

20. ULUSAL PAZARLAMA KONGRESİ Anadolu Üniversitesi - Eskişehir

Sosyal Medya Analitiği: Twitter için Büyük Veri Yaklaşımı

- Özet Bildiri -

Müge Çevik¹ - Selcen Öztürkcan² - Nihat Kasap³

Anahtar Kelimeler: Sosyal Medya, Twitter, Büyük Veri, Araştırma Yöntemi

Giriş ve Amaç

Bu çalışma, büyük veri analizini sosyal medya araştırmalarında kullanımı için yöntem önerisinde bulunmaktadır. Büyük veri araştırmalarında, tündengelim ya da tümevarım yerine alternatif olarak “*tümleşim (abduction)*” yaklaşımı kullanılabilir (Kitchin, 2014). Tümleşim, elde edilen veriden yola çıkarak veriyi en iyi açıklayacak hipotezlere ulaşılmasını amaçlayan bir anlam çıkarma biçimidir (Peirce, 1958; Magnani, 2001, sf 26). Tümleşim yaklaşımında, incelenecek olguya ilişkin veri elde edildikten sonra, veriye ait bulguları en iyi açıklayacak hipotezler oluşturulur; ve veri farklı açılardan da analiz edilerek oluşturulan hipotezler sınanır. Böyle bir çalışmayı çok büyük hacimli ve detaylı bir veri seti ile yapmak araştırmanın bilimsel niteliği açısından önemli bir unsurdur. Sosyal medya araştırmalarında veri hacmi büyüklüğü ve verinin oluşum hızının yüksekliği sebebiyle, büyük veri analizi benimsenen yöntemler arasında öne çıkmaktadır (Jacob, 2009). Bu bağlamda, sosyal medya araştırmalarında analitik ve yapısal bir yöntemin takip edilmesi, araştırmaların ve çıktılarının bilimsel niteliğinin korunması açısından önem taşımaktadır. Çalışmamızın amacı, sosyal medya analitiğine örnek oluşturmak ve özellikle Twitter analizi için büyük veri analizine dayanan bir araştırma yöntemi önerisinde bulunmaktadır.

Literatür Analizi

Kullanıcıların 140 karakterlik mesaj paylaşımlarıyla iletişimlerini sağlayan, katılımı ücretsiz çevrimiçi bir küresel sosyal ağ olarak tanımlanan Twitter (Russell, 2014), farklı özellikleriyle bilimsel araştırmalara konu olmaktadır. İstatistiksel verilere göre Twitter aylık 284 milyon aktif kullanıcı sayısı ile günde 500 milyon adet tweet kaydetmektedir (Twitter, 2014)⁴. Böylesine büyük ve güncel veri havuzu, daha küçük ölçeklerde farkedilmesi mümkün olmayan, tüketici davranışlarına yönelik gerçek örüntü ve modellerin ortaya çıkarılmasına imkan vermektedir (Kleinberg, 2008).

Twitter, eğlenceli bir paylaşım aracı olmasının yanı sıra, sosyal krizlere ait bilginin yayılmasına (Oh, Agrawal & Rao, 2013; Li, Vishwanath & Rao, 2014) ve toplumsal hareketlere ilişkin bilgi alımının ve dağıtımının gerçekleşmesine (Subramani & Rajagopalan, 2003; Varnali & Gorgulu, 2014; Theocharis, Lowe, vanDeth & Garcia-Albacete, 2014) geniş katılımlı sanal bir ortam sağlamaktadır. Örneğin, Arap Baharı kapsamında 2011 yılında Mısır devrimi sırasında yaşanan olayların, Twitter’a yansımaları üzerinden incelendiği bir çalışmada (Choudhary, Hendrix, Lee, Palsetia & Liao, 2012) 800.000 tweet incelenmiştir. Bu çalışmada, protestolar sırasında ön plana çıkan ana konular, bu konulara ilişkin bireysel olarak öne çıkan duygular, bu konuların medya kuruluşları tarafından nasıl paylaşıldığı ve protestolara yön veren etkili kişiler ve kurumlar belirlenmiştir. Benzer bir yaklaşımla, Japonya’da meydana gelen Fukuşima Nükleer

1 Sabancı Üniversitesi, mugecevik@sabanciuniv.edu

2 İstanbul Bilgi Üniversitesi, selcen.ozturkcan@bilgi.edu.tr

3 Sabancı Üniversitesi, nihatk@sabanciuniv.edu

4 Twitter tarafından kaydedilen bilgiler arasında kullanıcıların sisteme kayıtlı isimleri, attıkları tweetler, takipçi sayıları, takip ettikleri kullanıcılar ve paylaşmış oldukları diğer profil bilgileri yer almaktadır (<https://about.twitter.com/company>).

Santrali kazasında atılan 38.300 tweet arasından örneklem olarak seçilen 1.520 tweet analiz edilmiştir (Li, Vishwanath & Rao, 2014). Yapılan analiz ile felaket sırasında ve sonrasında ortaya çıkan duygular ve toplum ile hükümet arasındaki iletişim incelenerek, bu tür felaketler sırasında tehlike iletişiminin yönetimi konusunda önerilerde bulunulmuştur.

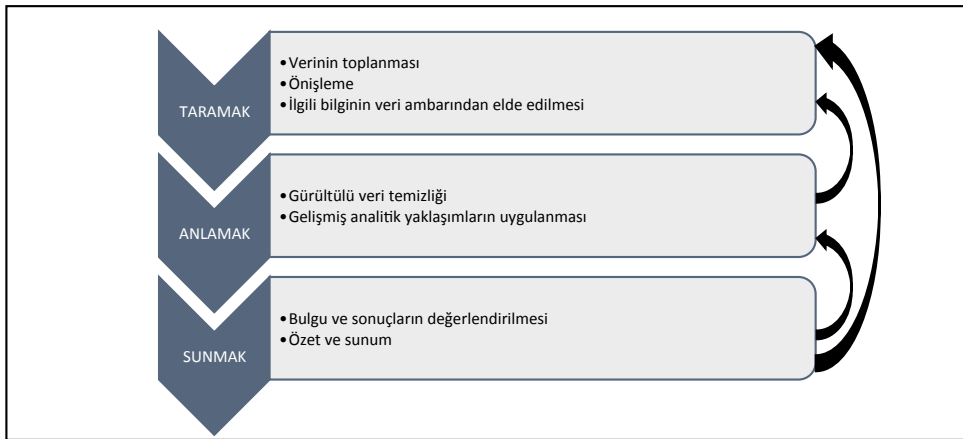
Sosyal medya iletişimi, günümüzde markalar için bütünlük iletişiminin ayrılmaz bir parçasıdır. Yazında (i) markaların çevrimdışı mecralarda yaptığı satış ve pazarlama çalışmalarının çevrimiçi ortamlara yansımaları, (ii) bu çalışmaların sosyal medya üzerindeki etkisi, ve (iii) kriz iletişimi gibi tüketicilerle ilişki yönetiminin yapı taşları olan çeşitli konularda araştırmalar bulunmaktadır (Burton & Soboleva, 2011; Kietzmann, Hermkens, McCarthy & Silvestre, 2011; Peters, Chen, Kaplan, Ognibeni & Pauwels, 2013). Twitter, takip etme modeli ve paylaşılan mesajların bulaşıcı bir etki ile dağılımını sağlayan *retweet* özelliğine (Russell, 2014; Shi, Rui & Whinston, 2014) sahiptir. Bu özelliği, Oh, Agarwal ve Rao (2013), söylenti teorisi (rumor theory) çerçevesinde tüketicilerin Twitter'daki bilgi paylaşımları üzerinden incelemiştir. Bu çalışmada, 2010 yılında Toyota'nın gaz pedalındaki sorun ile ilgili olarak araçları geri çağırıldığı döneme ilişkin atılan 5000 tweet incelenmiştir. Sonuçlarda, kurumsal bir kriz sırasında marka

kimliğine yönelik olası tehditleri belirlemek ve yönetmek ile ilgili çıkarımlarda bulunulmuştur. Başka bir çalışmada ise Twitter'da yaratılan ağızdan ağıza pazarlamanın (WOMM), filmlerin gişe hasılatı üzerine olan etkisi incelenmiştir. Bu çalışmada, takipçi sayısı ve filmler sinemada izlenmeden önce oluşan ağızdan ağıza pazarlama ile film satışları arasında pozitif korelasyon olduğu bulunmuştur (Rui, Liu & Whinston, 2013).

Twitter üzerine yapılan büyük veri araştırmaları, büyük kitlelerin davranışsal açılarından incelenmesine olanak sağlayarak, pazarlama başta olmak üzere farklı disiplinleri etkileyen sosyal olguları izlenebilmesine imkan vermektedir (Savage, 2011).

Tasarım ve Yöntem

Büyük verinin en belirgin özelliği zaman içerisinde sürekli tekrarlanan gözlemlerden oluşmasıdır (Jacob, 2009). Tekrarlı gözlem barındıran sosyal medya uygulamalarının araştırılmasında büyük veri analizi önem kazanmaktadır. (Peters, Chen, Kaplan, Ognibeni & Pauwels, 2013; Park, Baek & Cha, 2014). Çalışmamız, Fan ve Gordon (2014)'ün sosyal medya analitiği için önerdiği *Taramak*, *Anlamak* ve *Sunmak* adımlarından oluşan TAS çerçevesini kullanarak, Twitter verisine odaklanmaktadır (Şekil 1).



Şekil 1. Sosyal Medya Analitiği Süreci (Fan ve Gordon, 2014, sf. 77)

Twitter'ın kullanıcı sayısı büyüklüğü ve sık frekanslı veri akışı, veri ambarının yüksek hacimli veriden ziyade, hızlı veri depolayabilmesini gerektirir. MongoDB, bu ihtiyacı karşılayan bir veri ambarıdır (Kromer, 2014). *Taramak* adımı ile depolanan veri,

Twitter üzerinden paylaşılan kamuya açık; Tweet içerikleri, paylaşım zamanları, paylaşan kişinin numetik kullanıcı kimlikleri, paylaşımın yapıldığı cihaz bilgisi, paylaşım bütünlük bir uygulama ile yapılmışsa uygulama bilgisi ve paylaşım açığa kullanıcı

cının yer imi bilgisini içermektedir. Ambarda toplanan yapılandırılmamış verinin incelenmesi için, MongoDB'den sorgu yapılabilmesine imkân veren bir veri tabanı yönetim sistemi aracılığı ile önışleme tabii tutulması gerekir. Toplanan verilerin yapılandırılarak incelenebilir hale getirilmesi için SQL SERVER önerilen bir analiz ve raporlama sistemi olarak karşımıza çıkmaktadır (Qin and Li, 2013). Bu aşamada, eksik veri temizliği yapılmalıdır. Takip eden *anlamak* adımı, gürültülü veri için kümeleme yöntemi kullanılmalıdır. Kümeleme yönteminde, araştırılan olgu ile ilişkili anahtar terimler belirlenerek, analize ışık tutacak ana başlıklar gruplanmalıdır. Bu çerçevede veri madenciliği, içerik analizi, konu modellemesi, sosyal ağ analizi, ve eğilim analizi gibi alternatifler kullanılmalıdır. Twitter üzerinden elde edilen veriler kapsamında yapılan analizleri üç ana başlık altında toplamak mümkündür;

1. Temel frekans analizleri
2. İleri analizler
3. İçerik ve duygu analizleri

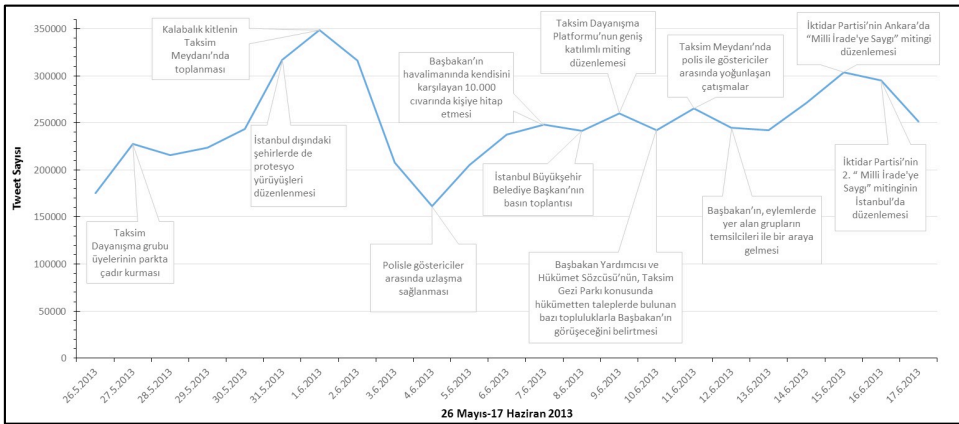
Temel frekans analizinde Twitter kullanıcı sayıları, günlük ortalama Tweet adetleri, paylaşımı tekrarlanan Tweet (Retweet) hacimleri ve kullanıcıların takipçi sayıları gibi betimsel istatistiksel dağılımlar ve zaman serisi yöntemi ile analiz edilir. Bu analizlerden yola çıkarak belirlenen örüntü kalıpları üzerinden Bayes modellemesi ve benzeri yaklaşımlarla geleceğe yönelik tahmine dayalı modelleme yapılabilmektedir. Diğer taraftan çıkarımsal istatistik yöntemleriyle değişkenler arasındaki ilişkilere ait hipotezler oluşturmak; içerik ve duygu analizleri ile ortaya atılan hipotezleri sınamak mümkündür.

Sunmak adımı, araştırılan olgu ile ilgili ana başlıkları içeren genel analizler, günlük Tweet hacimleri, Twitter kullanıcı sayıları, günlük ortalama Tweet adetleri, paylaşımı tekrarlanan Tweet (Retweet) hacimleri aktaran grafiklerle sunulur.

Bulgular ve Tartışma

Twitter, kullanıcılarının günlük hayata ilişkin yerel ve küresel bilgi paylaşımına ve edinimine olanak sağlayan bir sosyal medya ağıdır (Savage, 2011). Sanal bir platformda oluşan Twitter verisi, gerçek yaşama ilişkin olguların tespitine de olanak vermektedir. *Anlamak* adımıyla oluşturulan büyük veri eğilim grafikleri, tarihsel olarak gerçek hayatta meydana gelen olaylar ile eşleştirildiğinde yansıma görülmesi beklenmektedir. Araştırılan olguyla ilgili çevrimiçi ortamda gerçekleşmiş örüntülerin frekans analizi, açıklayıcı modelleme yapılabilmesini sağlar.

Çevrimdışı ortamlarda meydana gelen olayların çevrimiçi ortamlara yansıma biçimi ve sosyal medya araçlarının bireyler arasındaki iletişime ilişkin kullanılış şekli ile ilgili gerçek verinin Twitter üzerinden alınarak araştırma konusu olan olguya ilişkin detaylı analizlerin yapılması mümkün olmaktadır. Örneğin, 2013 yılında ülkemizde meydana gelen Gezi Parkı protestolarına yönelik yürütülen bir çalışmamızda protesto dönemi başında meydana gelen önemli olayların Twitter üzerine yansıması Şekil 2'de gösterilmektedir. Bu tarihler arasında söz konusu olaya ilişkin Twitter'da konuşulan temel kavramlar belirlerken Twitter'ın ne için ve nasıl kullanıldığını belirlemek mümkün olmaktadır. Ayrıca, bu süreçte boykot edilen markalara ilişkin sürecin çevrimiçi ortamlardaki yansıması da Twitter üzerinden yapılan analizlerle ortaya çıkarılabilmektedir.



Şekil 2. Sosyal Medya Veri Analizi Örneği: Gezi Parkı Protestosu'nun Twitter Yansıması

Benzer şekilde diğer bir çalışmamızda ise 2014 yılında yaşanan Soma Maden Kazası'nın ilk günlerinde ve sonrasında Twitter ortamındaki paylaşımlar konu edilmiştir (Şekil 3). Kar amacı gütmeyen yardım etkinlikleri şekil üzerinde kırmızı ile işaretlenmiştir. Twitter yansımaları sosyal pazarlama etkileri açısından incelenebilmektedir.



Şekil 3. Sosyal Medya Veri Analizi Örneği: SOMA Maden Kazası'nın Twitter Yansımaları

mümkündür. Çevrimdışı ortamlarda meydana gelen olay ve olgulara ilişkin yapılan daha detaylı büyük veri analizleri marka-müşteri ve hükümet-toplum gibi gruplar arasındaki iletişim yönetimi ve stratejik pazarlama kararları üzerine çıkarımlar yapılmasına olanak vermektedir.

Aynı zamanda, Twitter kapsamında yapılan büyük veri analizi ile elde edilen örüntüler, gelecek ile ilgili tahmine dayalı davranışsal modellemeler yapılabilmesini sağlamaktadır. Böylece birey ve markalar için itibarı tehdit edecek hareketlerin önceden belirlenmesiyle, başarılı itibar yönetimi ve iletişimi yapılabilmektedir. Diğer taraftan, herhangi bir markaya ilişkin kullanılan kelimelerin ve duygu kalıplarının incelenmesiyle, markanın hedef kitlesi olan tüketici gruplarının psikografik olarak betimlenmesine olanak sağlamaktadır.

Sonuç, Öneriler ve Kısıtlar

Dijital çağ ile günlük hayatımızın bir parçası haline gelen çeşitli sosyal iletişim ağlarından elde edilen büyük verinin işlenebilmesi, yönetilmesi ve analiz edilmesi ihtiyacı, sosyal bilimciler tarafından yürütülen araştırmalar için epistemolojik bir etkiye sahiptir. Sosyal medya araçları üzerinden elde edilen büyük veri, bireylerin çevrimiçi ortamlarda bilgiyi işleyiş biçimlerini ve diğer bireylerle hangi faktörlerin etkisi altında paylaşım yaptıklarını geniş bir bilgi tabanı üzerinden inceleyebilme imkanını araştırmacılara vermektedir (Oh, Agarwal & Rao, 2013; Shi, Rui & Whinston, 2014). Çalışmamız, büyük veri analizinde "tümleşim" yaklaşımını kullanarak Twitter sosyal ağında yapılan paylaşımları incelemeye yönelik araştırma yöntemi önerisinde bulunmaktadır. Twitter kullanıcı davranışları üzerine ülkemizde yapılan araş-

Twitter, kullanıcıları arasında paylaşılan mesajların onay gerektirmeden, başka kullanıcılar tarafından tekrarlı olarak iletilmesine olanak sağlamaktadır. Bu açıdan, mesajın içeriği, ilk paylaşan kişi, takipçi sayısı ve zamansal yayılım nitelikleri incelenerek toplumsal, sosyal veya pazarlamaya ilişkin herhangi bir olgunun bulaşıcılık etkisine yönelik çalışmalar yürütmek

tirmalar bulunmakla beraber, paylaşılan Tweet'lerin incelenmesine dair yöntemsel bir yaklaşım henüz benimsenmemiştir. TAS çerçevesi, Twitter uygulaması üzerinden sosyal medya analitiğine yönelik büyük veri incelemesi için önerilmektedir.

Twitter üzerinden elde edilen büyük veri, pazarlama başta olmak üzere çeşitli disiplinlerin araştırma konusu olan olgulara ait gerçek örüntülerin keşfedilerek modelleme yapılmasına imkan vermektedir. Twitter araştırmalarında büyük veri analizi yaklaşımı, gerçek zamanlı ve detaylı olarak inceleme olanağı vermesi açısından araştırmacılara büyük fırsatlar sunmaktadır. Alternatif araştırma yöntemlerine kıyasla, incelenmek istenen olgu çerçevesinde örneklem olarak çok sayıda bireye ve bu bireylere ait çok katmanlı davranış bilgisine, büyük veri üzerinden uygun maliyetle ve kısa zamanda ulaşmak mümkün olmaktadır. Ayrıca Twitter, gerçek davranışsal veri içermesi yönüyle, öz bildirim dayalı (self-response report) anket çalışmalarından daha güvenilir sebep sonuç ilişkileri ortaya konulmasına olanak verir.

Gelecek çalışmalarda da farklı olguların araştırmalarında kullanılarak, yöntem olarak güvenilirlik ve geçerlilik açısından sınanmalıdır. Twitter'da yapılan paylaşımlar kültürel ve sosyo-ekonomik farklılıklardan etkilenebileceği için önerilen araştırma yöntemi bu etkenlerin ışığında gelecek çalışmalarda dikkate alınmalıdır. TAS çerçevesi, kullanıcıları gizli olmayan profiller ile onay gerektirmeyen paylaşımlar yapabilmeye olanak veren, Foursquare, Instagram, ve Yelp gibi sosyal ağların incelenmesi için de kullanılabilir. Diğer taraftan, toplanan verinin, incelenen sosyal ağın sadece paylaşımları gizli olmayan kullanıcıların davranışlarını içermesi, analiz sonuçları açısından bir kısıt ortaya koymaktadır (Crawford, 2013).