

# Twitter Verilerinde Olay Bilgisi Odaklı Yer İsimleri Analizi: Ön Çalışma

## Event Information Based Location Name Analysis in the Twitter Data: A Preliminary Study

Gizem Abalı<sup>1</sup>, Ali Hürriyetoglu<sup>2</sup>, Enis Karaarslan<sup>1</sup>

<sup>1</sup>Bilgisayar Mühendisliği Bölümü, Muğla Sıtkı Koçman Üniversitesi, Muğla, Türkiye  
gizemabali93@gmail.com, enis.karaarslan@mu.edu.tr

<sup>2</sup>Radboud University, Dil Çalışmaları Merkezi, Nijmegen, Netherlands  
ali.hurriyetoglu@gmail.com

**Özetçe**—Sosyal medya, internet kullanıcılarının anlık ve güncel tepkilerini içerdiğinden, özellikle akıllı şehirler kavramı ve güvenlik konusunda önemli bir potansiyel içermektedir. Bu ön çalışmada, tweet verileri üzerinde olay bilgisi odaklı yer isim analizi yapılmış, elde edilen sonuçlar tartışılmış, sonuçların Akıllı Şehirler kavramına nasıl katkı sağlayabileceği anlatılmıştır. Çalışma için, Twitter güncel verilerinden Muğla ilinden atılan tweet verileri seçilmiş, veriler MSKÜ Bilgisayar Mühendisliği bölümünde kurulan bir Linux sunucudaki veritabanında toplanmış ve analiz edilmesi için çeşitli yöntemler denenmiştir. Yapılan bu ön çalışmada, metin yöntemleri esas alınarak tweet verilerinde analiz yapılmıştır. Analiz sonuçlarında belirli tarihlerde belirli yer isimlerinde ani artışlar olduğu gözlemlenmiş, artışların sebeplerinin, o bölgede gerçekleşen önemli olaylarla ilişkili olduğu tespit edilmiştir. Olay bilgisi odaklı yer analizi çalışmasının sonuçları, tartışmalarla beraber verilmiştir.

**Anahtar Kelimeler** — *metin madenciliği, Twitter, tweet, veri madenciliği, akıllı şehir*

**Abstract**—Social Media has an essential potential especially for the Smart City concept and security issues as it includes instant and current reactions of the Internet users. In this paper, event-detection based location name analysis is done on the tweet data, the obtained results are discussed and the way how these results can support Smart City concept is explained. Tweet data that was posted from Muğla region is collected through Twitter's current stream API, the data is stored in the database on a Linux server in the MSKU Computer Engineering Department and various methods are experimented to analyze the data. The analysis was performed by a rule-

based text mining technique. An instant increase in specific location names on specific dates is observed in the analysis results. The reason of the increase is observed as related to the important events in the region. The results of the event-detection based location analysis are given with the discussions.

**Keyword** — *text mining, Twitter, tweet, data mining; smart city*

### I. GİRİŞ

İnternet'in oluşumundan bu yana geçen sürede teknoloji gelişmiş, İnternet özellikle Twitter gibi insanların anlık bilgi paylaşabilecekleri çeşitli sosyal medya platformları ile evrimleşmiştir. Twitter, kullanıcılarına 'Neler olup bitiyor?' sorusunu sorar ve kullanıcıların 'tweet' adı verilen en fazla 140 harften oluşan kısa mesajlarla soruyu yanıtlamasına izin verir. Bunun yanı sıra, kullanıcılarına diğer kullanıcıları takip edebilme ya da diğer kullanıcılar tarafından takip edilebilme olanakları da sağlar. Twitter'ın esas amacı insanlarla iletişim kurmak değil, bilgi paylaşımı ve arayışdır[1]. Bu da Twitter'ı bilgi analizi için kullanılmak üzere uygun bir araç haline getirmektedir. Bu konuda yapılan bir çalışmada[2], 2011 yılında gerçekleşen ve Japon Devleti'nin "Büyük Doğu Japonya Depremi" olarak adlandırdığı depremin tweet verileri üzerindeki etkileri incelenmiş ve önemi gösterilmiştir.

İnternet teknolojilerinin mümkün kıldığı hayatımızdaki önemli uygulamalarından birisi de akıllı şehirlerdir. 'Akıllı Şehir' kavramı ve nasıl uygulanabileceği belli başlı çalışmalarda[3-5] ele alınmaktadır. Akıllı Şehir üzerine gerçekleştirilen bir araştırmada[6], şehir verisinin analiz

sonuçlarının vatandaşların günlük yaşayışlarına ve günlük kararlarına yardımcı olabileceği ve bunun yanında şehir gelişimine de katkıda bulunabileceğinden bahsedilmiştir.

Geçmişte yapılan araştırmalarda[7, 8], doğal afetler gibi acil durumlarda, bilgi alışverişi ve diğer insanlara iletişim sağlayan en önemli sosyal medya aracının Twitter olduğu gözlemlenmiş ve tweetlerin şehir ulaşımı, trafik, çevresel faktörler, vatandaşların güvenliği ve genel olaylar hakkında birçok bilgi içerdiği belirtilmiştir. Tweet analizi ile şehrin genel durumu hakkında daha fazla bilgi elde edilebilmesi ve akıllı şehircilik kapsamında kullanılması mümkündür. Bu verilerin kullanımını ile o an yaşanan sorunların en kısa sürede algılanması ve buna göre önlemler alınması gerçekleştirilebilir.

Akıllı Şehircilik için tweet analizinin önemi olarak aşağıdakilerden söz etmek mümkündür.

1. Gelecek nesillerin toplumdan uzak, sözsöz iletişiminde zayıf olarak yetişmesi,
2. Sosyal medya üzerinde insanların daha rahat ve kolay bir şekilde kendilerini ifade etmeleri,
3. Önlem alınması gereken durumlarda (su borusu patlağı, doğal afetler vb.) belediye ile iletişimde çıkabilecek sorunların (tefonda bekleme, yetkili kişilere ulaşamama, telefonun kitlemesi vb.) yaşanmamasıdır.

Bunlar göz önüne alındığında tweet atmanın günlük yaşamın neredeyse bir parçası haline gelmesi ile birlikte, tweet analizi ile daha hızlı önlemler alınabileceği düşünülmektedir.

Bu çalışmada, ikinci bölümde Twitter ve akıllı şehir kavramları ele alınacaktır. Üçüncü bölümde twitter verisinin akıllı şehirlerde kullanımına örnekler verilecektir. Dördüncü bölümde, twitter verilerinden yer ismi analizi konusunda bu çalışmada yapılan ön uygulama anlatılacaktır. Sonuç kısmında konu tartışılacak ve sonraki çalışmalar ele alınacaktır.

## II. TEMEL KAVRAMLAR

### A. Tweet Verisi

Tweet verisi olarak tanımlanan bilgi havuzu, Twitter kullanıcılarının anlık olarak paylaştıkları duygularından, düşüncelerinden ya da acil olarak iletmek istedikleri bilgilerden meydana gelir. Bu yüzden, tweet verisi, yaşanan zamanı analiz etmek veya analizlerden elde edilen bilgi çıkarımı ile gelecek hakkında öngörülerde bulunabilmek için kullanılmaya uygundur. Tweet verisi, sadece günlük dilde kullanılan cümlelerden meydana gelmez. Sadece cümlelerden meydana gelen tweet verisinin yanı sıra, kullanıcılara belirli bir konuda herkesin fikirlerini görebilmesi için sınıflandırmalar sağlayan "hashtag"lerden de meydana gelen tweet verileri gözlemlenebilir. Hashtag'ler, paylaşılan olayları ve bilgiyi belirlemek, bir sınıf altında toplamak için kullanılır ve başlarında "#" sembolü bulundurulur[9].

### B. Akıllı Şehir

Gelişen dünyada, birçok şey gibi kentsel nüfus da hızla artış göstermektedir. Tüm dünyada kentsel nüfus, 1950 yılından 2014 yılına kadar hızla artış göstermiş ve 746 milyondan 3.9 milyara ulaşmıştır[10]. Bu hızlı artışın devam etmesi beklenmekte ve 2050'de dünya nüfusunun %66'sının şehirlerde yaşaması öngörülmektedir[11]. Kentselleşmedeki bu hızlı artışı göz önünde bulundurduğumuzda, bu büyük insan kapasitesini en etkili şekilde yönetmek adına Akıllı Şehircilik kavramı oluşmaktadır. Bu kavram, şehrin fiziksel unsurları ile sosyal unsurlarını bir arada kullanıp, bölge vatandaşları için yüksek seviyede imkanlar sağlayan şehir oluşumunu simgeler.

Akıllı Şehircilik anlam olarak hala tamamlanmamış bir kavramdır ve bu konu üzerine gün geçtikçe yeni fikirler eklenmektedir. Akıllı şehirleri temel olarak tanımlayan aşağıdaki özelliklerden söz edilebilir[3]:

1. Günümüzdeki şehirlerle kıyaslandığında daha verimli, güçlü, yaşanabilir bir şehir,
2. Kritik altyapısı (enerji, su, internet vb.) ve servisleri için akıllı hesaplama teknolojilerini kullanan bir şehir,
3. Tüm altyapıları üzerinde kontrol sağlayabilen bir şehir (tüm yollar, tüneller, havaalanları, limanları, büyük binaları vb.),
4. Sosyal altyapıyı, fiziksel altyapıyı, İnternet altyapısını ve iş altyapısını birleştirip ortaklaşa çalışabilen bir şehirdir.

## III. TWEET VERİSİNİN AKILLI ŞEHİRCİLİKTE KULLANIMI

Sosyal medyanın akıllı şehircilikte kullanımında, vatandaşların çeşitli sorunları (kapalı yollar, kazalar, patlamış su kanalları gibi diğer birçok durum) anında haber vererek yardım etmesinden söz edilebilir. Buna ek olarak trafikte yoğun olan ya da trafiğe kapalı yollar saptanıp, vatandaşların o yollara yönelmemesi sağlanabilir[12, 13]. Bu bağlamda, uydu görüntülerini kullanan bazı uygulamalar bulunmaktadır. Ancak, uydu görüntülerini elde etmek için gereken maliyet yüksek olmakla beraber, uydu görüntülerinden detaylı çıkarımlar yapmak zaman alıcı ve masraflı bir işlemdir. Sosyal medya analizi, belediyeler tarafından karşılanamayacak kadar büyük ölçekli yatırımlara gerek duymamaktadır. Tweet verilerinin; trafik durumunu, sıklıkını, trafikte meydana gelen kazaları ve bunun gibi durumları saptamada kullanılabileceği çeşitli araştırmalarda[14-16] belirtilmiştir. Önlem alınması gereken durumlarla ilgili tweet örnekleri Şekil 1'de verilmiştir. Bahsedilen örnekler bu konudaki çalışmanın mümkün olduğunu göstermektedir. Vatandaşlar kazalardan ya da çeşitli olaylardan sonra Twitter gibi sosyal ağları kullanarak

bildirim vermektedirler. Bu eşzamanlı verilerin toplanıp analiz edilmesi durumunda gereken önlemlerin en kısa sürede alınması, olası kazaların önlenmesi, çevresel kaynakların (başlıca su) boşa harcanmasının önüne geçilmesi ve vatandaşların güvenliğinin sağlanması mümkündür. Tweet verileri gibi hali hazırda bulunan sosyal kaynakların kullanılması ile maliyet sıkıntısı çekmeden daha akıllı bir yapı sağlanması mümkündür[17].

#### IV. UYGULAMA

Twitter platformundaki veri akımına gerçek zamanlı olarak erişmek mümkündür. Fakat 2-3 haftadan daha eski tweetlere sorgulama yoluyla erişebilmek Twitter'a ancak ciddi ücretler ödenmesi ile mümkündür. Dolayısıyla, bu mesajları toplamak ve kullanıma açmanın bu zengin verilerin potansiyelini kullanmanın önemli olduğunu düşünürüz.

Bu çalışma için Muğla yerelinden atılan tweet verileri toparlanması ve işleminden geçirilerek veritabanına kaydedilmesi hedeflenmiştir. Belirli zaman aralıklarında atılan tweet metinlerinde geçen tüm yer isimlerinin bulunması ve miktarında ani artış gösteren bölgedeki sorunun en kısa sürede algılanmasını hedeflenmiştir.

Uygulama için, Muğla Sıtkı Koçman Üniversitesi'nin bulunduğu Muğla ili hedef alınmıştır. MSKÜ NetSecLab (<http://netseclab.mu.edu.tr>) bünyesinde bir Linux sunucu kurulmuştur. Programlama dili olarak Python kullanılmıştır.

Muğla Üniversitesi'nde bulunan araştırma grubu ve Radboud Üniversitesi'ndeki araştırmacıların ortak çalışmayla geliştirmiş olduğu açık kaynak kodlu yazılımla<sup>1</sup>, şu anda Muğla yöresinden atılmakta olan tweetler gerçek zamanlı olarak toplanmakta ve Akıllı Şehir ihtiyaçları kapsamında analiz edilmektedir. Akıllı şehir ihtiyacı olarak bölgenin her kesiminde kontrol sağlama ve ani değişikliklerden en hızlı şekilde haber alma faydası hedeflenmiştir. Bu doğrultuda tüm yer isimleri ve bu isimlerin günlük geçme adetleri analiz edilmiştir. Günlük geçme sayısında ani artış gösteren yer isimleri tespit edilmiş ve artışın gerçekleştiği gün hakkında araştırmalar yapılmıştır.

##### A. Veri Seti

Veri seti olarak Muğla yöresinden atılmakta olan sosyal medya verileri (tweetler) kullanılmaktadır. 27.2104, 36.5449, 29.1495, 37.3483 konumları arasından tweepy<sup>2</sup> kütüphanesi ile Twitter API ile ücretsiz olarak tweet verisi çekilmekte ve bilgi içermeyen alanları elenerek JSON formatında veritabanında tutulmaktadır. Twitter API,

<sup>1</sup> Sadece Muğla yöresinden tweet çekmek için kullanılan açık kaynak kodlu yazılım, <https://bitbucket.org/hurrial/musma/overview>

<sup>2</sup> Twitter API için Python kütüphanesi, <http://www.tweepy.org/>

Twitter'dan herkesin belirli bir miktarda ücretsiz olarak veri çekebilmesini sağlar. Sağlanan ücretsiz miktar yaklaşık olarak %1'dir[18]. Bu miktar koordinatlara bağlı veri çekiminde de değişmemektedir. Veritabanı olarak MongoDB<sup>3</sup> kullanılmaktadır. MongoDB, NoSql ve güçlü javascript desteği ile tutarlı bir sistem sunmaktadır.



Şekil 1. Önlem alınması gereken durumlarla ilgili başlıca tweetler

##### B. Yöntem

Bu çalışma ile tweet verilerinin akıllı şehircilikte kullanımına odaklanılmaktadır. Tweet verileri üzerinden gereken bilgilerin çıkarılması için önemli olan, hızlı ve verimli bir şekilde bu verilerin analiz edilmesidir. Analizlerdeki temel eğilim; tweet içinde geçen yer isimlerinin ani artış ve azalışlarını belirleyerek, bu duruma sebep olan faktörleri bulmak üzerinedir.

Belirli bir analiz esnasında odaklanılacak başlıklar;

1. tweetlerdeki yer isimlerinin ayırt edilmesi,
2. ayırt edilen bir yer isminin miktarındaki ani değişikliğin saptanması,

<sup>3</sup> MongoDB servisi, <http://www.mongodb.org>

3. ani deęişikliğe neden olan durumların çıkarımının yapılmasıdır.

Bu çalışmanın önemli bir kısmını kapsayan yer isimleri analizi, tweet içerisinde geçen yer adının tamamını ve metin içindeki konumunu bulmak üzerine geliştirilmiştir. Örneğin, “Dün gece Bodrum'da olay olmuş.” cümlesinde geçen yer ismi “Bodrum”dur ve bu kelimenin analizi ile elde ettiğimiz sonuç şu şekildedir [{{'tokens': 'Bodrum', 'location': 'bodrum', 'type': 'location', 'end': 15, 'start': 9}}]. Burada görüldüğü üzere yer ismi Bodrum'dur ve cümlenin 9. karakterinde başlayıp 15. karakterinde son bulmaktadır. Analiz için oluşturulan açık kaynak kodlu yazılımda<sup>4</sup>, Türkiye'deki yer isimlerini içeren Geonames dosyaları kullanılmaktadır. Bu yer isimlerini tweet metinleri içinde aramak için Python 'nın regular expressions ve pyparsing kütüphaneleri kullanılmıştır. Analiz konularının çeşitlendiği durumlara yönelik yeni söz dizimi kuralları da yazılmaya devam edilmektedir.

Anlamlarından biri yer ismi olan çok anlamlı kelimeler, yer ismi anlamı dışında da sıkça kullanılmaktaysa TR.txt dosyasından çıkarılmıştır. Bu kelimelere örnek olarak ‘Bedel’, ‘Yağmur’, ‘Akalm’, ‘Alan’, ‘Sandık’, ‘Siyah’ ve ‘Savcı’ verilebilir.

Olay odaklı yer isimleri analizinde etkin sonuç alınabilmesi için Türkiye’de meydana gelen önemli olaylar gözden geçirilmiş ve etkisinin en çok gözlemlenebileceği düşünülen 15 Temmuz olayı göz önüne alınarak ön çalışmada kullanılmak üzere 10 Temmuz - 19 Temmuz tarihleri arası seçilmiştir. Bu on günlük süreçte, Muğla ili genelinden toparlanmış tweet verisi (Twitter API'nin ücretsiz olarak sağladığı miktar) olan yaklaşık 166 bin tweet analiz edilmiş ve sonuçlar bu bölümde tartışılmıştır. Analiz kapsamında her gün için tweetlerde geçen yer isimleri ve miktarları bulunmuştur. Örnek günlerde en çok geçen 10 yer ismi tespit edilmiş ve bu isimlerin dağılımı Şekil 2’de verilmiştir. Grafiklerde gösterilen renkli kısımlar yer isimlerini, rakamlar ise o gün o yer isimlerinin geçtiği toplam sayıyı göstermektedir. Muğla'nın tatil beldesi olması sebebi ile genelde tatil bölgelerinde yoğunluk gözlenmiştir. Ancak 15 Temmuz gününden öncesi ve sonrası ele alındığında, ‘İstanbul’ ve ‘Ankara’ isimlerindeki artış göze çarpmaktadır. Bu da bize önemli olayların tweet verisine etkisini açıkça göstermektedir.

Bir sonraki analizde, Ankara ve İstanbul yer isimlerinin 10 – 19 Temmuz arasındaki durumu incelenmiştir. İstanbul yer isimlerinin dağılımı Şekil 3’de verilmiştir. İstanbul grafiğinde görülen ani artışlar 15 Temmuz olaylarının tweet verisine etkisini göstermiştir.

<sup>4</sup> Bir cümle içerisindeki yer isimlerini bulmak için geliştirilmiş açık kaynak kodlu proje

[https://bitbucket.org/hurrial/placenames/branch/turkish\\_location\\_recognizer](https://bitbucket.org/hurrial/placenames/branch/turkish_location_recognizer)

15 Temmuz 2016 Muğla yerelinde geçen İstanbul ile ilgili bazı kullanıcı tweetleri aşağıdaki gibidir;

1. #Darbe #Ankara #İstanbul #Turkey
2. Gunshots, F-16's, sirens. Yelling. Watching a country implode in front of my eyes. #istanbul #turkey
3. Bence darbeyi kesin #Polat Alemdar yapmıştır...\n#TSK #Turkey #Ankara #İstanbul

Ankara yer isimlerinin dağılımı Şekil 4’te verilmiştir. Ankara grafiğinde 2 büyük kırılma yaşanmıştır. Bunlardan birincisi 13 Temmuz tarihinde diğeri ise 15 Temmuz tarihinde gerçekleşmiştir. 13 Temmuz’da yaşanan kırılmanın sebebinin bu tarihte yayılan Fransa’nın Ankara Büyükelçiliği’nin kapatılma haberi olduğu gözlemlenmiştir. 15 Temmuz’da yaşanan kırılmanın sebebinin ise darbe girişimi olduğu gözlemlenmiştir.

13 Temmuz 2016 Muğla yerelinde geçen Ankara ile ilgili bazı kullanıcı tweetleri aşağıdaki gibidir:

1. DUYURUUU!!!\nFransa İstanbul ve Ankara Başkonsolosluklarını Geçici Olarak Kapatmış!!\nDemek Ki Gene Bir Patlama Canlı Bomba İhtimali Var.:(
2. Fransa'nın Ankara Büyükelçiliği ve İstanbul Başkonsolosluğu ikinci bir emre kadar kapatıldı.

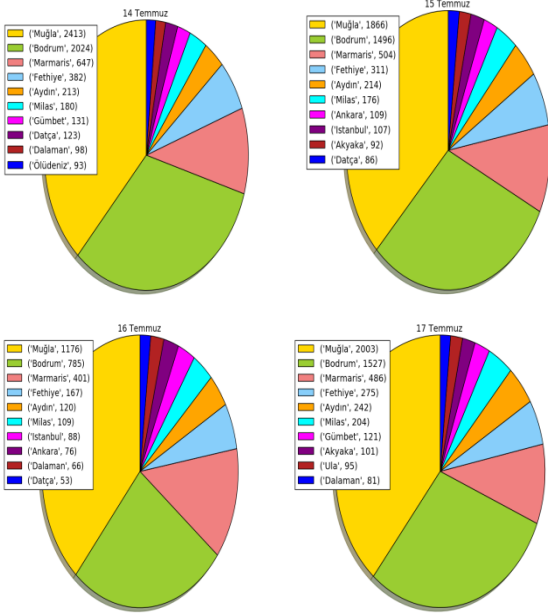
15 Temmuz 2016 Muğla yerelinde geçen Ankara ile ilgili bazı kullanıcı tweetleri aşağıdaki gibidir:

1. #Ankara'da ne oluyor? Darbe oldu mu yoksa bastırıldı mı?
2. Ankara lanet olsun lanet! Siyaset eli kanlı bir oyun ! Allahım sen tüm masumları koru
3. #Ankara olan Türkiye ve Türk halkına olcak

### C. Atılan Tweetlere Göre Ani Deęişiklik Gösteren Yerlerin Tespiti

Araştırmamız için, yer isimlerinden meydana gelen Geonames<sup>5</sup> dosyalarından Türkiye’deki yer isimlerini içeren TR.txt dosyasını kullanmaktayız. TR.txt dosyası, yer isimlerinin idlerini, isimlerini, ASCII isimlerini, deęişik versiyonlarını, ENLEM ve BOYLAM bilgilerini, ülke kodunu, admin 1, 2, 3, 4 kodlarını, zaman dilimi, eğitim bilgisi, deęişiklik tarihi bilgilerini içerir. Araştırma genelinde, Türkiye’deki yer isimlerini kapsayan TR.txt dosyasından tüm yer isimleri çekilmektedir.

<sup>5</sup> GeoNames Ücretsiz Coğrafi Veritabanı, <http://www.geonames.org/>



Şekil 2. Muğla tweet verisinde 14-17 Temmuz tarihleri için en çok geçen 10 yer ismi

Bu isimler arasından, hata payını azaltmak için Türkçe’de sıkça geçen kelimelerle (‘Tahta’ vb.) sesteş olan sözcükler belirlenmiş ve sistemin bu isimleri görmemesi sağlanmıştır. TR.txt dosyasındaki yer isimlerinin bir tweet içinde geçip geçmediğini bulmak için pyparsing ile oluşturduğumuz yeni dilbilgisi yapıları ile tweetler analiz edilmektedir. Her gün tweetlerde geçen tüm yer isimleri belirlenmekte, bu yer isimleri ve geçiş sayıları belirlenmektedir. Analiz sonucuna göre miktarında hızlı artış gösteren yer isimleri bulunmaktadır.

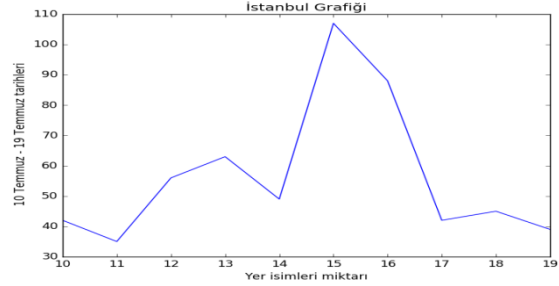
#### D. Tartışma

Çalışmada Muğla yöresinden 10 Temmuz’dan 19 Temmuz’a kadar olan her günün tweet verisi ayrı bir şekilde incelenmiştir. Söz konusu tweetlerde ağırlıklı olarak geçen yer isimleri tatil yerleri olarak tespit edilmiştir. Ancak, bazı günlerde bölge ile alakasız olan yer isimlerinde artış gözlenmiş, bu artışların sebebinin olağandışı durumlar olduğu göze çarpmıştır. Tespit edilen durumlara örnek olarak aşağıdakiler verilebilir:

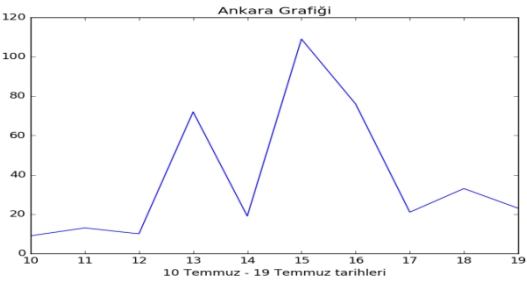
- 13 Temmuz: 13 Temmuz’da tweet metinleri içinde ‘Ankara’ yer ismi artış gözlenmiş, sebebinin Fransa’nın Ankara Başkonsolosluğu’nu kapatması olduğu görülmüştür.
- 15 Temmuz: 15 Temmuz’da tweet metinleri içinde geçen ‘Ankara’ ve ‘İstanbul’ yer isimlerinde artış gözlenmiş ve sebebinin 15 Temmuz darbe girişimi olduğu görülmüştür.

Bu çalışma olay odaklı yer isimleri analizinin önemini göstermektedir. İstatistiksel dağılımdaki

artışlardan edinilecek bilgilerin, özellikle daha büyük veri setleri incelendiğinde daha etkinleşeceği ortadadır.



Şekil 3. Muğla tweet verisinde 10-19 Temmuz tarihleri arasında ‘İstanbul’ yer isminin günlük geçme miktarı



Şekil 4. Muğla tweet verisinde 10 - 19 Temmuz tarihleri arasında ‘Ankara’ yer isminin günlük geçme miktarı

#### V. SONUÇLAR VE GELECEK ÇALIŞMALAR

Bu ön çalışmada metin madenciliği yöntemi ile yer analizi yapılmış ve olağan dışı durumların tweet verisi üzerinde etkisi olduğu gözlemlenmiştir. Ön çalışma sonucunda olağanüstü hal tespitlerinde tweet verilerinin kullanılabilirliği göz önüne alınmış, Akıllı Şehir uygulamalarında ani durumların gözlemlenmesi ve önlemlerin alınabilmesi için tweet veri analizinin önemi açığa çıkmıştır. Ön çalışmada elde edilen sonuçlar ışığında, olay odaklı yer isimleri analizinin, daha verimli ve gelişen teknoloji çağına yakışan bir hızla hizmet verebilen Akıllı Şehir uygulamaları kapsamında kullanılması mantıklı görülmüştür. Bu yöntemin, sadece şehirle kısıtlı kalmaksızın ülkede de oluşan ani değişiklikleri ve acil önlem alınması gereken durumları tespit etmede önemli olduğu kanaatindeyiz.

Gelecek çalışmalarda, elimizdeki verilerle beraber anlık veri üzerinde detaylı bir şekilde analiz yapılması amaçlanabilir. Genel anlamda elimizdeki verilerin tamamı ile beraber anlık veri üzerinde metin madenciliği, duygusal analiz ve çeşitli yöntemler kullanılarak, kurulmakta olan sistemin akıllı şehir kriterlerine tam olarak uyarlanmasını hedeflemekteyiz. Bunun için de, yukarıda sözü geçen yer ismi araştırmalarına ek olarak sistem, tweet verisi üzerinde Doğal Dil İşleme (Natural Language Processing) yöntemi olan duygu analizi

(sentiment analysis)[19] yöntemleri ile vatandaşların duygusal durumunun belirli aralıklarda analizi sağlanarak daha güçlü hale getirilebilecektir.

Analiz sonuçlarının kullanıcının anlayacağı şekilde görselleştirilmesi planlanmaktadır. Bu süreçte, elde edilen çıkarımlara göre gerekiyorsa bir web uygulaması üzerinden araştırmadan yararlanmak isteyen kişilerin rahatlıkla tweetlere ve tweet analizlerine erişebilmesi için uygun ortam hazırlanması hedeflenmektedir.

#### KAYNAKÇA

- [1] Liu, I. L., Cheung, C. M., & Lee, M. K. (2010, July). Understanding Twitter usage: What drives people to continue to tweet. Paper presented at the the 14th Pacific Asia Conference on Information Systems (PACIS), Taipei, Taiwan.
- [2] Takeshi Sakaki, Fujio Toriumi, and Yutaka Matsuo. Tweet trend analysis in an emergency situation. In Proceedings of the Special Workshop on Internet and Disasters, SWID '11, New York, NY, USA, 2011. ACM.
- [3] Canada, Taewoo Nam, Shawn Walker, J. Ramon Gil-Garcia, Sehl Mellouli, Karine Nahon, Theresa A. Pardo, Hans Jochen Scholl, Understanding Smart Cities: An Integrative Framework, Hafedh Chourabi Université Laval, 2012 45th Hawaii International Conference on System Sciences
- [4] J.M. Shapiro, "Smart Cities: Quality Of Life, Productivity, and the Growth Effects of Human Capital," The Review of Economics and Statistics 88:2 (2008).
- [5] Batty, M., Axhausen, K. W., Giannotti, F., Pozdnoukhov, A., Bazzani, A., Wachowicz, M., et al. (2012). Smart cities of the future. European Physical Journal Special Topics, 214(1), 481–518.
- [6] Kitchin R (2014) The real-time city? Big data and smart urbanism. GeoJournal 79: 1–14.
- [7] Brett D. M. PEARY, Rajib SHAW, Yukiko TAKEUCHI, Utilization of Social Media in the East Japan Earthquake and Tsunami and its Effectiveness, (July 10, 2012)
- [8] Anantharam, P. (2014). Extracting City Traffic Events from Social Streams. ACM Transactions on Intelligent Systems and Technology, 9(4), 1–27.
- [9] Zongyang Ma ve Aixin Sun ve Gao Cong, Will This #Hashtag Be Popular Tomorrow?, School of Computer Engineering Nanyang Technological University, Singapore
- [10] Department of Economic and Social Affairs, World Urbanization Prospects The 2014 Revision
- [11] Barresi, A., & Pultrone, G. (2013). European strategies for smarter cities. TeMA Journal of Land Use, Mobility and Environment, 6(1), 61e72.
- [12] T. Sakaki, Y. Matsuo, T. Yanagihara, N. P. Chandraisiri and K. Nawa, "Real-time Event Extraction for Driving Information from Social Sensors," in International Conference on Cyber Technology in Automation, Control, and Intelligent System (CYBER), Bangkok, 2012.
- [13] N. Wanichayapong, W. Pruthipunyaskul, W. Pattara-atikom and P. Chaovalit, "Social-based Traffic Information Extraction and Classification," in 11th International Conference on ITS Telecommunication (ITST), St. Petersburg, 2011.
- [14] M. Hasby and M. L. Kodra, "Optimal Path Finding based on Traffic Information Extraction from Twitter Social-based Traffic Information," in International Conference on ICT for Smart Society (ICISS), Jakarta, 2013.
- [15] S. K. Endarnoto, S. Pradipta, A. S. Nugroho and J. Purnama, "Extraction and Classification Traffic Condition Information Extraction & Visualization from Social Media Twitter for Android Mobile Application," in International Conference on Electrical Engineering and Informatics (ICEEI), Bandung, 2011.
- [16] Oostdijk, N., Hürriyetoglu, A., Puts, M., Daas, P., & van den Bosch, A. (2016). Information extraction from social media: A linguistically motivated approach. PARIS Inalco du 4 au 8 juillet 2016, 23.
- [17] IBM, "Social Media and The City", Part of a series of papers on IBM Smarter Cities 2012
- [18] Morstatter, F., Pfeffer, J., Liu, H., & Carley, K. M. (2013). Is the sample good enough? comparing data from twitter's streaming api with twitter's firehose. arXiv preprint arXiv:1306.5204.
- [19] H. Nizam and S. S. Akın, "Sosyal Medyada Makine Öğrenmesi ile Duygu Analizinde Dengeli ve Dengesiz Veri Setlerinin Performanslarının Karşılaştırılması," in XIX. Türkiye'de İnternet Konferansı, İzmir, 2014.