Publications

1. [Weakly Supervised Extraction of Computer Security Events from Twitter]

(https://aritter.github.io/twitter_security.pdf)

Twitter verilerini kullanarak sabit kategorilere bağımlı kalmadan kullanıcının istediği kategorileri ekleyebileceği ve çıktı olarak bu kategorilerdeki saldırılardan etkilenen kurbanları belirlenmesi üzerine bir çalışma.

Victim	Date	Category	Sample Tweet
namecheap	Feb-20-2014	DDoS	My site was down due to a DDoS attack on NameCheap's DNS server. Those are lost page hits man...
bitcoin	Feb-12-2014	DDoS	Bitcoin value dramatically drops as massive #DDOS attack is waged on #Bitcoin http://t.co/YdoygOGmhv
europe	Feb-20-2014	DDoS	Record-breaking DDoS attack in Europe hits 400Gbps.
barcelona	Feb-18-2014	Account Hijacking	Lmao, the official Barcelona account has been hacked.
adam	Feb-16-2014	Account Hijacking	@adamlambert You've been hacked Adam! Argh!
dubai	Feb-09-2014	Account Hijacking	Dubai police twitter account just got hacked!
maryland	Feb-20-2014	Data Breach	SSNs Compromised in University of Maryland Data Breach: https://t.co/j69VeJC4dw
kickstarter	Feb-15-2014	Data Breach	I suspect my card was compromised because of the Kickstarter breach. It's a card I don't use often but have used for things like that.
tesco	Feb-14-2014	Data Breach	@directhex @Tesco thanks to the data breach yesterday it's clear no-one in Tesco does their sysadmin housekeeping!

Table 6: Example high-confidence events extracted using our system.

Candidate eventleri aşağıdaki tablodaki gibi belirliyorlar.

Feature Category	Sample Feature	Event Category
keyword-context-left	security **breach**	Data Breach
victim-context-right	**X** admits	Data Breach
victim-context-right	**X** data breach affecting	Data Breach
keyword-context-both	DT **ddos** attack	DDoS
keyword-context-left	NNP NNP hit IN **ddos**	DDoS
victim-context-right	**X** getting ddos'd	DDoS
victim-context-both	PRP hacked **X** POS account	Account Hijacking
keyword-context-left	POS email was **hacked**	Account Hijacking
victim-context-right	**X** POS NNP account	Account Hijacking

Table 4: Examples of high-weight features. Context words other than nouns and verbs are replaced with their part-of-speech tags for better generalization.

Daha sonra bu eventlerden etkilenen mağdur kimse/kurum/programı bulmaya çalışıyorlar.

Victim	Date	# Mentions
spamhaus	2013/03/18	26
soca	2011/06/20	89
etrade	2012/01/05	76
interpol	2012/02/29	45
ustream	2012/05/09	911
virgin media	2012/05/08	15
pirate bay	2012/05/16	2,265
demonoid	2012/07/27	182
att	2012/08/15	2,743
sweden	2012/09/03	28
godaddy	2012/09/10	849
github	2013/07/29	102
reddit	2013/04/19	2,042
cia	2011/06/15	36
paypal	2010/12/10	57

Table 1: Seed instances for DDoS attacks

2. [Automatic Detection of Cyber Security Related Accounts on Online Social Networks: Twitter as an example]

(https://kar.kent.ac.uk/66861/7/SMSociety2018.pdf)

Yıldırım Beyazıt Üniversitesinden bir doktora öğrencisi ve bir Profesörün University of Kent Canterbury(UK) de çalışan bir profesörle yaptıkları bir çalışma. Machine learning teknikleri kullanarak Siber güvenlikle alakalı sosyal medya hesaplarının bulunup bulunamayacağını araştırmışlar. Twitterı da örnek sosyal medya hesabı olarak seçmişler.

3. [DDoS Event Forecasting using Twitter Data]

(https://www.ijcai.org/proceedings/2017/0580.pdf)

Twitter datalarını işleyerek henüz gerçekleşmemiş DDos saldırılarını tahmin etmeye yönelik bir yayın.

Post	Target
The Thai government is about to end their citizens Internet freedom.	Thai Government
soooooooooo basically they are all leaving soon, sony is mad as hell.	SONY

Table 1: Example tweets with attack targets.

Bu bilgiyi 6 popüler “supervised classification model” kullanarak elde etmeye çalışmışlar. Örneğin kullandıkları modellerden biri negative term count.

“Neg-Term-count is the baseline sentiment-based model, we count the negative words from tweets each day, forecasting an attack if the number of negative words is larger than a threshold α, which is the average number of negative words on training data.”

4. [Real Time Prediction of Drive by Download Attacks on Twitter](https://arxiv.org/pdf/1708.05831.pdf)

url kısaltma ile zararlı web sitelerini normal bir web sitesi olarak gösterip twitter'da kısaltılmış url'i paylaşarak zararlı web sitelerinin tıklanmasını önlemeye yönelik ne yapabilirizi araştırmışlar. Honeypot ile cpu ram kullanımı artışından detect etme gibi yöntemler denemişler.

5. [SONAR: Automatic Detection of Cyber Security Events over the Twitter Stream] (https://www.researchgate.net/publication/319048233_SONAR_Automatic_Detection_of_Cyber_Security_Events_over_the_Twitter_Stream)

Sonar adında otomatik olarak kendi kendine öğrenebilen bir framework geliştirmişler. Siber güvenlikle ilgili olayları twitter verilerini işleyerek yakalayabiliyor. Sisteme takip etmesi gereken bazı keywordleri veriyorlar, system siber güvenlikle ilgili olabilecek diğer keywordleri verdiğimiz keyworlerin yardımıyla bulabiliyor.

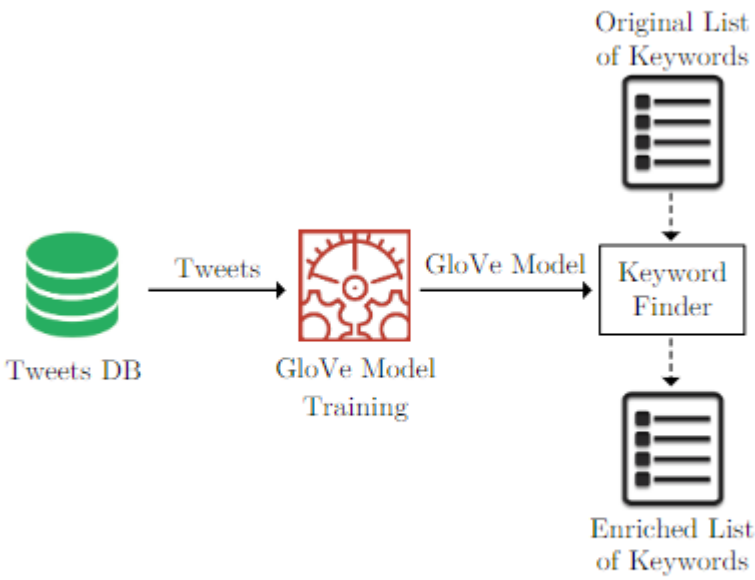

Figure 2: Architecture of the keyword finder component.

Araştırmalarını yapabilmek için birçok big data teknolojisinden de faydalanmışlar.

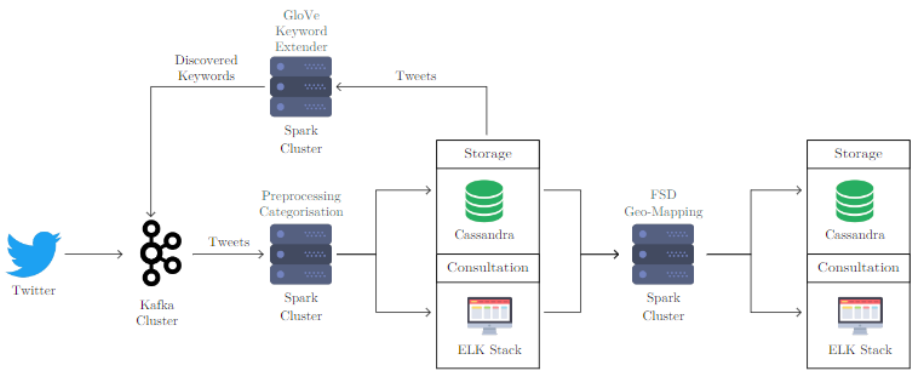

Figure 3: Technical overview of SONAR.

6. [Sec-Buzzer: cyber security emerging topic mining with open threat intelligence retrieval and timeline event annotation]

(https://link.springer.com/article/10.1007/s00500-016-2265-0)

Yayının full-text ine erişemedim. Research gate aracılığıyla erişim izni istedim. Preview den anladığım kadarıyla sosyal media streamlerini işleyerek siber güvenlikle ilgili saldırıları detect edip bu saldırılardan korunmak için otomatik olarak çözümler öneren bir system üzerine araştırma yapmışlar.

7. [Crowdsourcing Cybersecurity: Cyber Attack Detection using Social Media]

(http://people.cs.vt.edu/naren/papers/case1872-khandpurA.pdf)

Twitter verilerini işleyerek siber güvenlik saldırılarını detect etme üzerine bir çalışma daha. Daha önceki yapılan çalışmalara benzer bir çalışma olduğunu Kabul ediyorlar fakat daha başarılı sonuçlar elde ettiklerini iddia ediyorlar.

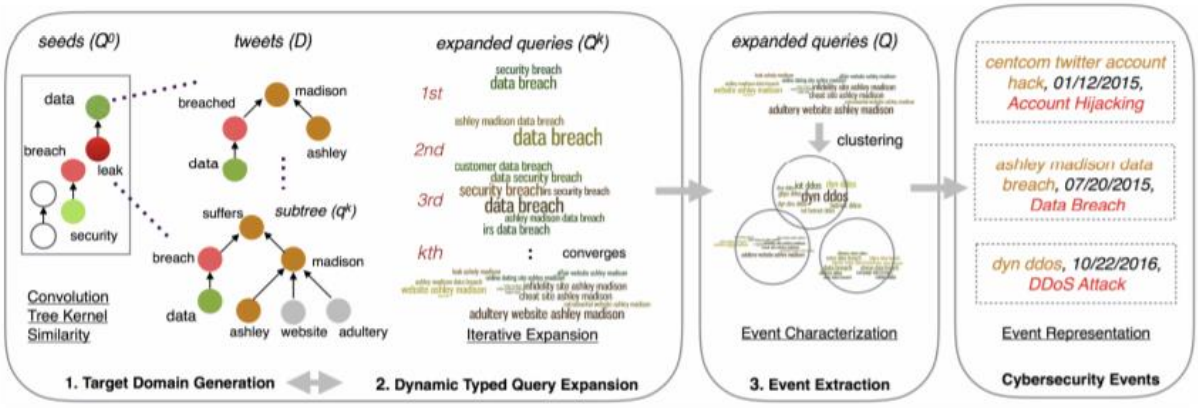

Figure 1: A schematic overview of the cybersecurity event detection system.

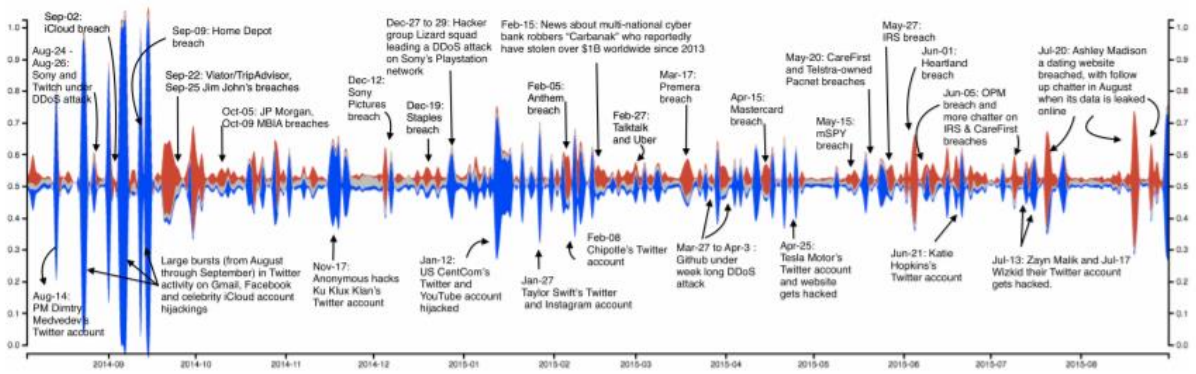

Figure 3: Streamgraph showing normalized volume of tweets (August 2014 through August 2015) tagged with data breach (red), DDoS activity (grey) and account hijacking (blue) types of cyber-security events.

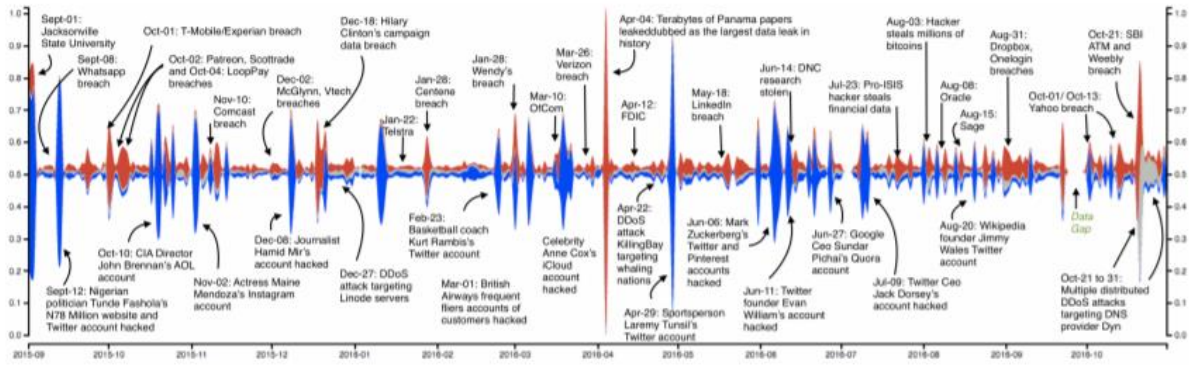

Figure 4: Streamgraph showing normalized volume of tweets (September 2015 through October 2016) tagged with data breach (red), DDoS activity (grey) and account hijacking (blue) types of cyber-security events.

8. [Identifying offenders on Twitter: A law enforcement practitioner guide](https://www.researchgate.net/publication/320005848_Identifying_offenders_on_Twitter_A_law_enforcement_practitioner_guide)

Twitter da saldırganları ve yasadışı içerik paylaşanlara karşı twitter streamlerini analiz ederek tespit edilmesine yönelik bir çalışma. Yayının full text ine erişim iznim henüz yok, researchgate üzerinden erişim izni talebinde bulundum.