SOSYAL BOT ALGILAMA TEKNİKLERİ VE ARAŞTIRMA YÖNLERİ ÜZERİNE BİR İNCELEME

Arzum KARATAŞ, Serap ŞAHİN
İzmir Yüksek Teknoloji Enstitüsü Bilgisayar Mühendisliği Bölümü, İzmir, Türkiye
arzumkaratas@iyte.edu.tr, serapsahin@iyte.edu.tr

ÖZET

Facebook, Twitter, LinkedIn gibi çevrimiçi sosyal ağların (OSN) popülerliği ve web servislerinin yaygınlığı, bu alanlarda sosyal bot olarak nitelendirdiğimiz yazılımsal sosyal aktörlerin ortaya çıkmasına ve yaygınlaşmasına neden olmuştur. Ancak çoğunlukla bu aktörler kötü rollerde karşımıza çıkmaktadırlar. Örneğin, sosyal botlar insanmış gibi sohbetlere katılma, başka hesapları çalarak üzerinden dolandırıcılık yapma, yanlış bilgi yayma, borsayı manipüle etme, sahte halk tabakası oluşturarak propaganda yapma gibi ciddi problemlerde karşımıza çıkmaktadırlar. Bununla beraber, istenmeyen postaları ve zararlı yazılımları yaymanın en etkin araçları haline gelmişlerdir. Dahası, botlar gerçek hesapları ele geçirerek "zombi bilgisayar ağı" (botnet attack) saldırıları düzenlemekte de kullanılmaktadırlar. Öte yandan; sosyal botların, sosyal paylaşım ağları üzerindeki yaygınlığı ve önemi inkâr edilemez bir gerçekliktir. Bu çalışmada, kötü niyetli sosyal botların potansiyel tehlikeleri vurgulanmış, literatürdeki bot tespit yaklaşımları metodolojik bir sınıflandırma içerisinde gözden geçirilmiş, bu yaklaşımların sınırları ve açık problemler sunulmuş ve bu problemleri çözmeye yönelik iki yeni yaklaşım önerilmiştir.

Anahtar Kelimeler: Sosyal botlar, sosyal ağlar, çevrim içi ağlar, sosyal bot algılama, Sybil hesaplar.

Review on Social Bot Detection Techniques and Research Directions

ABSTRACT

The rise of web services and popularity of online social networks (OSN) like Facebook, Twitter, LinkedIn etc. have led to the rise of unwelcome social bots as automated social actors. Those actors can play many malicious roles including infiltrators of human conversations, scammers, impersonators, misinformation disseminators, stock market manipulators, astroturfers, and any content polluter (spammers, malware spreaders) and so on. It is undeniable that social bots have major importance on social networks. Therefore, this paper reveals the potential hazards of malicious social bots, reviews the detection techniques within a methodological categorization and proposes avenues for future research.

Keywords: Social bots, OSN, Sybils, social bot detection, Sybils.

1. GİRİŞ (INTRODUCTION)

Facebook, Twitter ve LinkedIn gibi çevrimiçi sosyal ağlar (OSN) günlük hayatımızda belki de en yoğun olarak kullanılan internet alanları haline geldiler. Sosyal ağlar; kullanıcılarına yeni arkadaşlıklar kurmaları, ortak ilgi alanları ile meşgul olmalarını teşvik ederek, kullanıcılar arasında bilgi paylaşımını arttıran bir platform oluşturmaktadır. Bu platformlar yeni bir tür iletişim kanalı olarak hayatımızda önemli bir rol oynamaktadır. Bu özellikleri ile müşteri profil ve güncel hareketleriyle oluşan büyük veriye sahiptirler. Sosyal ağlar, bu özellikleriyle beraber kolay ve zengin programlama ara yüzünden dolayı, aynı zamanda sosyal botların istilası için açık ve cazip

bir hedef haline gelmektedir [1].

En basit tanımıyla, sosyal botlar kullanıcı faaliyetlerini otomatik olarak yerine getiren yazılımlardır. Bu faaliyetler, (i) sanki bir insan tarafından yazılmış gibi görünen, anlamlı gönderiler oluşturmak, (ii) başka gönderileri, fotoğrafları, durum bildirimlerini veya güncellemeleri yeniden göndermek (reposting), (iii) yapılan gönderilere yorum yapmak veya beğenide bulunmak ve (iv) ağdaki diğer kişiler ya da sayfalarla bir kullanıcı gibi bağlantı kurmak şeklinde olabilir. Anlaşılacağı üzere, botların yerine getirebildiği görevlerin karmaşıklık seviyeleri farklıdır. Örneğin [2, 3] botlar, haberler, hava durumu, ya da internet günlüğü (blog) gönderilerinden bilgiyi alıp, daha sonra bunları sosyal paylaşım ağında yeniden yayınlayacak basitlikte olabilir ya da kişiler arası sohbetleri algılayıp, sızmaya çalışacak kadar sofistike de olabilirler. Bu yeteneklerin elbette sosyal ağ kullanıcıları için farklı artı ve eksileri vardır. Dolayısıyla iyi ve kötü niyetli kullanımları olabilir:

(i) Öncelikle, botlar iyi niyetlerle tasarlanabilirler. Örneğin, kişi kendi gizliliğini korumak, ya da zamandan kazanmak amacıyla bazı sıradan faaliyetleri otomatik hale getirmek için kullanmak isteyebilir. Bu faaliyetlere örnek verecek olursak; kişi sürekli aldığı güncel haberleri, hava durumunu, sürekli takip ettiği bir sayfa veya hesaplardan gelen güncellemeleri kendisi de sayfasında anlık olarak paylaşmak isteyebilir. Ya da istediği bir konu üzerinde bilgi toplayabilir [4], yeni takipçilerine otomatik teşekkür mesajı gönderebilir. Bunların dışında botlar; Siri [1] gibi, kullanıcıları için sanal asistan ya da Microsoft'un Tay [5] yapay zekâ botu gibi, bireysel sohbet botu veya şirketler için kullanıcı dostu müşteri hizmetleri [6] sunmak için tasarlanabilirler.

(ii) Madalyonun diğer yüzünde ise sosyal botların zararlı yazılım yayma, istenmeyen elektronik posta gönderme, kimlik hırsızlığı, botnet saldırısı başlatma ve benzeri kötü faaliyetlerde kullanılmak için tasarlanabildiği gerçeği duruyor.

- Sosyal botların kötü niyetli kullanımlarından biri güçlü bir yanlış bilgilendirme yayıcısı olarak kullanılmasıdır. Örneğin, 2013 yılında Suriye Elektronik Ordusu, ABD haber ajansının (Associated Press) Twitter hesabını ele geçirdi. Beyaz Saray'ın saldırıya uğradığını ve Obama'nın yaralı olduğunu açıkladı. Bu sahte haber, önce kısa süreli de olsa büyük bir paniğe ve sonuç olarak borsada 136,5 milyar dolar gibi büyük miktarda maddi kayba yol açmıştır [7].
- Sosyal botlarla yapılan başka bir kötü niyetli faaliyet ise propagandadır. Sosyal botlar propagandanın en etkili ve kolay yoludur. Literatürde bu faaliyet "suni kamuoyu oluşturma" (astroturfing) olarak adlandırılır. Başka bir deyişle, bir politikayı, bireyi, ya da ürün kampanyasını desteklemek için kamuoyu üzerinde sahte bir izlenim verme girişimidir [8]. Ratkiewicz ve arkadaşlarının yaptığı araştırma [9], 2010 ABD ara seçimleri sırasında Twitter'ın bu tür kampanyalar tarafından nasıl kullanılabileceğini göstermektedir. Boshmaf ve arkadaşlarına göre [10], sosyal ağlar 2011'de Ortadoğu'da Arap Bahar'ının ortaya çıkmasındaki en önemli unsurlardan biridir. Buna ek olarak, otomatik propagandanın Trump ve Clinton arasındaki 2016 seçimini etkileyip etkilemediği konusu da bir endişe konusudur [11]. Bu olasılıklara göre, sosyal botların toplumsal devrimler için çok güçlü araçlar olabileceğini söyleyebiliriz.
- Başka bir olumsuz yön ise; sosyal botların sahte reyting ve değerlendirme yazıları elde etmek için kullanılabilir olmasıdır. Örneğin, bu amaca hizmet eden nüfuz (influence) botları vardır. Bununla birlikte, arama motorları aracılığıyla sahte takipçi ve beğeni sağlayan, hatta aralarında ücretsiz de olan pek çok servis bulmak mümkündür. Subrahmanian ve arkadaşları [3] bazı politikacıların sosyal medya üzerinde nüfuz satın almaktan dolayı suçlandığını belirtmektedir.
- Bu faaliyetlere ek olarak; sosyal bot gerçek kişi veya organizasyon üzerinden kimlik hırsızlığı

[1] http://www.apple.com/ios/siri/

yapabilir. Kimlik hırsızlığının kullanıldığı kötü amaçlardan biri de bazı ideolojileri teşvik etmektir. Bu teşvik yoluyla saldırganlar, sosyal ağdaki kişileri yanıltma veya gerçek görünümlü sahte kimlikler oluşturma gücü kazanırlar. Daha sonra, bu hesapların takipçilerini dolandırma [3,12] ya da basit sosyal hesaplardan oluşan bot orduları olarak, Sybil saldırıları gibi kötü niyetli faaliyetlerde kullanılabilirler [13].

Sonuç olarak, literatürde odaklanılan temel soru "çevrim içi ağlarda kötü niyetli faaliyetlerin nasıl tespit edileceği"dir. Bu konu üzerine yapılan pek çok çalışma vardır. Bu makalede, ikinci kısımda kullanılan yaklaşımları inceledikten sonra, onları metodolojik olarak sınıflandırılmıştır. Üçüncü kısımda araştırmacıları bilgilendirmek ve açık problem alanlarına motive etmek için her bir yaklaşımın sınırlayıcı zorluklarını ve olası araştırma alanlarını sunulmuştur. Çalışmamızın dördüncü bölümünde, problemin çözümüne yönelik iki yeni yaklaşım önerilmiştir.

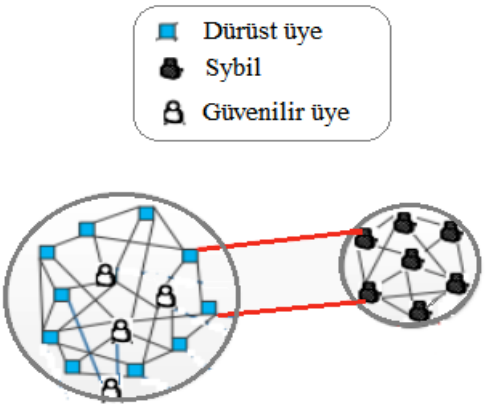
2. SOSYAL BOT ALGILAMA TEKNİKLERİ (SOCIAL BOT DETECTION TECHNIQUES)

Sosyal botların yukarıda belirtilen tüm kötü niyetli kullanımları nedeniyle, araştırmacılar onların en kesin şekilde tespit edilebilmesi için ileri seviyeli teknikler geliştirmektedir. Genel olarak, bu algılama tekniklerini üç sınıfa ayırmak mümkündür: (A) Sosyal ağ topolojisine (yani yapı temeline) dayalı bot algılama sistemleri, (B) kitle kaynak kullanımına dayalı sistemler ve (C) öznitelik tabanlı makine öğrenme yöntemlerine dayalı sistemler.

A. Sosyal Ağ Topolojisinde (Yapı Temelli) Bot Algılama Teknikleri

Sybil hesapları; dışardan bir kişi tarafından kontrol edilen çok sayıda sosyal ağ hesaplarıdır. "Sybil" terimi adını, çoklu kimlik bozukluğu tanısı konan bir kadının isminden almaktadır [14]. Bu kontrol altında tutulan hesaplar, sosyal ağlara sızmak, özel verileri çalmak, yanlış bilgilendirmeleri ve zararlı yazılımları yaymak için kullanılır. Bu nedenle, Sybil saldırıları sosyal ağlar için temel tehditlerden birisidir [15-17]. Örneğin, Facebook 2015 yılında Sybil hesapları olarak yaklaşık 170 milyon sahte hesabı belirliyor ve siliyor [18]. Bazı Sybiller, kullanıcı anonimliği korumak gibi iyi niyetli amaçlarla bilinçli olarak oluşturulabilir. Ancak bu çalışma kapsamında sadece kötü amaçlı olan Sybilleri değerlendiriyor ve iyi niyetli olanları dikkate almıyoruz.

Bu sınıftaki bot algılama tekniğinde, Sybil hesapların ağ üzerinde nasıl yayıldığını bilmek, onları algılamak için oldukça önemlidir. Yapı temelli algılama tekniklerinin temel varsayımı çevrimiçi sosyal ağların genellikle *benzer türlük eğilimi* (*homophily tendency*) göstermesidir [17]. Başka bir deyişle, sosyal ağ üzerinde birbiri ile ilişkisi olan iki hesap, benzer özellikler gösterme eğilimindedir. Yani Sybil hesaplar kendi gibi Sybil hesaplarla, dürüst hesaplar (gerçek ve iyi niyetli bir kişiye ait) diğer dürüst hesaplarla bağ kurma eğilimi gösterirler. Aslında bu varsayım Şekil 1'deki sezgiden kaynaklanır. Şekle baktığımızda dürüst ve Sybil hesaplar arasındaki kutuplaşmayı ve bu iki tür arasında sınırlı sayıda bağ kurulduğunu görebiliriz. Ancak, Sybil toplulukları kurdukları çok sayıda bağlantı sayesinde ağdaki dürüst üyeler üzerinde sahte bir güvenirlik etkisi bırakır. Şekil 1'de güvenilir olarak belirtilen hesapların zaten dürüst olduğunu ve bu sınıftaki bot algılama tekniklerinde referans hesaplar olarak kullanıldığını belirtmemiz faydalı olabilir.



Şekil 1. Dürüst, güvenilir ve Sybil üyeleri içeren bir çevrimi sosyal ağ)

Ağ topolojisinin analizi yerel toplulukların belirlenmesinde kullanılan bir yoldur. Sosyal ağ yapısını (topolojisini) kullanarak Sybil hesapların

algılanması problemini çözmek amacıyla yapılmış pek çok çalışmaya literatürde rastlamak mümkündür. Yapılan çalışmaları kullandıkları yöntemlere göre özetleyebiliriz:

- **Rasgele Yürüyüş (Random Walk) [19]**

Genel olarak, Sybil hesapları dürüst hesaplardan ayırt etmek için rasgele yürüyüş metodunun kullanılmasının arkasındaki önsezi, sosyal ağların "hızlı ayrıştırılması"dır (fast mixing). Bu kapsamda ağın hızlı ayrışması şu anlama gelir: Eğer başlangıç noktası olarak dürüst bir hesap seçer ve kısa rasgele yürüyüşler yaparsak çabukça diğer dürüst hesaplara bir grup olarak erişip ayrıştırabiliriz; ancak eğer başlangıç noktası olarak bir Sybil hesap seçersek ve kısa rasgele yürüyüşler yaparsak bir dürüst hesaba erişmemiz zordur [20]. Daha üst seviyeden bakacak olursak, rasgele yürüyüşü kullanan metotlar güvenilir bir hesabın bakış açısından diğer hesapları Sybil ya da dürüst diye etiketler, ayrıştırır diyebiliriz.

Rastgele yürüyüş metodunu kullanan uygulamalardan biri olan SybilInfer [21], dürüst hesapları ve Sybil hesapları tahmin etmek için Bayes çıkarım ve Monte-Carlo örnekleme tekniklerini harmanlayarak kullanmaktadır. SybilInfer dürüst kullanıcı ve Sybil bölgeler arasında bir darboğaz kesiti tespit eder ve bu kesitten itibaren bu bölgeleri dürüst ya da Sybil diye etiketler. SybilGuard [22], kötü niyetli kullanıcının birçok Sybili oluşturabileceği varsayımını benimser, ancak Sybil hesapların Şekil 1'deki gibi dürüst hesaplara sınırlı sayıda bağlantısı olabilir. Benzer şekilde, SybilLimit [23], Sybilleri rasgele yürüyüşlere dayalı olarak izole etmeye çalışır. SybilLimit, SybilGuard ile aynı görüşü benimser, ancak optimum doğruluğu garanti eder. SybilRank [24] hesapları, algılanan Sybil olma olasılığına (kısa rastgele yürüyüşlerin iniş olasılığına) göre sıralar. Çünkü, güvenilir bir hesaptan başlayarak kısa rastgele bir yürüyüşle Sybil bölgesine geçme olasılığı sınırlıdır.

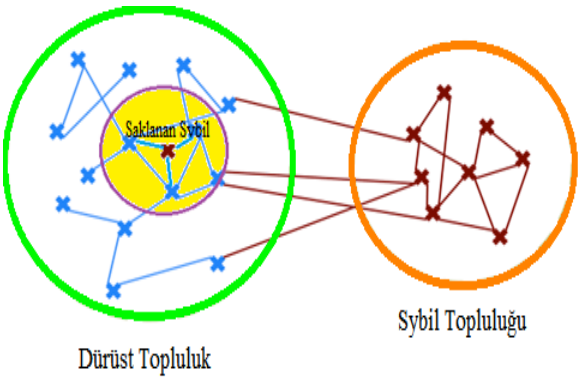
- **Markov Rastgele Alan (Markov Random Field) [25] ve Döngüsel İnanç Yayılımı (Loopy Belief Propagation) [26]**

Sosyal ağların hızlı bir biçimde ayrıştırıldığı varsayımı, ağın büyük bir topluluk veya küme olmasını gerektirir. Bununla birlikte, Mohaisen ve arkadaşları [27], çevrimiçi sosyal ağın genel olarak hızlı ayrıştırılmadığını gösterdiler. Benzer şekilde, Leskovec ve arkadaşları [28], ağın bir büyük küme (topluluk) oluşturmak yerine çok sayıda küçük topluluk oluşturduğunu gösterdiler. Bu nedenle Viswanath ve arkadaşları [29], Sybil tespit probleminin topluluk tespit problemi olarak görülebileceğini belirtmektedirler. Ayrıca, Boshmaf ve arkadaşları [30], yapı temelli Sybil algılama algoritmalarının daha iyi performans göstermesi için güvenilir hesaplar etrafında dinamik olarak yerel toplulukları bulmak üzere tasarlanması gerektiğini belirtirler.

Ayrıca, Viswanath ve arkadaşları [29] dürüst hesap topluluklarının birbirleriyle sıkı bağlar kurduğu ağlarda, topluluk bulmaya yönelik bot algılama tekniklerinin olası Sybil saldırılarına karşı daha savunmasız olduklarını keşfettiler. Çünkü; Sybiller bağ kurmak için dürüst hedefleri akıllıca seçerek dürüst hesapların topluluğuna sızıp, kendilerini bu topluluğun elemanı olarak gösterebilirler. Başka bir deyişle, Sybiller sadece az sayıda dürüst hesapla ilişki kurarak kendisini dürüst hesap topluluğu içerisinde gizleyebilir. Bu zayıflık Şekil 2'de resmedilmiştir. Yazarlar, Sybil hesapların rasgele hesaplarla bağlantı kurması yerine, hedeflerini güvenilir hesaplara yakın olacak şekilde seçmelerine izin verdikleri bir deney yaparlar. Burada yakınlıktan kasıt; kullanılan topluluk bulma algoritmasının, güvenilir hesabın bakış açısından baktığında, dürüst hesap olma sıralamasında, önde olması demektir. Dolayısı ile bu şekilde Sybiller kendilerini sıralamada üstlere çıkartabiliyorlar. Doğal olarak, bu tür Sybilllerin algılanma olasılığı oldukça azalır, çünkü Sybiller güvenilir hesabın bakış açısıyla yerel topluluğun üyeleri olarak gözükürler.

Sosyal ağların hızlı ayrıştırılabilir olmadığının kanıtlanmasından dolayı, yeni çalışmalar yapılmaya devam edilir. SybilBelief [17] ve SybilFrame [15] ağdaki hesapların dürüst olma olasılıklarını tahmin etmek için ağ üzerinde rasgele yürüyüşler yerine Markov rastgele alan ve döngüsel inanç yayılımı

yöntemlerini kullanırlar.



Şekil 2. Saldırganların topluluk algılama ile hedeflerini akıllıca seçme zayıflığı

SybilBelief bilinen dürüst ve Sybil hesaplar hakkındaki bilgileri bünyesinde birleştirirken, SybilFrame kullanıcıların yerel bilgilerini ve global ağ yapısındaki ilişkileri de kullanarak çok aşamalı bir sınıflandırma mekanizması kullanır. Yapı temelli bot algılama tekniklerini kullanan çalışmalar içerisinde en yüksek bot algılama oranı %68,2 ile SybilFrame'e aittir [15].

Buna ek olarak, SybilRank gibi yapı temelli bot algılama sistemlerinden bazıları "ilişki sebebiyle masum" (innocent by association) paradigmasını da kullanmaktadır [31]. Bu paradigmaya göre, masum bir hesapla ilişki kuran diğer hesap ta bu ilişki sebebiyle masumdur. Yukarıda anlattığımız şekilde dürüst topluluğun içine akıllıca sızan Sybiller için geçersiz bir yaklaşımdır. Bu paradigmanın etkinliği, dürüst kullanıcıların LinkedIn ağında olduğu gibi, bilmediği hesaplarla ilişki kurmayı reddetmesi ile sınırlıdır. Öte yandan, Twitter ve Renren (Çin'deki en büyük çevrimiçi sosyal ağ) gibi gerçek dünyadaki bazı sosyal ağlar, hesapların birbirleriyle ilişki kurabilmesi için böyle güçlü bir güven varsayımı gerektirmiyorlar. Birbirlerini hiç tanımayan kişiler bile birbirlerini izleyebilir, arkadaş olabilir, birbirlerinin iletilerini görebilir ve yorum yazabilirler. Bu nedenle, "ilişki sebebiyle masum" paradigması güçlü bir güven ağı olmadığı sürece uygulanabilir değildir ve yüksek yanlış negatif (false-negative, FN) oranı üretir.

B. Kitle Kaynak Kullanımına Dayalı Sosyal Bot Algılama Teknikleri

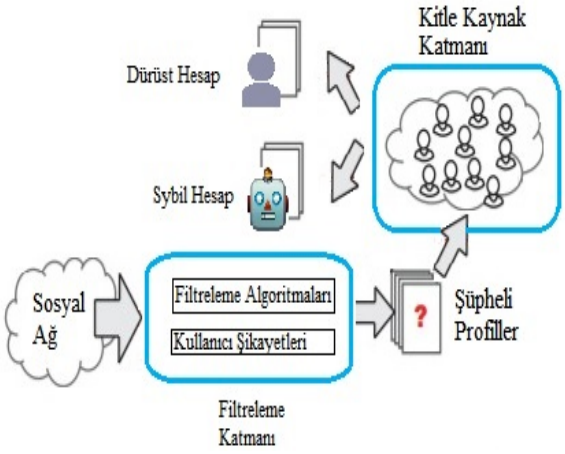
Jiang ve arkadaşları [32], Sybillerin nadiren diğer Sybiller ile bağlantı kurduğunu göstermişlerdir. Bunun yerine, Sybiller güvenilir kullanıcıların topluluklarına sızmayı hedeflemektedirler [33]. Dolayısı ile yukarıda açıkladığımız gibi; yapı temelli Sybil algılama tekniklerinin başarısı (i) saldırgan topluluk algılama yeteneğine karşı, Sybiller için dürüst hedefleri akıllıca seçebilme zayıflığı (Viswanath ve arkadaşları [29]) ve (ii) botların makine öğrenmesi ile insan hesaplarını daha gerçekçi taklit edebilmeleri nedeniyle azalmıştır.

Wang ve arkadaşları [34] tıpkı Amazon'un Mekanik Türk'ü [35] gibi insan emeğinin (kitle kaynak kullanımının) sosyal ağlardaki hesapları dürüst ya da bot diye etiketlemesi yaklaşımını önermişlerdir. Yazarların bu yaklaşımı önermelerinin arkasında yatan sezgileri; makinelere göre dikkatli insanlar tarafından, hesapların ağda yaptığı paylaşımlardaki ve profillerindeki en küçük tutarsızlıkların daha kolay yakalanabilmesidir.

Yazarlar, filtreleme ve kitle kaynak kullanımını içeren iki katmanlı bir sistem önermekteler. Önerilen bu sistem Şekil 3'de görülmektedir. Yazarlar şüpheli hesapları belirlemek için filtreleme katmanında topluluk bulma, ağ temelli özellik seçmeye dayalı makine öğrenmesi algoritmaları ve kullanıcı şikâyetleri gibi bilinen otomasyon tekniklerini kullanmayı önermişlerdir. Sonrasında, bu şüpheli hesaplar bir analist kitle tarafından incelenir ve bu hesapların dürüst ya da Sybil olduğuna ilişkin son sınıflandırma bu analistler tarafından yapılır. Yazarlara göre stratejileri simüle ettikleri bin tanesi dürüst, bin tanesi Sybil hesaptan oluşan sistemde hem yanlış pozitif hem de yanlış negatif oranı olarak %1'in altındadır.

Bununla birlikte, sosyal bot algılama için kitle kaynak kullanımı hakkında göz önünde bulundurulması gereken üç temel husus vardır. Birincisi, sosyal ağ kullanıcılarının mahremiyeti göz önünde bulundurulmalı ve analist kitle ile paylaşmadan önce

hesap sahiplerinin kişisel bilgileri gizlenmelidir. İkincisi, büyük çevrimiçi sosyal ağ şirketleri sadece bu iş için ilave uzman analistler çalıştırmak durumundadır. Küçük sosyal ağ şirketleri de bu ek masrafı karşılayamayabilir. Sonuç olarak, kitle kaynak kullanımının büyük sosyal ağlar için uygulaması kolay, pratik ve uygulanabilir değildir; çünkü bu ağların insan gücünün baş edebileceğinden çok daha fazla sayıda üyesi vardır. Analist kitle, hesapları dürüst ya da Sybil diye etiketleme sürecinde çok fazla zamana ihtiyaç duyacaktır.

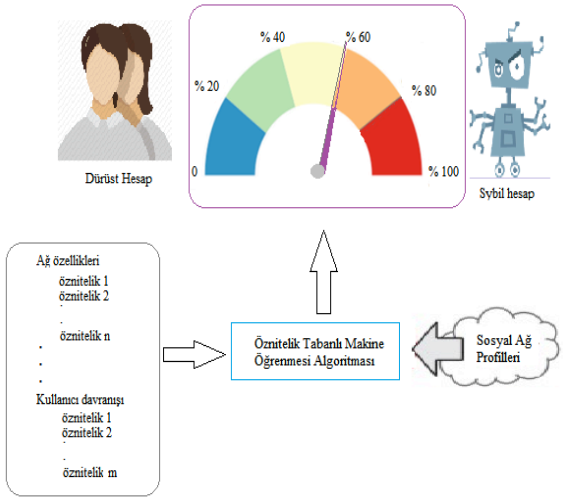


Şekil 3. Kitle kaynak kullanımına dayalı bir Sybil algılayıcı

C. Öznitelik Tabanlı Makine Öğrenmesi Yöntemlerine Dayalı Bot Algılama Teknikleri

Yapay Zekâ (AI) alanındaki büyük gelişmelerle sosyal botlar daha karmaşık ve gerçek bir kullanıcıdan ayırt edilemeyecek kadar benzetimde başarılı olmuşlardır. Bu nedenle, sosyal ağlardaki bu yapay hesapları saptamak oldukça zorlu bir iştir. Hatta, DARPA sosyal botların algılanması ile ilgili bir yarışma düzenlemiştir. Yapısal ağ tabanlı bot algılama teknikleri bu yarışma sonucunda yeterince faydalı bulunmamışlardır [3]. Yapay zekânın alt dalı olarak makine öğrenmesindeki ilerlemeler sosyal bot algılama problemine de yeni bir çözüm yaklaşımını beraberinde getirmiştir. Bu yaklaşımdaki temel fikir; bir insan tarafından kontrol edilen hesap ve bir makine öğrenmesiyle oluşturulan Sybil hesap arasındaki ayırt edici sınırı çizebilmek için, bu iki hesap karakterini oluşturan özniteliklerin bulmasıdır. Şekil 4'te bu yaklaşımın işleyişi görülmektedir. Bulunan özniteliklere göre, sosyal ağ profilleri öznitelik tabanlı bir makine öğrenmesi algoritmasına gönderilerek hesaplar dürüst ya da Sybil hesap olarak sınıflandırılmaktadır. Literatürdeki bu teknikle ilişkili seçilmiş bazı çalışmaların özetleri aşağıda bulunabilir.

Chu ve arkadaşları [2], insan, bot ve sayborg [2] (Cyborg) hesaplarının profillerini çıkarmak amaçlı bir çalışma yapmışlardır. Yazarlar, hesapların attıkları tweet içerikleri, tweet atma davranışları ve harici sayfalara link verme vb. gibi bazı özelliklerde bu üç tipin farklılık gösterdiğini gözlemlemişlerdir. Lee ve arkadaşları [36] sosyal balküpü (honeypot) oluşturmak



Şekil 4. Öznitelik tabanlı makine öğrenmesi yöntemlerine dayalı bot algılama

için bir çalışma ortaya koymuşlardır. Balküpü, istenmeyen elektronik posta ve zararlı yazılım yayıcıları gibi sosyal medyadaki içerik kirleticilerin profillerinin çıkartılması ve bu profil özelliklerine göre filtreleme işleminin yapılması için kullanılır. Yang ve arkadaşları [33] öncelikle Renren ağındaki bilinen sosyal bot hesaplarını referans (ground-truth) olarak toplarlar. Sonrasında, öbeklenme katsayısı (clustering coefficient), gelen ve giden arkadaşlık talep oranı gibi ağ ve yapısal özellikleri kullanarak kullanıcı hesaplarını analiz ederler.

[2] Sayborg hesaplar: insan destekli botlar veya bot destekli insan hesapları

Sentibot [37] sosyal ağlarda birey hesaplarının ve Sybillerin sınıflandırılması problemini adresleyen bir çerçevedir. Sentibot, ortalama konu etiketi (hashtag) sayısı gibi sözdizimi ve ortalama topik duygu polaritesi gibi semantikler, tweet atma gibi kullanıcı davranışı ve bazı ağ merkezli kullanıcı özelliklerine dayanır. Sentibot'un geliştiricileri, duygularla ilişkili öznitelikleri ağdaki botların algılanması için anahtar bileşen olarak değerlendirirler ve çalışmalarında karmaşık duygu analizi tekniklerini kullanırlar.

"Bor or Not?" [38], ilk halka açık sosyal bot algılama çerçevesidir. Twitter ağı üzerinde 2014 yılından beri hizmet vermektedir. "Bot or Not?", altı ayrı kategoride binden fazla sayıda özniteliği analiz eder. Bu kategoriler, ağ, kullanıcı, arkadaş listesi, zamana bağlı içerikler ve duygu kategorisidir. Bu çerçevede, bu temel özniteliklere bağlı bir makine öğrenmesi uygulaması gerçekleştirilir. Uygulamanın basit ve karmaşık sosyal botlar için başarım oranı 2017 yılı için %86'dır ve öznitelik tabanlı makine öğrenmesine dayalı teknikler için en yüksek başarım oranıdır [39]. Makine öğrenmesi ve sosyal ağ yapısına ait özelliklerin birlikte kullanımıyla bu başarım oranının ancak elde edilebildiğini not etmekte fayda vardır.

3. SUNULAN YÖNTEMLERDEKİ AÇIK PROBLEMLER (OPEN PROBLEMS OF PRESENTED TECHNIQUES)

Çevrimiçi sosyal ağlarda botların algılanması zorlu bir konudur ve tam anlamıyla çözülmüş değildir. İkinci bölümde bahsettiğimiz her bir sınıfa özgü açık sorular ve olası araştırma yönleri vardır.

Çevrimiçi sosyal ağlar doğası gereği büyük miktarda veri içerirler ve dinamik olarak büyürler. Yapısal temelli bot algılama teknikleri genellikle durağan ortamda bile yüksek çalışma maliyetine sahiptirler. Bilinen en az çalışma maliyeti rasgele yürüyüş kullanan çalışmalarda bile $O(nlogn)$'dir. Bu nedenle çalışma maliyeti bakımından daha verimli, gerçek zamanlı ve büyük veri işleyebilecek çizge algoritmaları geliştirmek iyi bir araştırma alanı olabilir. Diğer bir problem, yapısal temelli tekniklerin günümüzde öznitelik tabanlı makine öğrenme yöntemlere dayalı tekniklere göre oldukça düşük doğrulukla, yüksek yanlış pozitif oranı vermesidir. Örneğin, ikinci kısımda belirttiğimiz gibi, SybilFrame %68,2'lik başarımı %4,2 yanlış pozitif oranıyla verir. Bu çalışma kapsamında, bot olarak etiketlenen gerçek kullanıcı oranı yanlış pozitifleri ifade ederken, gerçek kişi hesabı olarak etiketlenen bot oranları yanlış negatifleri ifade eder. Bot algılama sistemleri için yanlış pozitifler kullanıcı deneyimini olumsuz yönde etkiler; çünkü gerçek kullanıcı bot olarak etiketlenirse olumsuz tepki verip hesabını kapatabilir. Yanlış negatifler ise sistemin başarımını doğrudan etkiler. Bu nedenle sosyal ağ çizgesinde yanlış pozitif ve yanlış negatif oranlarını düşürmek için yeni makine öğrenmesi yaklaşımları uygulanabilir. Topluluk bulma yoluyla bot algılayan çalışmalar için, güvenilir hesapların başlangıçta bilinmesi gereksiniminin kaldırılıp, algılama sırasında zarar verebilecek mekanizmaların geliştirilmesi başka bir açık sorudur. Öyle ki; sosyal ağın yapısı ve güvenilir hesapların belirlenmesi, yapı temelli bot algılama tekniklerinin başarısının merkezinde yer alır. Kaynak kitle kullanımına dayalı algılama çalışmaları yapay zekâ ile donatılmış sofistike botlara karşı insan zekasını kullanır. Sosyal ağ kullanıcılarının gizliliğini sağlamak, bu tip algılama teknikleri için göz önünde bulundurulması gereken önemli bir husustur. Bu nedenle, hesapları değerlendirecek olan analist kitlenin etik anlayışını arttırmak için ön bir çalışma yapılabilir ve gizliliği koruyan veri madenciliği teknikleri de kullanılarak kullanıcı hesapları anonim hale getirilebilir. Ancak, çevrim içi sosyal ağların doğası gereği hem dinamik hem de çok büyük miktarda veri içerdiğini değerlendirdiğimizde, bu tekniğin hem zaman hem de maliyet açısından etkin ve uygulanabilir olmadığı görülür.

Yapay zekâ, hem botların geliştirilmesinde hem de onların algılanmasında kullanılan, çift taraflı bir araçtır. Başka bir deyişle, yapay zekânın ilerlemesiyle botlar, insani davranışları daha iyi taklit etmekte sürekli gelişirken, bot algılama sistemleri de yapay zekâyı kullanarak ağdaki hesapları daha iyi analiz edip botların profillerini çıkarmada iyileşirler.

Tablo 1. Sosyal Bot Algılama Teknikleri ve Araştırma Yönlerinin Özetlenmesi

Bot Algılama Teknikleri	**İlgili Çalışmalar**	**Doğruluk (%)**	**Sınırlamalar**	**Açık Araştırma Alanları**
Ağ Yapı temelli teknikler	-SybilInfer [21] -SybilGuard [22] -SybilLimit [23] -SybilRank [24] SybilBelief [17] -SybilFrame [15]	68,2	-Çevrimiçi sosyal ağların büyük miktarda veri içermesi -Çevrimiçi sosyal ağların dinamik bir ortam olması -FN, FP oranlarının düşürülme ihtiyacı -Yüksek çalışma süresi maliyeti -En az bir tane güvenilir hesabın referans olarak bilinmesi gerekliliği	-Daha verimli büyük veri işleme algoritmaları geliştirme -Dinamik ya da gerçek zamanlı bot algılama metotları geliştirme -FN, FP oranlarının düşürülmesine yönelik yeni metotlar geliştirme -Güvenilir hesapların çalışma anında belirlenmesi
-Kitle kaynak kullanımına dayalı teknikler	-Wang ve arkadaşlarının çalışması [34]		-Örnekleme yöntemi ile sınırlı boyutta analiz imkânı -Kullanıcı mahremiyeti ilgili sorunlar -Kitlenin yüksek çalışma süresi maliyeti -Kitle işgücünün maliyeti	
-Öznitelik tabanlı makine öğrenme yöntemlerine dayalı teknikler	-Chu ve arkadaşlarının çalışması [2] -Lee ve arkadaşlarının çalışması [36] -Yang ve arkadaşlarının çalışması [33] -SentiBot [37] -Bor or Not? [38]	86	-Çevrimiçi sosyal ağların büyük miktarda veri içermesi -Çevrimiçi sosyal ağların dinamik bir ortam olması -Botların da makine öğrenmesi kullanması -FN, FP oranlarının düşürülme ihtiyacı -Sybilleri yöneten kaynağın bilinmiyor olması	-Dinamik ya da gerçek zamanlı bot algılama metotları geliştirme -FN, FP oranlarının düşürülmesine yönelik yeni metotlar geliştirme -Bit botu dürüst hesaptan ayıracak özniteliklerin belirlenmesinde derin öğrenme gibi büyük veri de işleyebilecek güncel makine öğrenmesi tekniklerinin kullanımı -Sybil hesapların kaynağının bulunması (Source detection)

Öznitelik tabanlı makine öğrenme tekniklerine gelince, var olan özniteliklere ek olarak botların profil çıkartımının daha da iyileştirilmesi için derin öğrenme gibi en güncel makine öğrenmesi yöntemleri kullanılarak yeni öznitelikler araştırılabilir. Sybil hesaplar bir saldırganın elindeki sahte hesaplar olduğuna göre, başka bir açık soru ise bu hesapların kimin tarafından yönetildiğidir. Dolayısıyla, Sybillerin kaynak tespiti, önemli açık sorulardan biridir.

4. SUNULAN YÖNTEMLERİN KARŞILAŞTIRIMI VE DEĞERLENDİRİLMESİ (COMPARISON AND EVALUATION OF REPRESENTED TECHNIQUES)

Çevrim içi sosyal ağlar dünyanın her yerindeki milyonlarca insanı birbirine bağlayan güçlü araçlardır; bu sebeple sosyal botlar için de oldukça caziptirler. Çoğunlukla sosyal ağlarda kötü aktivitelerde bulunmaları sebebiyle, sosyal bot hesaplarının gerçek kullanıcılardan ayırt edilme zorunluluğu doğmuştur. Bu kötü kullanım alanlarından bazıları kimlik hırsızlığı, sahte halk tabakası oluşturarak propaganda yapmak, zararlı yazılımları ağda yaymak, takipçi dolandırıcılığı ve yanıltıcı haber yaymak vb. şeklinde söylenebilir.

Tablo 1'de sosyal bot algılama tekniklerinin sınıflandırması, her bir sınıf için literatürdeki ilgili çalışmalar, sınırlayıcı zorluklar ve olası araştırma yönleri özetlenmiştir. Çevrim içi ağların hem dinamik bir ortam hem de büyük veri içeriyor olmasından dolayı, olası çözümlerin hem büyük veriyi işleme hem de botların dinamik olarak algılanması için çalışma süresi bakımından verimli bir şekilde ele alınması gerekir. Bununla birlikte, çözüm adayları, mevcut çözümlerin yanlış negatif ve yanlış pozitif oranlarını düşürmeli ve doğruluğu mümkün

olduğunca arttırmalıdır. Yapı temelli bot algılama tekniklerinde en az bir güvenilir (referans) hesaba ihtiyaç duyulduğundan, bu referans noktayı ya da noktaları çalışma anında belirlemek bir başka araştırma yönü olabilir. Yapı temelli teknikler için bilinen en yüksek başarım oranı SybilFrame tarafından sağlanan %68,2'dir.

Kitle kaynak kullanımına yönelik teknikler hem zaman hem de analist kitle maliyeti bakımından yüksek olması nedeniyle uygulanabilir değildir. Dolayısıyla, bu teknik için tabloda ne başarım oranı ne de açık araştırma alanı kısmı doldurulmamıştır. Öznitelik tabanlı makine öğrenmesi tekniklerine gelince, bunların ana sınırlayıcı zorluğu güçlerini yapay zekâdan alıyor olmalarıdır, ancak botlar da güçlerini yapay zekâdan almaktadırlar. Bununla beraber, yine çözüm; Sybil hesaplarla dürüst kullanıcı hesaplarını ayırt edebilecek öznitelikleri keşfedebilecek "derin öğrenme" gibi gelişmiş yapay zekâ yöntemleridir. Aynı zamanda, Sybillerin kaynaklarının bulunması diğer bir araştırma alanıdır. Öznitelik tabanlı makine öğrenmesi teknikleri arasında bilinen en yüksek başarım oranı "Bot or Not" tarafından sağlanan %86'dır.

5. ÖNERİLEN YENİ ÇÖZÜM VE ARAŞTIRMA YÖNLERİ (PROPOSED SOLUTION AND RESEARCH DIRECTIONS)

Sosyal bot algılama tekniklerindeki gelişmelerle birlikte, gelinen en yüksek başarım oranının (%80 den fazla) sosyal ağın yapısal özellikleri ve denetimsiz makine öğrenmesinin birlikte kullanımıyla elde edildiği görülür. Bunların dışında, bot algılama başarım oranını arttırmak için aşağıda önerdiğimiz konularda da araştırma yapmak faydalı olacaktır. Bunlar; (i) özerk akıllı ajan temelli yaklaşımlar ve (ii) kimliklendirme temelli yaklaşımlar olabilir. Bunlar aşağıda kısaca tanıtılmıştır.

(i) Özerk Akıllı Ajan Temelli Yaklaşımlar
Bu yaklaşımdaki ana fikir, sosyal ağlardaki toplulukları meydana getiren üyelerin algılanması ve kimliklendirilmesinin merkezi değil dağıtık bir ortamda gerçekleştirilmesidir. Başka bir deyişle, sosyal ağın yapısı göz önünde bulundurularak üyelerini algılama ve analiz işleri toplulukların kendilerine bırakılmalıdır. Aynı zamanda, bu dağıtık ortamda topluluk başına akıllı ajanlar olmalıdır. Hem topluluğun sınırları hem de üyelerinden haberdar olan bu ajanlar, sosyal ağ uzmanlarınca belirlenen karakter özelliklerine göre daimî olarak sorumlu olduğu topluluğun aktivitelerini gözetmelidirler. Blok-zincir yaklaşımının bu alanda projelendirilmesi değerli bir katkı sunabilir.

(ii) Kimliklendirme Temelli Yaklaşımlar
Sosyal bot algılama problemine çözüm olarak teknolojik veya yöntemsel olarak geliştirilmiş bir metot içerisinde, ağdaki hesapları kimliklendirme bileşeni bulundurmadığı sürece saldırganlar tarafından sömürülme potansiyeli vardır. Bu nedenle, Sybil hesapların otomatik ve dinamik olarak algılanması için sosyal ağda konuşlandırılmış akıllı ajanlar mutlaka güvenilir kimliklendirme mekanizmaları ile desteklenmelidir. Eğer bu mekanizmalar olmazsa, akıllı ajanlar saldırganlar tarafından hedef alınacaktır ve saldırganın kimliği belirlenemeyecektir. Ne yazık ki, bu saldırı sonucunda dürüst hesaplar Sybil, Sybil hesaplar dürüst olarak etiketlenebilir. Dolayısıyla, bu durum başka bir saldırı problemi ile sonuçlanacaktır.

6. SONUÇ (CONCLUSION)

Bu çalışmada, sosyal ağlarda bot algılamanın üç sınıfı (sosyal ağ topolojisine dayalı bot algılama sistemleri, kitle kaynak kullanımına dayalı sistemler ve öznitelik tabanlı makine öğrenme yöntemlerine dayalı sistemler), her bir sınıfın başarım oranları, sınırlayıcı zorlukları ve araştırma yönleri incelenmiştir. İnceleme sonucunda, bu sınıflardan öznitelik temelli bot algılama tekniğinin %86 başarım oranı ile bilinen tekniklerden en etkili ve yaygın kullanıma sahip olduğu görülmüştür. Diğer yandan makine öğrenmesindeki ilerleme, gerçek zamanlı büyük veri işleme ve Sybil hesapların kaynağının bulunması için küresel bir kimliklendirme sistemine duyulan ihtiyaç bizi yeni bir çözüm önerme konusunda motive etmiştir. Bu yüzden, araştırmacılar için hem var olan tekniklerde iyileştirme yapılabilecek yönler belirtilmiş hem de yeni iki yaklaşım önerilmiştir.

KAYNAKLAR (REFERENCES)

[1] Dehade, S. K. ve Bagade, A. M., "A review on detecting automation on Twitter accounts," ***Eur. J. Adv. Eng. Technol,*** Cilt 2, 69-72, 2015.

[2] Chu, Z., Gianvecchio, S., Wang,H., ve Jajodia, S., "Detecting automation of twitter accounts: Are you a human, bot, or cyborg?," ***IEEE Transactions on Dependable and Secure Computing,*** Cilt 9, 811-824, 2012.

[3] Subrahmanian, V., Azaria, A., Durst, S., Kagan, V., Galstyan, A. ve Lerman, K.,"The darpa twitter bot challenge," ***arXiv preprint*** *arXiv:1601.05140,* 2016.

[4] Haugen, G., **Manipulation and Deception with Social Bots: Strategies and Indicators for Minimizing Impact,** Master Tezi, Norwegian University of Science and Technology, Faculty

of Information Technology and Electrical Engineering, 2017.

[5] Ferrera, E.,"The Rise of Social Bots," 2016. https://vimeo.com/166538072 (Erişim Tarihi: 13 Eylül 2016).

[6] Freitas, C., Benevenuto, F., Ghosh S., ve Veloso, A., "Reverse engineering socialbot infiltration strategies in twitter," ***2015 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining***, Paris, France, 25-32, 25-28 Ağustos 2015.

[7] Mail, D., "Syrian Electronic Army linked to hack attack on AP Twitter feed that 'broke news' Obama had been injured in White House blast and sent Dow Jones plunging," 2013. http://www.dailymail.co.uk/news/article-2314001/Syrian-Electronic-Army-linked-hack-attack-AP-Twitter-feed-broke-news-Obama-injured-White-House-blast-sent-Dow-Jones-plunging.html (Erişim Tarihi: 5 Mayıs 2016)

[8] Bienkov, A., "Astroturfing: what is it and why does it matter?," 2012. https://www.theguardian.com/commentisfree/2012/feb/08/what-is-astroturfing (Erişim Tarihi: 5 Mayıs 2016)

[9] Ratkiewicz, J., Conover, M., Meiss, M., Gonçalves, B., Flammini, A., ve Menczer, F., "Detecting and Tracking Political Abuse in Social Media," **Fifth International AAAI Conference on Weblogs and Social Media (*ICWSM-11)*,** Barcelona, Spain, 297-304, 17-21 Temmuz, 2011.

[10] Boshmaf, Y., Muslukhov, I., Beznosov, K., ve Ripeanu, M., "The socialbot network: when bots socialize for fame and money," ***27th Annual Computer Security Applications Conference***, Orlando, Florida, 93-102, 5-9 Aralık 2011.

[11] Schreckinger, B.,"Inside Trump's 'cyborg' Twitter army". http://www.politico.com/story/2016/09/donald-trump-twitter-army-228923, 2016. (Erişim Tarihi: 30 Eylül 2016).

[12] Goga, O., Venkatadri, G., ve Gummadi, K. P., "The doppelgänger bot attack: Exploring identity impersonation in online social networks," ***2015 ACM Conference on Internet Measurement Conference*****,** Tokyo, Japan, 141-153, 28-30 Ekim 2015.

[13] Abokhodair, N., Yoo, D. ve McDonald, D. W. "Dissecting a social botnet: Growth, content and influence in Twitter", ***18th ACM Conference on Computer Supported Cooperative Work & Social Computing***,Vancouver, Canada, 839-851, 14-18 Mart, 2015.

[14] Mulla, M., Sambare, S., "A Survey on Sybil Attacks with Solutions in Mobile Adhoc Networks", ***International Journal of Innovations & Advancement in Computer Science(IJIACS)***, Cilt 4, No 1, 47-53, 2015.

[15] Gao, P., Gong, N. Z., Kulkarni, S., Thomas, K., ve Mittal, P., "Sybilframe: A defense-in-depth framework for structure-based sybil detection," ***arXiv preprint*** *arXiv:1503.02985,* 2015.

[16] Mulamba, D., Ray, I., ve Ray, I., "SybilRadar: A Graph-Structure Based Framework for Sybil Detection in On-line Social Networks," **ICT Systems Security and Privacy Protection**, cilt 471, Hoepman JH. Ve Katzenbeisser S.(editörler), Springer, Cham, 179-193, 2016.

[17] Gong, N. Z., Frank, M., ve Mittal, P., "Sybilbelief: A semi-supervised learning approach for structure-based sybil detection," ***IEEE Transactions on Information Forensics and Security,*** Cilt 9, 976-987, 2014.

[18] Parsons, J., "*Facebook's War Continues Against Fake Profiles and Bots*," Eylül 12 2015. http://www.huffingtonpost.com/james-parsons/facebooks-war-continues-against-fake-profiles-and-bots_b_6914282.html (Erişim Tarihi: 7 Mayıs 2016)

[19] Pearson, K., "The problem of the random walk," ***Nature,*** Cilt 72, 294, 1905.

[20] Carminati, B., Ferrari, E., ve Viviani, M., "Security and trust in online social networks," ***Synthesis Lectures on Information Security, Privacy, & Trust,*** Cilt 4, 1-120, 2013.

[21] Danezis, G. ve Mittal, P., "SybilInfer: Detecting Sybil Nodes using Social Networks," **16th Annual Network & Distributed System Security Symposium** (***NDSS)***, San Diego, California, 1-15, 8-11 Şubat February 2009.

[22] Yu, H., Kaminsky,M., Gibbons, P. B. ve Flaxman, A., "Sybilguard: defending against sybil attacks via social networks," ***ACM SIGCOMM Computer Communication Review***, 267-278, 2006.

[23] Yu, H., Gibbons, P. B., Kaminsky, M. ve Xiao, F., "Sybillimit: A near-optimal social network defense against sybil attacks", ***IEEE/ACM Transactions on Networking (TON)***, Cilt 18, Sayı 3, 885-898, 2010.

[24] Cao, Q., Sirivianos, M., Yang, X., ve Pregueiro, T., "*SybilRank,"* 2016. http://www.tid.es/research/areas/sybil-rank (Erişim Tarihi: 15 Mayıs 2016)

[25] Cross, G. R. ve Jain, A. K., "Markov random field texture models," ***IEEE Transactions on Pattern Analysis and Machine Intelligence****,* 25-39, 1983.

[26] Murphy, K. P., Weiss,Y., ve Jordan, M. I., "Loopy belief propagation for approximate inference: An empirical study," ***Fifteenth conference on Uncertainty in artificial intelligence(UAI'99)***, Stockholm, Sweden, 467-475, 30 Temmuz – 1 Ağustos 1999.

[27] Mohaisen, A., Yun, A., ve Kim, Y., "Measuring the mixing time of social graphs," ***10th ACM SIGCOMM conference on Internet measurement(IMC 2010)***, Melbourne, Australia, 383-389, 1-3 Kasım 2010.

[28] Leskovec, J., Lang, K. J., Dasgupta A., ve Mahoney, M. W., "Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters," ***Internet Mathematics****,* Cilt 6, 29-123, 2009.

[29] Viswanath, B., Post, A., Gummadi, K. P., ve Mislove, A., "An analysis of social network-based sybil defenses," ***ACM SIGCOMM Computer Communication Review****,* Cilt 40, 363-374, 2010.

[30] Boshmaf, Y., Beznosov, K., ve Ripeanu,M., "Graph-based sybil detection in social and information systems," ***Advances in Social Networks Analysis and Mining (ASONAM 2013)****,* Niagara Falls, Canada, 466-473, 25-28 Ağustos 2013.

[31] Xie, Y., Yu, F., Ke, Q., Abadi,M., Gillum, E. ve Vitaldevaria, K.,"Innocent by association: early recognition of legitimate users,"***2012 ACM conference on Computer and communications security (CCS'12)***, Raleigh, North Carolina, 353-364, 16 - 18 Ekim 2012.

[32] Jiang, J., Wilson, C., Wang, X., Sha, W., Huang, P., Dai,Y., "Understanding latent interactions in online social networks," ***ACM Transactions on the Web (TWEB)****,* Cilt 7, 18, 2013.

[33] Yang, Z., Wilson, C., Wang,X., Gao,T., Zhao,B. Y., ve Dai, Y., "Uncovering social network sybils in the wild," ***ACM Transactions on Knowledge Discovery from Data (TKDD)****,* Cilt 8, 2, 2014.

[34] Wang, G., Mohanlal, M., Wilson, C., Wang,X., Metzger, M., Zheng, H.,"Social turing tests: Crowdsourcing sybil detection," ***arXiv preprint*** *arXiv:1205.3856,* 2012.

[35] *Overview of Mechanical Turk*, 2016. http://docs.aws.amazon.com/AWSMechTurk/latest/RequesterUI/OverviewofMturk.html (Erişim Tarihi: 30 Eylül 2016)

[36] Lee, K., Eoff, B. D., ve Caverlee, J., "Seven Months with the Devils: A Long-Term Study of Content Polluters on Twitter," **Fifth Annual Conference on Weblogs and Social Media** (***ICWSM 2011)***, Barcelona, Spain, 185-192, 17-21 Temmuz 2011.

[37] Dickerson, J. P., Kagan, V., ve Subrahmanian, V., "Using sentiment to detect bots on Twitter: Are humans more opinionated than bots?," ***Advances in Social Networks Analysis and Mining (ASONAM 2014)****,* Beijing, China, 620-627, 17-20 Ağustos 2014.

[38] Davis, C. A., Varol,O., Ferrara, E., Flammini, A., ve Menczer, F., "Botornot: A system to evaluate social bots," ***25th International Conference Companion on World Wide Web(WWW'16)*****,** Montreal, Canada, 273-274, 11-15 Nisan 2016.

[39] Varol, O., Ferrara, E., Davis, C. A., Menczer, F., ve Flammini, A., "Online human-bot interactions: Detection, estimation, and characterization," ***arXiv*** *preprint arXiv:1703.03107,* 2017.