında bulunmaktadır.**

OBD temel olarak emisyon ölçümleri için tasarlanmış bir sistem olup, zamanla araçların daha fazla elektronik sistemler içermesi nedeni ile araç hakkında kritik parametrelerin araç üzerindeki algılayıcılardan toplandığı bir sisteme dönüşmüştür. Bu sistemin dışarıya veri aktarabilmek için portu bulunmaktadır. Bu porta takılacak bir OBD protokolünü anlayan adaptör ve yazılım ile araç iç sistemleri hakkındaki bilgilerin gözlenmesi ve hata tespiti vb. hususların yapılabilmesi mümkün olabilmektedir.



Şekil 2 OBD Bağlantısı (5)

Araç takip cihazı OBD portu üzerinden araç iç sistemlerine ait bilgileri gerçek zamanlı olarak toplayabilmektedir. Bu sayede aracın mevkisine ek olarak, yakıt durumu, motorun çalışma durumu, hareket bilgisi, ivmesi, hızı vb. bilgileri araçların üzerinde bulunan algılayıcılardan gelen verileri GSM ağı üzerinden servis sağlayıcı altyapısına gönderebilmektedir. Bu veriler, mevki verilerinden farklı olarak çok çeşitlidir ve yoğundur. Müşterinin ihtiyacına göre filtrelenerek, istenilen detay seviyesinde ve sıklıkta müşteriye gönderilmektedir.

- Araç takip sistemlerine ek algılayıcılar bağlamak mümkündür. Örnek olarak sıcaklığa duyarlı mal taşıyan kamyonlarda soğutucu bölümlere araç takip sistemleri ile irtibatlı algılayıcılar takılabilmektedir. Araç takip sistemleri, soğutucu bölmelerin sıcaklığını raporlayabilmektedir. Ek olarak, güvenlik maksatları ile sadece varış noktalarında yetkili kişilerce açılması gereken bölmelerin durum kontrolleri, benzer şekilde araç takip sistemlerine bağlı algılayıcılar ile yapılabilmektedir. **Bu çalışmamızda bu tür iş süreçlerine yönelik özel veriler ile ilgilenilmeyecektir, sadece mevki ve standart OBD üzerinde alınan veriler dikkate alınacaktır.**

Sonuç olarak araç takip sistemi, araç üzerinde bulunan ve aracın OBD portuyla irtibatlı araç takip cihazı vasıtası ile, otomatik olarak aracın mevkisi ve iç durumu hakkında kapsamlı veriler toplayarak, anlık olarak GSM şebekeleri üzerinden Araç Takip Sistemi Servis sağlayıcısına iletmektedir.

CEVA Lojistik araçlarından toplanan veriler bu proje maksatları ile kullanılacaktır. Söz konusu veride hukuki bir sorun oluşturacak hiçbir özel bilgi bulunmayacak, tüm kişisel bilgiler anonim hale getirilerek kullanılacaktır.

1.1.3 Araç Takip Sistemi Verilerinin Büyük Veri Niteliği

Mevki verilerine ek olarak özellikle OBD üzerinden toplanan verilerin, proje hedefi olan güzergah anormallik tespiti ve şoför davranışlarındaki farklılıkların tespiti analizlerine yönelik hususlar Yöntem bölümünde bahsedilecektir. Bu hedeflerin de ötesinde, söz konusu verilerin analizi genel manada, **gizli kalmış değerleri de ortaya çıkararak, pek çok açıdan ilgili kuruma katkı sağlama potansiyeli bulunmaktadır.**

Ancak gizli değerleri içeren veri setlerinin yüksek frekanslı olması, aktarım ve depolamanın yüksek maliyetli olması, analizinin yüksek bir uzmanlık gerektirmesi nedeni ile çoğu lojistik şirketi tarafından kullanılmamakta ve atılmaktadır. Söz konusu verinin; toplanmanın, depolamanın, işlemenin ve analizinin maliyeti nedeni ile büyük verinin hız, miktar, değişiklik, güvenilirlik ve değer gibi kriterleri göz önüne alındığında **büyük veri** kapsamına girdiği değerlendirilmektedir.

2209/B SANAYİYE YÖNELİK LİSANS ARAŞTIRMA PROJELERİ DESTEĞİ PROGRAMI ARAŞTIRMA ÖNERİSİ FORMU

1.1.4 Güzergah Anormallik ve Sürücü Davranış Analizi

Araç takip sistemleri ile toplanan mevki ve araç iç bilgileri ile aracın seyrettiği güzergahlara ve aracı kullanan şoförün davranış kalıplarına yönelik analizin yapılması, hali hazırda güncel bir araştırma konusu olup lojistik firmaları gibi pek çok araç filosu yöneten işletmelerde önemli bir ihtiyaç olarak ortaya çıkmaktadır. Bu konuya ilişkin olarak Yöntem bölümünde daha kapsamlı açıklama yapılacaktır.

1.1.5 Akan Veri Üzerinde Gerçek Zamanlı Çıkarım Yapılması

Mevki ve OBD verilerin araç bazında, yüksek sıklıkla toplanması, ileride analiz edilmek üzere saklanması, alınan verinin büyük veri niteliği taşıması nedeni ile maliyetlidir. Sık miktarda güncelleme, GSM şebekelerine yüksek abonelik ücretlerine neden olmaktadır. Tam olarak nasıl fayda sağlanacağı bilinmeden yüksek miktarda verinin saklanması da ciddi bir depolama maliyeti getirmektedir. Günümüzde, pek çok şirket, verilerini bulut ortamında tutmakta, depolama ve işlem gücü için para ödemektedir. Bu nedenle, aktif olarak kullanılmayan veri atılmaktadır.

Ancak gerçek zamanlı veya yakın gerçek zamanlı olarak söz konusu verilerin, **geldikçe analiz edilmesi ve geçmişe yönelik verinin ihtiyaç olduğu kadar tutulması**, söz konusu verinin tamamen çöpe atılmasından çok daha fayda sağlayacağı ve aynı zamanda verinin sürekli tutulmaya çalışılmasından daha verimli olacağı aşikardır.

Ek olarak; ***bir problemin varlığının tercihen, olmadan önce veya oluyorken tespit edilmesi, sonradan fark edilmesinden çok daha uygun maliyetli olduğu yine bilinen bir gerçektir.*** Lojistik verileri özelinde, araçların anormal güzergahlarda seyretmemesi, garip güzergahlara yönelmesinin anlık tespiti, araç sürücülerinin genel davranış paternlerinden çok daha farklı araç kullanması, örnek olarak aracın çalındığına, ciddi trafik problemlerine, sürücünün rahatsızlandığına, sürücünün araştırılmaya ihtiyaç duyulan bir problem içinde olduğuna yönelik, erken müdahale ile tedbir almayı sağlayabilecek bir ikaz sağlayabilir. **Zamanında doğru bir ikaz ve erken müdahale, çok maliyetli sonuçların oluşmasının engellenmesine veya hasarın azaltılmasına, gecikmelerin önceden tespit edilerek tedbir alınmasına katkı sağlayacaktır.**

Veri tabanlarında bulunan verilerin analizi gibi durağan veri üzerinde analiz yapmaktansa, **akan veri üzerinden analiz hali hazırda önemli bir araştırma konusudur ve bu yönde yoğun akademik çalışmalar bulunmaktadır. Bu çalışmaları destekleyecek yazılım ve altyapı araçları süratli bir şekilde gelişmekte ve yaygınlaşmaktadır (6).**

1.2. Yenilikçi Yönü ve Teknolojik Değeri

Mevcut durumda lojistik firmalarına böyle bir hizmet sağlayan bir altyapı veya servis bulunmamaktadır. Bu çalışma, **her birisi kendi alanında başlı başına araştırma sahası olan büyük veri, akan veri analitiği, güzergah anormallik analizi, sürücü davranış analizi ve tüm bu araştırmaların yapılacağı güncel bir büyük veri analitik platformunu ve araçlarını içermektedir.**

Bu çalışmanın ortaya çıkaracağı sistem ve içindeki analiz yaklaşımları ile

- Genelde ayrı birer araştırma alanı olarak kabul edilen güzergah anormallik tespiti (anomaly trajectory detection) ile normal olmayan sürücü davranışları (anomaly driver behaviour detection) eşgüdüm içinde ve birbirini destekler mahiyette entegre bir çözüm oluşturulmaya çalışılacaktır. Her iki konuyu bünyesinde bulunduran bir çalışma tespit edilememiştir.
- Bu çözüm, gerçek zamanlı veya yakın gerçek zamanlı olarak dağıtık sistem mimarisi ve büyük veri altyapısı üzerinde sağlanacaktır. Problemin mahiyetine uygun genel olarak açık kaynak kodlu araçlardan oluşan bir altyapı kurulması hedeflenmektedir.

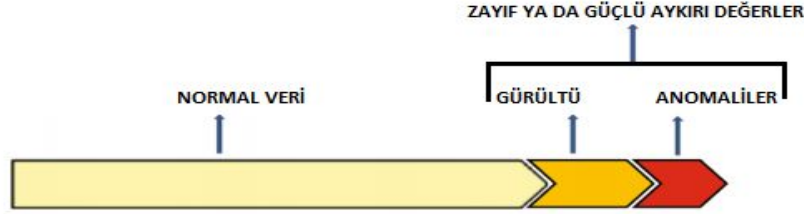
2. YÖNTEM

2.1 Anormallik Analizi

Anormallik tespiti, bilinmeyen aramak olarak ifade edilmektedir [4]. Genel olarak veri analizinde, analiz ettiğimiz verinin kaynağı olan sistemi ifade eden bir modelimiz vardır. Modelimiz ve bu modele uyan veriler normal sistem davranışlarını ifade etmektedir. Anormal veri, sistemin durumundaki bir anormalliğe, aykırılığa veya sistem hakkında modelimiz tarafından bilinmeyen yeni duruma işaret edebilmektedir. Anormallik tespiti, önceden görülmeyen, sistemin normal davranışının dışında olduğu değerlendirilen durumların tespit edilmesidir [5]. Tespit edilen durumlar analiz edilerek, tespit edilen husus eğer bu ana kadar keşfedilmemiş bir sistem davranışı ise

2209/B SANAYİYE YÖNELİK LİSANS ARAŞTIRMA PROJELERİ DESTEĞİ PROGRAMI ARAŞTIRMA ÖNERİSİ FORMU

modele eklenerek modelin güncellenmesi sağlanır, değilse buna neden olan durumun oluşmasına yönelik gerekli tedbirler alınır. Bu nedenle, anormal veriler genellikle sistemin durumu hakkında önemli bilgiler sunmaktadır [6].



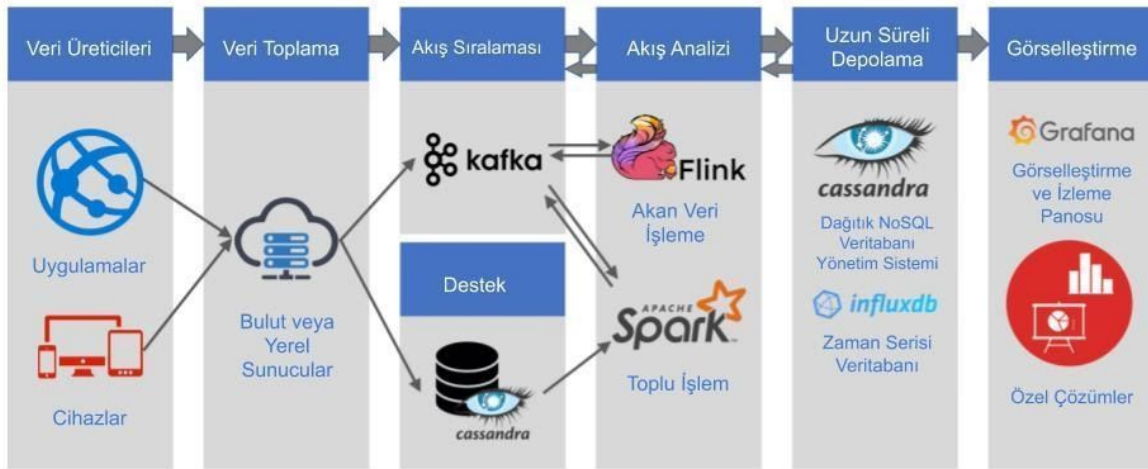
Şekil 4 Anormallik Analizi

2.2 Akan Büyük Veri Analitik Platformu

İlk aşamada çalışmaların yapılabilmesi için CEVA Lojistik tarafından tahsis edilecek sunucu(lar) üzerinde bir Akan Büyük Veri Analitik Altyapısı oluşturulacaktır. Nihai altyapının aşağıdaki şekle uygun olarak şekilleneceği öngörülmektedir.

CEVA Lojistik tarafından bilindiği kadarı ile hali hazırda Araç Takip Sistemi Verileri geçmişe yönelik olarak tutulmamakta, aracın en son bulunduğu mevki gibi sadece en güncel veri tutulmaktadır. Bu çalışma kapsamında, söz konusu veriler CEVA'dan gerçek zamanlı olarak geldikçe alınacak, Apache Cassandra gibi bir büyük veri depolama ortamında kaybolmamaları için araştırma maksatları ile yedeklenerek tutulacak ve aynı zamanda diğer servislerin erişimi için Apache Kafka veri simsarcısı benzeri bir ortama aktarılacaktır. **Gelen verinin tutulması, akan veri üzerindeki analizlerin, durağan veri ile teyit edilmesi ve desteklenmesi içindir.** Apache Kafka yayınla-abone ol kalıbı kapsamında sistem içindeki veri akışını yönetecektir. Müteakip bölümde belirtilecek analizler gerçek zamanlı olarak Apache Flink ve/veya Apache Spark Analiz platformları ile yapılacaktır. Apache Flink/Spark gerçek zamanlı olarak Kafka üzerinden aldıkları verileri analize tabi tutacak, Apache Cassandra üzerinde bulunan geçmişe ait durağan veriler kullanılan algoritmaların ihtiyacı kapsamında kullanılacak, sonuçlar yine müteakip analizlerde kullanılmak üzere aynı ortamda depolanacaktır. Özetle aynı ortamda ham veriler ve işlenmiş veriler bulunacaktır.

Sonuçların kullanıcı etkileşimine imkan sağlayan Coğrafi Bilgi Sistemi üzerinde çalışan bir arayüz sağlanması öngörülmektedir.



Şekil 4 Uygulanması Planlanan Akan Veri Analitik Platformu

Apache Cassandra üzerinde toplanan ham veriler ayrıca çevrimdışı (durağan veri) olarak veri analizine tabi tutulacak, böylelikle gerçek zamanlı analizin başarısı sayısal olarak ortaya çıkarılacaktır.

2209/B SANAYİYE YÖNELİK LİSANS ARAŞTIRMA PROJELERİ DESTEĞİ PROGRAMI
ARAŞTIRMA ÖNERİSİ FORMU

2.3. Analiz

2.3.1 Veri Seti

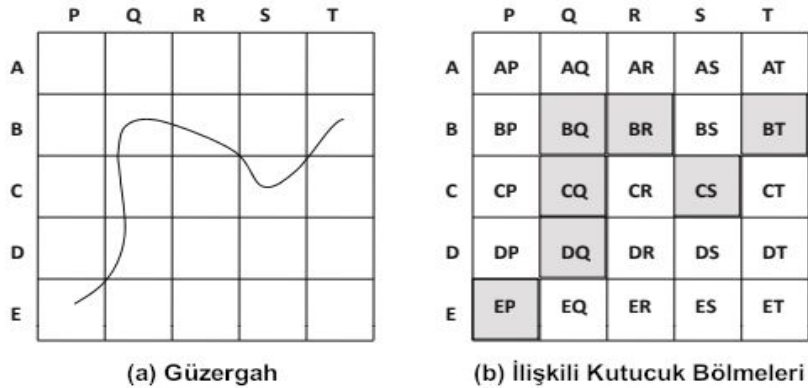
CEVA Lojistikten alınacak veri setinin bazı alanları aşağıda sunulmuştur.

VERİ	Açıklama
GPS_ID	Bilginin hangi GPS'den geldiğini ifade eder.
DEVICE_NO	Araca ait numara bilgisini tutmaktadır.
LICENSE_PLATE	Aracın plakası bu alanda kayıt edilmektedir.
DRIVER	Sürücünün kimliğidir. Her sürücü için farklıdır. Anonim hale getirilecektir.
DATE_TIME	Veri kaydının zaman bilgisidir.
SPEED_KM_H	Aracın anlık hız bilgisidir.
DISTANCE	Teslimat mesafesini vermektedir.
ADDRESS	Aracın anlık adresini vermektedir.
LATITUDE	Sürücünün anlık enlem bilgisidir.
LONGITUDE	Sürücünün anlık boylam bilgisidir.
GEOGRAPHICAL_REGION	Aracın bulunduğu konumun coğrafi bölge bilgisidir.
PAUSE_DURATION	Aracın duraklama süresini belirtmektedir.
IDLING_DURATION	Aracın boşta kalma süresini belirtmektedir.
IGNITION_DURATION	Aracın ateşleme süresini belirtmektedir.
USER_ID	Kullanıcı id bilgisidir.
INSERT_DATE	Sefer giriş tarihi bilgisini vermektedir.
UPDATE_DATE	Sefer güncelleme tarihi bilgisini vermektedir.

2.3.2. Anormal Güzergah Tespiti

Anormal güzergah tespitine yönelik kapsamlı çalışmalar bulunmaktadır. Söz konusu çalışmalar istatistiksel yöntemlerden makina öğrenmesine kadar çok farklı metotları kullanmaktadır. Örnek olarak, (7)'te izolasyon tabanlı bir metot önerilmektedir. Bu metotla, başlangıç ve bitiş noktalarına göre güzergahlar gruplanmakta, az ve farklı olanlar bazı istatistiksel ağırlık hesaplamalarına göre anormal olarak atanmaktadır. (8)'de başlangıç ve bitiş noktaları yerine alt güzergahlar dikkate alınmaktadır. Anormallikler alt güzergahlar da sürücünün ani hareketleri ve ivmelenmelerine bakılmaktadır.

Güzergah takibinde genelde kullanılan bir yaklaşım aşağıda sunulmuştur. WGS84 koordinatları, matematiksel modellerde daha kolay kullanım sağlayacak özelleştirilmiş koordinat sistemlerine dönüştürülmektedir. Şekil 2'de (a) seçeneğinde belirtilen güzergah bir grid tabanlı koordinat sistemine dönüştürülmüştür. (b)'de gösterilen kutucuklar genel olarak koordinatları göstermektedir.



Şekil 3 Güzergah çıkarılması (9)

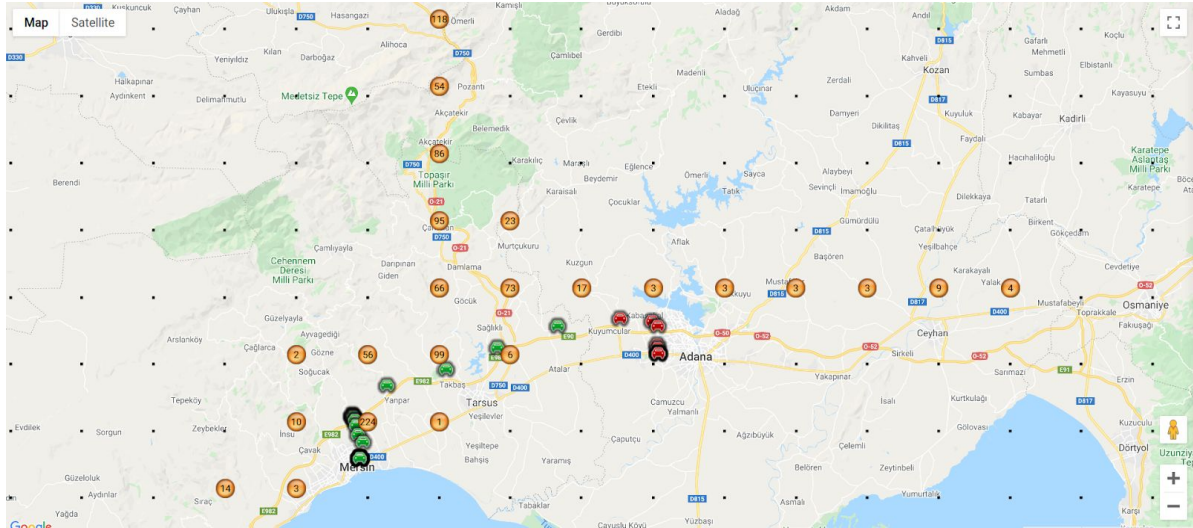
CEVA veri setinin kullanımında ana güzergahların Türkiye'yi 3-6 km civarında kenarları olan hücreleri içeren bir grid sistemine çevirerek takip edilebileceği değerlendirilmektedir. Aşağıdaki 6 km'lik grid uygulaması gösterilmektedir.

2209/B SANAYİYE YÖNELİK LİSANS ARAŞTIRMA PROJELERİ DESTEĞİ PROGRAMI ARAŞTIRMA ÖNERİSİ FORMU



Aşağıdaki örnek bir güzergah üzerinde anormallik tespit yaklaşımı sunulmuştur. Söz konusu yaklaşım henüz yayınlanmış taslak olan bir çalışmadan alınmıştır. Sarı noktalar grid merkezlerini, içindeki sayılar ise bu gridden seçilen güzergah kapsamında kaç aracın daha önceden geçtiğini belirtmektedir.

Bu yaklaşıma göre belirli bir eşik değeri dikkate alınarak, çok fazla tercih edilen yoldan geçen araçlar normal güzergahta kabul edilmekte olup yeşil ile gösterilmiştir. Diğer taraftan çok az kullanılan yoldan geçen araçlar ise kırmızı ile işaretlenmiştir.



2.3.3. Anormal Sürücü Davranışı Tespiti

(10)'de Akıllı telefonlar üzerindeki GPS verileri kullanılarak 50000 km boyunca, 37 sürücü, iki ay süresince takip edilmiştir. Söz konusu çalışmada ustalık, sürüş hızı veya manevralarının kullanıcıları birbirlerinden ayırdığı ifade edilmektedir. Bu çalışmada OBD verileri gibi destekleyici veriler kullanılmamış, sadece GPS verisi kullanılmıştır. Ek belirleyici veriler olmadan %82 civarında bir başarı yakalandığı ifade edilmektedir.

(11)'de OBD verileri kullanılarak yapılmıştır. Aynı araç, aynı güzergah ve 10 farklı sürücü ile 1 ay boyunca denemeler yapılmış, OBD tarafından sağlanan 60 küsur verinin istatistiksel olarak sürücülerini tanımlamada etkileri olup olmadıklarına bakılmış, 12 civarında verinin sürücülerini tanımlamada çok etkin olduğu tespit edilmiştir. Bu

2209/B SANAYİYE YÖNELİK LİSANS ARAŞTIRMA PROJELERİ DESTEĞİ PROGRAMI ARAŞTIRMA ÖNERİSİ FORMU

çalışmada değişik karar ağacı yöntemleri kullanılarak %80 üzerinde doğrulukla sürücü kimlikleri OBD verilerine bakılarak tespit edilebilmiştir.

Bu proje kapsamında (11)'de 60 civarındaki OBD verilerinden 6 tanesinin kullanıcı davranışı belirleme konusunda önemli etkisi olduğu ifade edilmektedir. Bu veriler aşağıdaki görselde sunulmuştur. Elde edilen bilgiler kapsamında söz konusu verilerin, karar ağacı (CART algoritması) kullanılarak aracı kullanan şoförün tanımlanabileceği değerlendirilmektedir.

```
Feature: TYPE, Score: 0.04475
Feature: SPEED_KM_H, Score: 0.00829
Feature: DISTANCE, Score: 0.31559
Feature: ADDRESS, Score: 0.46761
Feature: LATITUDE, Score: 0.07739
Feature: LONGITUDE, Score: 0.08637
Feature: GEOGRAPHICAL_REGION, Score: 0.00000
Feature: PAUSE_DURATION, Score: 0.00000
Feature: IDLING_DURATION, Score: 0.00000
Feature: UPD_USER_ID, Score: 0.00000
Feature: STATUS, Score: 0.00000
```

2.3.4. Çok Modelli Yaklaşım ve Sahadan Verilerin Kullanımı

Bu araştırma kapsamında güzergah anormallik tespiti ve sürücü davranışlarındaki anormallik yöntemlerinin bir arada kullanılması, sürücü yönelik olarak mevki ve OBD verilerinin bir arada kullanılması hedeflenmektedir. Ayrıca referans verilen çalışmalar, ideal kontrollü ortamlar ve veri setleri kullanmıştır.

3. PROJE YÖNETİMİ

3.1. İş- Zaman Çizelgesi

Aşağıda sayısal başarı ölçütü verilememiştir çünkü başarı ölçütü veri setlerinin dengeli olmasını ve elimizde olmasını gerektirmektedir. Ayrıca kullanılacak metodolojiler nihai olarak seçilmemiştir. Literatür taraması neticesinden seçilecektir .

İŞ-ZAMAN ÇİZELGESİ

No	İş Paketlerinin Adları ve Hedefleri	Kim(ler) Tarafından Gerçekleştirileceği	Zaman Aralığı (.... Ay)	Başarı Ölçütü ve Projenin Başarısına Katkısı
1	Akan Veri Analitik Platformunun Tesisi Literatür Taraması Araç Takip Sistemi Verilerinin Kayıt Edilmeye Başlanması	Ali EKEN, Furkan YILDIZ, Melike OĞUZ, Şeyma Nur MUTLU	1 Ay	Çalışmaların yapılabilmesi için bu altyapıya ihtiyacı vardır. Kritik ihtiyaçtır.

2209/B SANAYİYE YÖNELİK LİSANS ARAŞTIRMA PROJELERİ DESTEĞİ PROGRAMI
ARAŞTIRMA ÖNERİSİ FORMU

2	Anlık olarak anormal güzergahların tespitine yönelik seçilen/geliştirilen metodolojilerin kodlanması ve kayıtlı verilerin, akan veri olarak simüle edildiği ortamda denenmesi	Ali EKEN, Furkan YILDIZ	3 Ay	Gerçek zamanlı olarak normal ve anormal güzergahların sergilenmesi, algoritma başarımının ölçülmesi
	Anlık olarak farklı sürücü davranışlarının tespitine yönelik seçilen/geliştirilen metodolojilerin kodlanması ve kayıtlı verilerin, akan veri olarak simüle edildiği ortamda denenmesi	Melike OĞUZ, Şeyma Nur MUTLU	3 Ay	Gerçek zamanlı olarak beklenen sürücü veya beklenmeyen sürücü olarak araçların markalanması, algoritma başarımının ölçülmesi
3	İki kabiliyetin birbirini destekler biçimde eş güdümlü çalışmasının sergilenmesi	Ali EKEN, Furkan YILDIZ, Melike OĞUZ, Şeyma Nur MUTLU	Proje Sonuna kadar	Anormal güzergahı takip eden bir aracın şoförü için otomatik inceleme yapılabilmesi, farklı davranış sergileyen şoför için otomatik güzergah anormalliliğinin incelenmesinin başlatılabilmesi
4	Proje sonuçlarının raporlanması ve yayımlanması	Ali EKEN, Furkan YILDIZ, Melike OĞUZ, Şeyma Nur MUTLU	20 Gün	Konu hakkında makale yayımlanması

1.2. Risk Yönetimi

Araştırmanın başarısını olumsuz yönde etkileyebilecek riskler ve bu risklerle karşılaşılması durumunda araştırmanın başarıyla yürütülmesini sağlamak için alınacak tedbirlerle (B Planı) ilgili iş paketleri ana hatlarıyla belirtilerek aşağıdaki Risk Yönetimi Tablosu'nda ifade edilir. B planlarının uygulanması, araştırmanın temel hedeflerinden sapmaya yol açmamalıdır.

RİSK YÖNETİMİ TABLOSU*

No	En Büyük Riskler	Risk Yönetimi (B Planı)
1	CEVA lojistiğin birden fazla Araç Takip Sisteminde hizmet aldığı öğrenilmiştir. Farklı servis sağlayıcılarından gelen OBD verilerinin aynı olması ve birbiriyle uyumlu olmamasının sorun çıkarabileceği değerlendirilmektedir.	Sadece belirlenen servis tarafından rapor edilen araçlara yönelik çalışma daraltılabilir.
2	Araç takip sistemlerinden gelen veri sıklığının servisten servise değişebileceği, veri akış frekansını artırmanın maliyet getireceği öğrenilmiştir.	Güncelleme zamanının uzun olması durumunda, sık güncelleme gerektirmeyen, şehirler arası hatlara yönelik çalışma daraltılabilir.

(*) Tablodaki satırlar gerektiği kadar genişletilebilir ve çoğaltılabilir.

3.3. Araştırma Olanakları

Bu bölümde projenin yürütüleceği kurum ve kuruluşlarda bulunan ayrıca projede kullanılacak olan

2209/B SANAYİYE YÖNELİK LİSANS ARAŞTIRMA PROJELERİ DESTEĞİ PROGRAMI
ARAŞTIRMA ÖNERİSİ FORMU

altyapı/ekipman (laboratuvar, araç, makine-teçhizat vb.) olanakları belirtilir.

ARAŞTIRMA OLANAKLARI TABLOSU (*)

Altyapı/Ekipmanın Bulunduğu Kuruluş	Kuruluşta Bulunan Altyapı/Ekipman Türü, Modeli (Laboratuvar, Araç, Makine-Teçhizat vb.)	Projede Kullanım Amacı
CEVA Lojistik	CEVA Lojistik bünyesinde bulunan birkaç iş istasyonu veya sunucuyu bu çalışmaya tahsis edecek ve uzaktan erişim için gerekli internet hattını sağlayacaktır.	Akan Büyük Veri Analitik Platformu için gereklidir. Algoritmalar bu makinede çalıştırılacak ve veriler bu makinede saklanacaktır.
CEVA Lojistik	CEVA Lojistik Araç Takip Sisteminden gelen verilerin aktarılması için gerekli teknik/yazılım desteği verecektir.	Verilerin kurulan Akan Büyük Veri Analitik Platformuna aktarılması için ihtiyaç bulunmaktadır. Firma tarafından kurulan analitik platforma gelen veriler yönlendirilecektir.
CEVA Lojistik	CEVA Lojistik tarafından sağlanacak Araç Takip Sistemi Verileri	Güzergah ve Sürücü Davranış Analizi için kullanılacaktır. Kullanıcı adı gibi özel bilgiler anonim hale getirilecektir.

4. SANAYİ ODAKLI ÇIKTILAR ve YAYGIN ETKİLER

CEVA Lojistik firmasına ait kapsamlı araç filosunun anlık durumları hakkında işletme personelinin farkındalığı arttırılacak, elde edilen farkındalıkla maliyet uygun hale gelecek ve rekabet gücü artacaktır. Örnek olarak güzergahlardaki anormal değişiklikler fark edilecek, ek olarak güzergah analizi ile paralel olarak aracın tahsisli şoförü tarafından kullanılıp kullanılmadığı veya şoförün kullanım usullerinde farklılıkların olup olmadığı anlık tespit edilebilecektir. Diğer taraftan yapılan anlık analizler, geçmiş veri ile daha kapsamlı analizlerin yapılmasına rehberlik edecektir. Örneğin sistem, zaman içindeki anormallik tespitlerini (yeni yol yapımı ve güzergahlar oluşması vb.) ek bir güncelleme ihtiyacı olmadan normalleştirecektir. Bu kapsamda eski yoldan gitmek zaman içinde anormal bir davranış haline gelebilecektir.

Her ne kadar CEVA Lojistik tarafından sağlanacak veri seti üzerinde çalışılacak olsa da, **firmadan bağımsız olarak herhangi bir araç filosu işleten tüm kurum ve kuruluşların yararlanabileceği, gelişime açık, akademik gelişmeler ile analiz teknikleri geliştikçe güncellenebilen, hizmet ettiği kurumun ihtiyaçlarına yönelik olarak genişleyebilen (scalability), dağıtık mimaride çalışan bir sistem oluşturulacaktır. Özetle algoritmik yaklaşımlar kendi içinde servis mantığı ile çalışacak, yeni yaklaşımlar ek bir düzenleme getirmeden sisteme dahil edilip çıkarılabilecek veya başarımlar için paralel çalıştırılarak konsensüs tabanlı bir sisteme evrilmesi sağlanabilecektir.**

Özellikle araştırma sonucunda oluşturulacak prototip sistem ulusal ve uluslararası lojistik firmalarına hizmet verebilen, bulut mimarisi üzerinde çalışan, ülke ekonomimize ihracat katkısı sağlayabilecek bir ürüne dönüşebilme potansiyeli de bulunmaktadır.

5. BELİRTMEK İSTEDİĞİNİZ DİĞER KONULAR

Her ne kadar, oluşturulacak çözüm ve sistem, kara filolarına özel bir çözüm olarak gözükse de, herhangi bir takip cihazı taşıyan hava, deniz araçlarına uygulanma potansiyeli bulunmaktadır. Anormal güzergah tespitinde kullanılan yaklaşımlar, özellikle deniz araçlarında güncellenerek uygulanabilir. 300 Groston'dan büyük deniz araçları bir tür gemi takip sistemi olan Otomatik Takip Sistemi barındırmak zorundadır. Bu cihazlar gemi adı, bandıra, çağrı adı, kalkış, varış limanları gibi gemilerin temel bilgilerini ve mevkilerini rapor etmektedir. Bu bilgiler, çeşitli resmi ve özel kurumlar tarafından toplanmakta ve servis edilmektedir. Benzer yaklaşımların altyapıda hiç değişikliğe gidilmeden deniz araçları içinde kullanılabileceği öngörülmektedir.

6. EKLER

6.1. Kaynaklar

1. Brick and Mortar Stores vs. Online Retail Sites. [Çevrimiçi] <https://www.thebalancesmb.com/compare-brick-and-mortar-stores-vs-online-retail-sites-4571050>.
2. COVID-19, The Ultimate Wake-up Call for Bricks and Mortar Stores. [Çevrimiçi] <https://www.aoe.com/en/blog/covid-19-the-ultimate-wake-up-call-for-bricks-mortar-retail.html>.
3. <https://www.arvento.com/i/assets/images/content/howfinal.gif>
4. On-board Diagnostic. [www.wikipedia.org](https://en.wikipedia.org/wiki/On-board_diagnostics). [Çevrimiçi] https://en.wikipedia.org/wiki/On-board_diagnostics.
5. http://antalyaguvenliksistemleri.com/uploads/files/Ara%C3%A7%20Takip%20Sistemi%20Nedir_antalyaguvenlik_sistemleri_com.jpg
6. András A. Benczúr, Levente Kocsis, Róbert Pálovics. Online Machine Learning in Big Data Streams. [Çevrimiçi] 16 Feb 2018. <https://arxiv.org/abs/1802.05872>.
7. IBAT: Detecting Anomalous Taxi Trajectories from GPS Traces. Zhang, Daqing, Nan Li, Zhi-Hua Zhou, Chao Chen, Lin Sun, and Shijian Li. Beijing, China: ACM Press, 99. Cilt In Proceedings of the 13th International Conference on Ubiquitous Computing - UbiComp '11, 99. Beijing, China: DOI <https://doi.org/10.1145/2030112.2030127>.
8. Inferring Drivers Behavior through Trajectory Analysis. Eduardo M. Carboni, Vania Bogorny. s.l. : Intelligent Systems, Springer, 2014.
9. Aggarwal, Charu C. Outlier Analysis. New York : Springer, 2017.
10. Investigations on driver unique identification from smartphone's GPS data alone. Arijit Chowdhury, Tapas Chakravarty, Avik Ghose, Tanushree Banerjee,1 and P. Balamuralidhar2. s.l. : Journal of Advanced Transportation, Cilt 2018.
11. *Human behavior characterization for driving style recognition in vehicle system*. Fabio Martinellia, Francesco Mercaldea, AlbinaOrlandob, Vittoria Nardonec, Antonella Santoned, Arun KumarSangaiah. s.l. : Computers & Electrical Engineering, Cilt 83.
12. Ellen, Ted Dunning Friedman. *Practical Machine Learning: A New Look at Anomaly Detection*. s.l. : O'Reilly Media, 2014.
13. Saxena, Shilpi ve Gupta, Saurabh. *Practical Real-Time Data Processing and Analytics*. Birmingham - Mumbai : Packt Publishing, 2017.
14. Annielytics. Annielytics. [Çevrimiçi] [Alıntı Tarihi: 25 10 2019.] www.annielytics.com.
15. Petrov, Christo. Big Data Statistics. <https://techjury.net>. [Çevrimiçi] 22 03 2019. [Alıntı Tarihi: 11 10 2019.] <https://techjury.net/stats-about/big-data-statistics>.
16. Morphosis. Morphosis. [Çevrimiçi] 16 08 2016. [Alıntı Tarihi: 20 10 2019.] <http://imorphosis.com/big-data-in-logistics/>.
17. Gupta, Manish, et al., et al. *Outlier Detection for Temporal Data: A Survey*. s.l. : IEEE, 2014. s. 2250–2267.
18. Mehrotra, Kishan, Mohan, Chilukuri ve Huang, HuaMing. *Anomaly Detection Principles and Algorithms*. New York : Springer, 2017.
19. Babcock, Brian, et al., et al. *Models and issues in data stream systems*. New York : Madison,Wisconsin:ACMPress, 2002.
20. Bifet, Albert ve Morales, Gianmarco De Francisci. *Big Data Stream Learning with SAMOA*. China : IEEE, 2014.