

Türkiye 2023 Cumhurbaşkanlığı seçimlerinde adayları kim takip ediyor?

Who follows Turkish Presidential Candidates in 2023 Elections?

Onur Varol
Center of Excellence in Data Analytics
Sabanci University
onur.varol@sabanciuniv.edu

I. GİRİŞ

Özetçe —Sosyal medya kullanımı özellikle seçim dönemlerinde siyasi tartışmaların yapılması için önemli bir platform oluşturmaktadır. Siyasiler seçmenlere ulaşmak ve kampanyalarını paylaşmak için platformu kullanırken farklı seçmen grupları için de bir tartışma ortamı yaratmaktadır. Destekleyici ve eleştirel söylemleri bu platformlarda eş zamanlı görmek mümkün olmaktadır. Özellikle seçim tarihi yaklaştıkça kararsız seçmen farklı söylemleri de duyabilmek için siyasileri çevrimiçi platformlarda takibe almaktadır. Siyasilerin takipçilerindeki değişimler özellikle sonuçlarda belirleyici olabilecek kadın ve genç seçmenlerin analiz edilmesiyle detaylı olarak gözlemlenebilmektedir. Bu çalışmada 2023 Cumhurbaşkanlığı seçimlerinde yarışan adayların Twitter platformu üzerinde sahip oldukları takipçiler ve bu takipçilerin demografik özelliklerinin zaman içerisinde değişimi incelenmektedir. Cinsiyet ve yaş grubu gibi demografik özelliklerin analiz edilmesinden önce bot hesapların veri kümesinde tespit edilip ayrıştırılması da sağlanmıştır. Adayların takipçi sayılarının değişiminde önemli olayların rolü ve gözlenen anomalilerin demografideki değişim ile ilişkisi incelenmiştir.

Anahtar Kelimeler—hesaplamalı sosyal bilimler, sosyal medya analizi, seçim analizi, bot tespiti, demografi analizi

Abstract—Social media serves as an important platform for political discussions, especially during election periods. While politicians use the platform to reach out to voters and share their campaigns, it also creates a debate environment for different voter groups. It is possible to see supportive and critical speeches simultaneously. Particularly as the election date approaches, undecided voters follow politicians on online platforms to learn about their campaigns. Changes in politicians' followers can be observed in detail by analyzing women and young voters who may be decisive in the results. In this study, we analyzed candidates for the 2023 Presidential elections of Türkiye and their followers on the Twitter platform. We examined how the demographic of these followers change over time. Prior to analyzing demographics such as gender and age group, we identified bot accounts and isolated them from the dataset. Significant events are reflected in the number of followers and the relationship between the observed anomalies and changes in the demographics have also been studied in this work.

Keywords—computational social science, social media analysis, election analysis, bot detection, demographics analysis

979-8-3503-4355-7/23/\$31.00 ©2023 IEEE

Türkiye Cumhuriyeti'nin tarihindeki en önemli seçimlerden biri olarak görülen 2023 Türkiye genel seçimleri 14 Mayıs 2023 tarihinde gerçekleştirilecektir. Cumhurbaşkanlığı seçiminde yarışan 4 aday bulunmaktaydı ve bu adayları desteklemek için oluşturulan ittifaklar ve diğer destek olan partiler aynı zamanda Türkiye Büyük Millet Meclisinin 28. döneminin 600 yeni üyesinin belirleneceği seçimler için de işbirliğine girmiştir. Seçimlere 3 gün kala adaylardan biri seçimden çekildiğini açıklamıştır ancak bu çalışma kapsamında bu adayın da bilgileri analiz edilmeye devam edilmiştir.

Sosyal medya kullanımının artmasıyla siyasi tartışmalar için önemli bir platform oluşmakta ve gazete, radyo ve televizyon gibi klasik medya kanallarıyla rekabet edecek etkinliğe ulaşmıştır [1]–[3]. Özellikle yakın zamanda sosyal medyanın seçim döneminde kullanıcı davranışlarına etkileri ve platformun yanlış bilgi yayılımında kullanılarak bir siyasi manipülasyon aracı olarak kullanılması gittikçe daha çok araştırılan konulardan olmaktadır [4]–[7].

Sosyal medyanın seçim döneminde manipülasyonunda sahte takipçiler, bot ve troll hesaplar aktif olarak bilgi akışının kontrolünde kullanılmaktadır [8]–[12]. Platformların kullandığı algoritmalarındaki zaafiyetler ile de trendlerin ve arama sonuçlarının manipülasyonu mümkün olmaktadır [13]–[15]. Bu kordinasyonlu faaliyetlerin gerçekleştirilmesinde kullanılan hesaplar siyasiler ile onlara sahte takipçi kazandırarak veya paylaşımlarını popüler göstermek için paylaşarak etkileşmektedir.

Seçim tarihleri yaklaştıkça özellikle kararsız seçmeni etkilemek, seçim politikaları ve vaatlerle farklı partilerin seçmenlerini kendi tarafına çekmek için siyasiler aktif olarak miting düzenlemekte ve çevrimiçi mecraları kullanmaktadırlar. Bu çalışmada Türkiye 2023 Cumhurbaşkanlığı seçimi öncesi adayların kazandıkları takipçilerin demografik analizleri gerçekleştirilmekte ve son 2 aylık süreçte siyasilerin erişmeyi başardığı takipçi grupları analiz edilmektedir. Özellikle seçimin sonucunu etkileme potansiyeli yüksek olan kadın seçmenler ve ilk kez oy kullanacak genç seçmene erişmek siyasiler açısından önemli bir strateji olduğundan bu çalışmada ortaya çıkan sonuçlar da hangi siyasilerin bu konuda daha etkili olduğunu yansıtmaya potansiyeli taşımaktadır. Bu çalışmada siyasilere yö-

nelik oluşturduğumuz kapsamlı veriseti üzerinde gerçekleştirilen analiz ile siyasilerin takipçi demografisine daha detaylı bakma imkanımız oldu.

II. VERİ VE YÖNTEM

A. Twitter Seçim Veriseti

Bu çalışmada Twitter'in araştırmacılar için sağladığı uygulama programlama arayüzleri (application programming interfaces, API) üzerinden sosyal ağlardan veriye erişim sağlanmaktadır. Bu kaynak üzerinden gönderilen mesajlar ve kullanıcı profilleri sistematik olarak toplanabilmektedir.

Bu çalışmada yapılan demografik analizlerde cumhurbaşkanlığı adaylarının takipçileri Twitter API ile toplanmış ve analizlerde kullanılmıştır. Adayların sahip oldukları takipçilerin verisi ise açık kaynaklı olarak paylaşılmakta olan #Secim2023 projesi bünyesindeki veri kaynağı kullanılarak Harvard Dataverse platformu üzerinden erişilmiştir [16]. Bu verisetinde paylaşılmakta olan 48 milyon sosyal medya hesabı belirlenmiş ve profil bilgileri Twitter API aracılığı ile toplanmıştır.

Çalışmada analiz edilen 3 aday i) Recep Tayyip Erdoğan (@RTERdogan), ii) Kemal Kılıçdaroğlu (@kilicdaroglu), iii) Sinan Oğan (@DrSinanOgan) ve seçimlere 3 gün kala adaylıktan çekilen Muharrem İnce (@vekilince) hakkında veri sistematik olarak derlenmiş ve analiz edilmiştir. Bu hesaplara ait profil bilgileri toplanmıştır ve takipçilerine ait bilgiler demografik ve bot hesap tespit etmek için de kullanılmıştır.

B. Takip zamanı çıkarımı

Twitter platformunda takipçi bilgileri API tarafından takip etme zamanına bağlı sıralı olarak sağlanmaktadır. Bu takip etme sırası ve hesapların yaratılma tarihine dair bilginin bir arada kullanılması ile takip zamanına dair tutatlı çıkarımlar yapmak mümkün olmaktadır. Bu yaklaşım özellikle popüler hesapların takipçilerini oldukça tutarlı bir şekilde tahmin edilebildiğini göstermektedir [17]. Buradaki yaklaşım ile takip sırası ve o zamana kadar görülmüş olan en güncel hesap yaratılış tarihi takip zamanı için bir üst limit oluşturacak şekilde izlenmekte ve önceki takipçiler için hesaplanan üst limitleri de güncel takipçinin takip zamanı alt limiti olarak kabul edebilmektedir. Özellikle bahsi geçen siyasi hesaplar gibi milyonlarca takipçili ve günde yüzlerce yeni takipçi kazanan hesapların takipçileri saatlik hatalarla belirlenebilmektedir.

C. Bot tespiti

Sosyal medyanın kullanımının artması, platformlar üzerinde otomasyon teknolojileri kullanılarak hesapların görünürlüğü ve popüleritesi manipüle edilebilmektedir [18]. Geçtiğimiz on yıl içerisinde botların kullanılmasıyla dünyanın farklı coğrafyalarında seçim manipülasyonları üzerine bu teknolojilerin kullanımını gösteren çalışmalar yapılmıştır [19], [20].

Twitter üzerinde otomasyon kullanan hesaplar ciddi sorunlara yol açmaktadır. Bu hesapların faaliyetlerini tespit etmek amacıyla geliştirilen makine öğrenmesi yaklaşımları içerisinde en yaygın kullanılan araç Botometer isimli servistir. Sosyal bot tespitinde kullanılan bu sisteminin kullanıcı profil bilgileri ile çıkarımlar yapabilen bir versiyonu da bulunmaktadır ve API üzerinden erişimi bulunmaktadır [21]. Bu aracın kullanılması

ile siyasilerin sahip olduğu takipçilere ait bot skorlarının hesaplanabilmesi mümkün olmuştur. Hesapların profil bilgilerini kullanarak eğitilmiş bir makine öğrenmesi modeli bir hesabın otomasyon kullanma riskini $[0, 1]$ aralığında bir skor olarak sağlamaktadır ve yüksek skorlar modelin otomasyon kullanma riskinin yüksek olduğunu söylemektedir.

D. Demografi tespiti

Sosyal medya kullanıcılarına ait bilgilerden faydalanılarak yaş ve cinsiyet gibi demografik bilgileri tahmin etmek mümkün olabilmektedir. Bu amaçla geliştirilen *m3inference* sistemi Twitter kullanıcılarının isim, profil resmi ve kullanıcıların kendilerini tanıttıkları alandaki metin kullanılarak yaş grubu ve cinsiyet çıkarımını Türkçenin de dahil olduğu 32 dil için yapılabilmektedir [22]. Yaş grubu için 4 farklı kategori tanımlanmıştır: 18 yaş ve altı, [19, 29), [30, 39), ve 40 yaş üstü. Model ayrıca organizasyon hesaplarını da tahmin edebilmektedir ve analizlerde organizasyon olmayan hesaplara dair demografik bilgilerinden faydalanılmıştır.

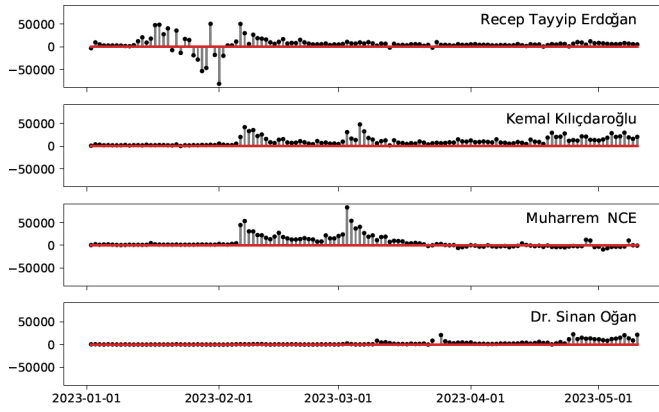
Özellikle cinsiyet ve yaş gibi özelliklerde derin öğrenme yaklaşımları ile tahminlenmesinde yanlışlıklar olabileceği düşünüldüğünden kendi veri setimiz üzerinden yapılan bir örneklem ile model performansı ölçülmüştür. İlk olarak 5 farklı araştırmacı tarafından hesapların yaş ve cinsiyet için etiketleme çalışması yapılmıştır. Bu çalışma ile her bir etiketleyici 100 hesaba dair bilgileri toplamıştır. Bu 100 hesap içerisinde 25 hesap etiketleyiciler arasındaki tutarlılığın ölçülmesi için ortak kullanımdadır. Etiketleyiciler arasındaki tutarlılık Cohen's Kappa skoru ile ölçülmüştür ve cinsiyet için 0.88 ve yaş grupları için 0.48 olarak ölçülmüştür. Buradan cinsiyet için yüksek tutarlılık ve yaş grupları için makul bir tutarlılık ölçülmüştür. Bu analiz ile cinsiyet tespitinin yaş gruplarının belirlenmesinden daha kolay olduğu ancak iki kategoride de tatmin edici sonuç alınabileceği görülmektedir. Etiketlenmiş verilerin model tahminleri ile kıyaslanması ile de model performansına dair çıkarımlarda bulunmamız mümkün olmaktadır. Burada etiketleyicilerin çoğunlukla oluşturduğu çıkarım üzerinden yapılan kıyasmalarda cinsiyet için %88 tutarlılık gözlenmiş ve yaş grupları için olan 4 kategoride de %52 doğruluk gözlenmiştir ve karışıklık matrisi incelendiğinde komşu yaş grupları arasında gerçekleşen hatalar olabildiği görülmüştür ve yaş gruplarına ait sonuçların bu bulgular ışığında yorumlanması önerilmektedir.

III. BULGULAR

A. Takipçi sayılarındaki değişimler

Seçime yaklaşan zaman içerisinde siyasilerin artan aktiviteleri, seçmenlerin söylemlere karşı artan ilgisi ve sahte takipçiler ile popülerite manipülasyonu yapılması sebebiyle siyasiler takipçi kazanabilmektedir. Bunun yanında platform üzerinde tespit edilen bot hesaplar kapatılmakta, satın alınan sahte takipçiler ise takipten çıkabileceğinden takipçi sayısında kullanıcıların kendilerin gerçekleştirdiği takipten çıkarma sonucu oluşan farklara ek olarak takipçi kaybedilebilmektedir.

Cumhurbaşkanlığı adayı hesapların takipçi sayılarındaki günlük değişimleri 2023 yılının başından bu yana incelediğimizde Figür 1 ile gösterildiği şekilde değişimler gözlenmiştir. Özellikle 6 Şubat'ta yaşanan deprem felaketini takip eden dönemde pek çok kullanıcı bilgiye ulaşmak için sosyal medya



Şekil 1: **Günlük takipçi değişimleri.** Siyasilerin profillerindeki takipçi sayılarının günlük olarak toplanması ve değişimleri ile kazanılan ve kaybedilen takipçilerin miktarları incelenmektedir.

platformlarına yönelmiş ve bununla eş zamanlı olarak da siyasileri takip etmeye başlamıştır.

Günlük değişimlerde gözlenen değişimlerden siyasi olaylarla açıklanması mümkün olan bir diğer değişim Mart ayı başında Meral Akşener'in Millet İttifakı ile yaşadığı anlaşmazlığın Muharrem İnce için yarattığı olumlu değişim de görmek mümkün olmaktadır. Recep Tayyip Erdoğan için Ocak ayı sonunda gözlenen takipçi kayıpları ve Kemal Kılıçdaroğlu'nun Nisan sonundan bu yana yaşanan artan olumlu büyümeyi de gözlemek mümkün olmaktadır.

B. Takipçilerin demografi özellikleri

Cumhurbaşkanlığı için aday olan siyasetçilerin hesapları uzun bir zamandır siyasette aktif olduklarında belli bir takipçi kitlesine sahip olmaktadır. En fazla takipçi sayısı ile 12. Cumhurbaşkanı @rterdogan hesabının 20.3 milyon takipçisi bulunmaktadır. Bu sayıyı 9.9 milyon ve 6.9 milyon ile @kilicdaroglu ve @vekilince takip etmektedir. ATA ittifakı adayına ait @DrSinanOgan hesabının ise yaklaşık 1.5 milyon takipçisi bulunmaktadır. Oransal olarak belirtilen analizlerin gerçek kullanıcı sayılarına dönüştürülmesinde bu takipçi sayıları göz önüne alınmalıdır.

Bu hesapların takipçileri içerisinde aktif hesaplar tweet sayısı, arkadaş ve takipçi sayılarının herbirinin 25'ten büyük olanları belirlenmiştir ve demografik özellikleri profil resimleri ve metaveriler kullanılarak incelenmiştir. Bu aktif hesaplar Figür 2'de öncelikle bot olma olasılıklarına göre incelenmiştir. Adaylar içerisinde çok takipçiye sahip olan @rterdogan ve @kilicdaroglu hesapları gözlemlenmiştir. Demografik özelliklere bakıldığında oransal olarak en fazla kadın takipçiye sırasıyla Muharrem İnce ve Kemal Kılıçdaroğlu'nun sahip olduğu görülmüştür. Bu oran %14 ile en az Sinan Oğan için gözlenmiştir. Yaş dağılımında da Sinan Oğan %43 ile kırk yaşının üzerinde takipçilere sahip olan hesap olarak diğer adaylardan ayrılmaktadır. En yüksek oranda genç takipçiyi ise Muharrem İnce'nin hesabı taşımaktadır. Seçmen olabilecek 19-29 yaş grubunda da oransal olarak diğer adayların önünde gözükmemektedir ve onu Recep Tayyip Erdoğan izlemektedir.

C. Seçim öncesi demografi değişimleri

Seçim yarışının hız kazandığı ve adayların online platformlarda ve miting alanlarında daha aktif kampanya çalışması yürüttüğü dönemde takipçilerinin demografilerindeki değişimleri gözlemek önemli olmaktadır. Bu değişimlerin doğrudan oy davranışına yansıtacağını beklemek oldukça yanıltıcı olacaktır ancak adayların paylaşımlarına ve söylemlerine olan artan ilgi önemli olabilmektedir. Özellikle adayların takipçi demografisindeki değişimleri incelemek önemli olduğunda bu kısımda cinsiyet ve yaş kırımlarındaki değişimi son bir kaç aylık süreç için analiz etmekteyiz ve sonuçları Figür 3'de gösterilmiştir.

Takipçi sayılarındaki 2022 Eylül ayından bu yana yaşanan artışın farklı demografik gruplara göre analizinde Recep Tayyip Erdoğan'ın takipçilerinde daha yavaş bir artış gözlenirken, Muharrem İnce ve Kemal Kılıçdaroğlu için ciddi yükseliş gözlenmiştir. Kemal Kılıçdaroğlu oransal olarak en fazla yeni kadın seçmen kazanan aday olurken genç kullanıcılara erişme konusunda Muharrem İnce öne çıkmıştır.

IV. SONUÇLAR VE TARTIŞMA

Bu deskriptif çalışmada sosyal medya takipçilerin üzerinden yapılan sade bir analiz paylaşılmıştır. Türkiye 2023 Cumhurbaşkanlığı seçiminde yarışan 4 aday için takipçi sayılarında ve demografilerindeki değişimleri inceledik.

Sosyal medya davranışı ve seçim sonuçları arasında ilişki kurmak iddialı olmaktadır ancak literatürde farklı seçimler için denenmiştir ve başka çalışmalarda da zorluklarına değinilmiştir [23], [24]. Özellikle sosyal botlar kullanılarak yürütülen kampanyalar, koordinasyonlu trol faaliyetleri ve değişen gündem sebebiyle zaman içindeki trendleri analiz etmek seçim güvenliği açısından daha önemli olmaktadır [12], [18].

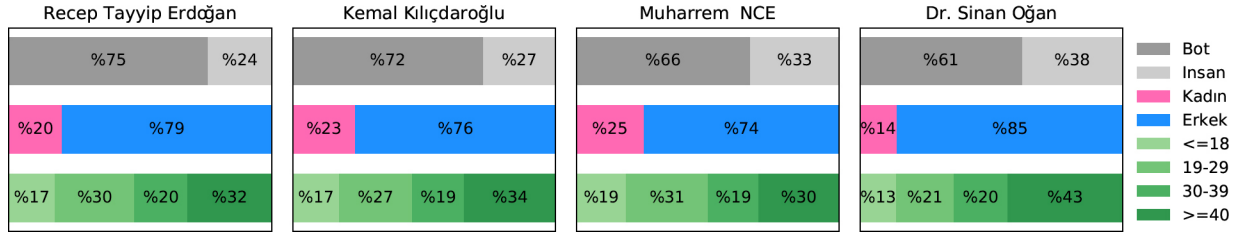
Yaptığımız analizde siyasilerin kazandığı takipçilerin otomasyon kullanma olasılığı yüksek olanları belirlenmiştir ve kalan hesaplar için demografik bir analiz yapılmıştır. Özellikle seçime gidilen son 2 aylık sürede kadın ve genç sosyal medya kullanıcılarına hitap edebilen adayları gözlemlenmiştir. Adayların söylemlerini ağırlıklı olarak takip eden kullanıcı tiplerini gözlemleme fırsatımız olmuştur. Bu analizin seçim sonuçlarının yorumlanmasında demografinin rolü sebebiyle fayda sağlaması amaçlanmaktadır.

TEŞEKKÜR

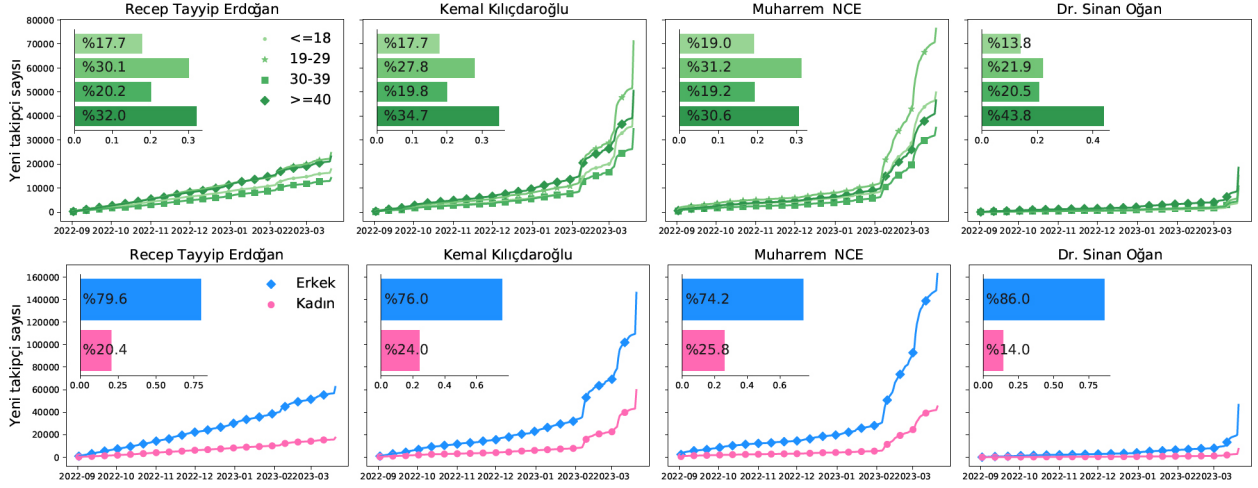
Bu çalışmada #Seçim2023 projesine verdikleri katkılardan dolayı VRL Lab üyelerine ve değerli geri bildirimleri için SİU'nun uzman hakemlerine teşekkür ederiz. Bu çalışma TÜ-BİTAK 121C220 numaralı proje tarafından desteklenmektedir.

KAYNAKLAR

- [1] M. Morris and C. Ogan, "The internet as mass medium," *Journal of Computer-mediated communication*, vol. 1, no. 4, p. JCMC141, 1996.
- [2] P. T. Metaxas and E. Mustafaraj, "Social media and the elections," *Science*, vol. 338, no. 6106, pp. 472–473, 2012.
- [3] N. Newman, "Mainstream media and the distribution of news in the age of social media," 2011.
- [4] S. Morgan, "Fake news, disinformation, manipulation and online tactics to undermine democracy," *Journal of Cyber Policy*, vol. 3, 2018.
- [5] S. Aral and D. Eckles, "Protecting elections from social media manipulation," *Science*, vol. 365, no. 6456, pp. 858–861, 2019.



Şekil 2: **Takipçilerin demografik kırılımları.** Cumhurbaşkanlığı adaylarının sahip olduğu takipçilerin cinsiyet ve yaş analizleri ile bot analizine göre yüzdesel dağılımları.



Şekil 3: **Güncel takipçilerin yaş (üst) ve cinsiyet (alt) kırılımları.** Seçime yaklaşıldıkça kazanılan yeni takipçilerin kategorilere dağılımı ve zaman içinde artış miktarları gösterilmiştir.

- [6] S. Bradshaw and P. N. Howard, "Challenging truth and trust: A global inventory of organized social media manipulation," *The computational propaganda project*, vol. 1, pp. 1–26, 2018.
- [7] S. C. Woolley and P. N. Howard, *Computational propaganda: Political parties, politicians, and political manipulation on social media*. Oxford University Press, 2018.
- [8] E. Saka, *Social media and politics in Turkey: A journey through citizen journalism, political trolling, and fake news*. Lexington Books, 2019.
- [9] E. Manuoğlu and B. Öner-Özkan, "Sarcastic and deviant trolling in turkey: Associations with dark triad and aggression," *Social Media+ Society*, vol. 8, 2022.
- [10] S. Zannettou, T. Caulfield, E. De Cristofaro, M. Sirivianos, G. Stringhini, and J. Blackburn, "Disinformation warfare: Understanding state-sponsored trolls on twitter and their influence on the web," in *Companion Proc. of the 2019 WEB Conf.*, 2019, pp. 218–226.
- [11] S. Cresci, R. Di Pietro, M. Petrocchi, A. Spognardi, and M. Tesconi, "Fame for sale: Efficient detection of fake twitter followers," *Decision Support Systems*, vol. 80, pp. 56–71, 2015.
- [12] C. Shao, G. L. Ciampaglia, O. Varol, K.-C. Yang, A. Flammini, and F. Menczer, "The spread of low-credibility content by social bots," *Nature communications*, vol. 9, no. 1, pp. 1–9, 2018.
- [13] Y. Zhang, X. Ruan, H. Wang, H. Wang, and S. He, "Twitter trends manipulation: a first look inside the security of twitter trending," *IEEE Trans. on Information Forensics and Security*, vol. 12, 2016.
- [14] T. Elmas, R. Overdorf, A. F. Özkalay, and K. Aberer, "Ephemeral astroturfing attacks: The case of fake twitter trends," in *2021 IEEE European Symposium on Security and Privacy*, 2021.
- [15] K. Lee, P. Tamilarasan, and J. Caverlee, "Crowdturfers, campaigns, and social media: Tracking and revealing crowdsourced manipulation of social media," in *Proc. of the Intl. AAAI Conf. on Web and Social Media*, vol. 7, no. 1, 2013, pp. 331–340.
- [16] A. Najafi, N. Mugurtay, E. Demirci, S. Demirkiran, H. A. Karadeniz, and O. Varol, "# secim2023: First public dataset for studying turkish general election," *arXiv preprint arXiv:2211.13121*, 2022.
- [17] B. Meeder, B. Karrer, A. Sayedi, R. Ravi, C. Borgs, and J. Chayes, "We know who you followed last summer: inferring social link creation times in twitter," in *Proceedings of the 20th international conference on World wide web*, 2011, pp. 517–526.
- [18] E. Ferrara, O. Varol, C. Davis, F. Menczer, and A. Flammini, "The rise of social bots," *Communications of the ACM*, vol. 59, no. 7, 2016.
- [19] T. R. Keller and U. Klinger, "Social bots in election campaigns: Theoretical, empirical, and methodological implications," *Political Communication*, vol. 36, no. 1, pp. 171–189, 2019.
- [20] S. Cresci, "A decade of social bot detection," *Communications of the ACM*, vol. 63, no. 10, pp. 72–83, 2020.
- [21] K.-C. Yang, O. Varol, P.-M. Hui, and F. Menczer, "Scalable and generalizable social bot detection through data selection," in *Proc. of the AAAI Conf. on Artificial Intelligence*, vol. 34, 2020.
- [22] Z. Wang, S. Hale, D. I. Adelani, P. Grabowicz, T. Hartman, F. Flöck, and D. Jurgens, "Demographic inference and representative population estimates from multilingual social media data," in *The World Wide Web Conference*. ACM, 2019, pp. 2056–2067.
- [23] P. T. Metaxas, E. Mustafaraj, and D. Gayo-Avello, "How (not) to predict elections," in *2011 IEEE third international conference on privacy, security, risk and trust and 2011 IEEE third international conference on social computing*. IEEE, 2011, pp. 165–171.
- [24] D. Gayo-Avello, "i wanted to predict elections with twitter and all i got was this lousy paper—a balanced survey on election prediction using twitter data," *arXiv preprint arXiv:1204.6441*, 2012.