

Lateral Astroturfing Attacks on Twitter Trending Topics

Tuğrulcan Elmas

LSIR & EPFL

Information War on Twitter Recap

- Anonymity is allowed
- Being public is mainstream
- Automation is easy & supported



Twitter bots are being weaponized to spread information on the French presidential campaign hack

Five percent of the accounts tweeting #MacronGate make up 40 percent of Tweets.

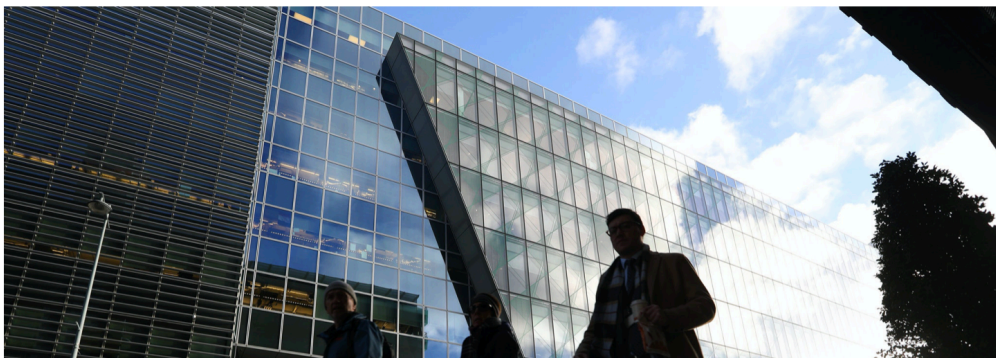
By April Glaser | @aprilaser | May 6, 2017, 2:42pm EDT

f t SHARE



The New York Times

The Fake Americans Russia Created to Influence the Election



Published on May 25th, 2016 | by EJC

Manipulation Of Public Opinion In Venezuela Using Political Bots



Politics How pro-government trolls are using a sexy Twitter bot to sway Turkey's election

By Efe Kerem Sozeri

Oct 31, 2015, 1:52pm CT | Last updated Oct 31, 2015, 1:59pm CT

Revealed: Iran's 'clumsy' troll army pushed one million tweets in effort to destabilise region

► Twitter releases extent of Tehran's Twitter activity, showing hundreds of accounts set up to spread regime's propaganda. But efforts were fairly crude and ineffective, analyst says



Bots on Twitter Recap

The image shows a screenshot of a Twitter profile for @johnabrahan89. The profile features a blue egg icon as a profile picture, no bio, and a 'Follow' button. The 'About' section shows 381 tweets, 1,347 following, 209 followers, and 0 listed. The 'Following' list shows several accounts, including @RodneyHess and @JCenter. The 'Timeline' shows four tweets from @johnabrahan89, all dated 25 Feb, with spammy content like 'Myspace friend bomber', 'The Simplest Forex Pips', 'wpCloaker', and 'Push Button Marketer'. Red callout boxes with arrows point to these features: 'No photo' points to the egg icon; 'No Bio' points to the empty bio field; 'Following a ton of people who aren't following back' points to the 'Following' count; 'Spammy posts' points to the tweets; and 'Not Listed' points to the 'Listed' count.

No photo

No Bio

Following a ton of people who aren't following back

Spammy posts

Not Listed

About @johnabrahan89

381 Tweets 1,347 Following 209 Followers 0 Listed

Connections

You both follow @RodneyHess, @JCenter, ...

Following 1,347

Timeline Favorites Following Followers Lists

johnabrahan89 john abraham
Myspace friend bomber Unlimited License
<http://howcanigetmoneyonline.blogspot.com/>
25 Feb

johnabrahan89 john abraham
The Simplest Forex Pips
<http://howcanigetmoneyonline.blogspot.com/>
25 Feb

johnabrahan89 john abraham
wpCloaker Cloak WordPress Blogs
<http://howcanigetmoneyonline.blogspot.com/>
25 Feb

johnabrahan89 john abraham
Push Button Marketer
25 Feb

Bots on Twitter Recap

The image shows a screenshot of a Twitter profile for @johnabraham89. The profile is crossed out with a large red 'X'. Several red callout boxes with arrows point to specific features:

- No photo**: Points to a placeholder image of a white egg.
- No Bio**: Points to the bio field, which is empty.
- Following a ton of people who aren't following back**: Points to the 'Following' count of 1,347.
- Spammy posts**: Points to a tweet that says 'Myspace friend bomber Unlimited License' with a link to a blogspot site.
- Not Listed**: Points to the 'Listed' count of 0.

The profile statistics are: 381 Tweets, 1,347 Following, 209 Followers, and 0 Listed. The profile name is 'john abraham' and the bio is empty. The timeline shows several tweets, all with the same link: <http://howcanigetmoneyonline.blogspot.com/>.



Detecting Social Bots is
Difficult Even For Humans!!!

What is Astroturfing?



- Grassroots movement: starts from people
 - Yellowvests
- Astroturfing: as though starts from people
 - Paid online reviewers
- On Twitter it is equal to
 - Boost content with retweets
 - Game Twitter's Trends

Astroturfing **Attacks** on Twitter's Trending Topics

Turks manipulate trending topics in spite of Twitter

ISTANBUL



A Turkish company can push any hashtag to the top of the Turkish TT list within 10 minutes and to the worldwide TT list in 15 minutes. REUTERS Photo

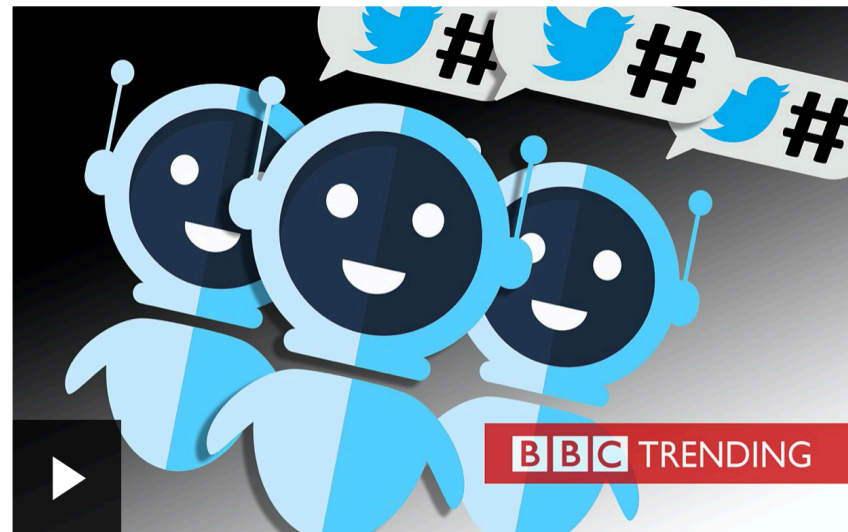
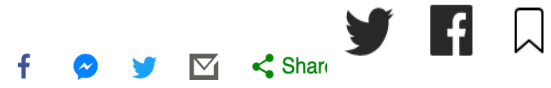
BBC Trending

How much to fake a trend on Twitter? In one country, about £150



By BBC Trending
Going in-depth on social media

2 March 2018



K-Pop Goes Info Ops

How an online army of K-pop boy band fans exposes Twitter's multiple account problem



@DFRLab

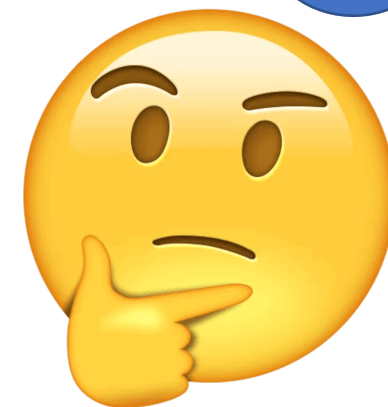
Jun 7, 2018 · 6 min read



Lateral Astroturfing Attack



Been trending for hours...



Lateral Astroturfing Attack

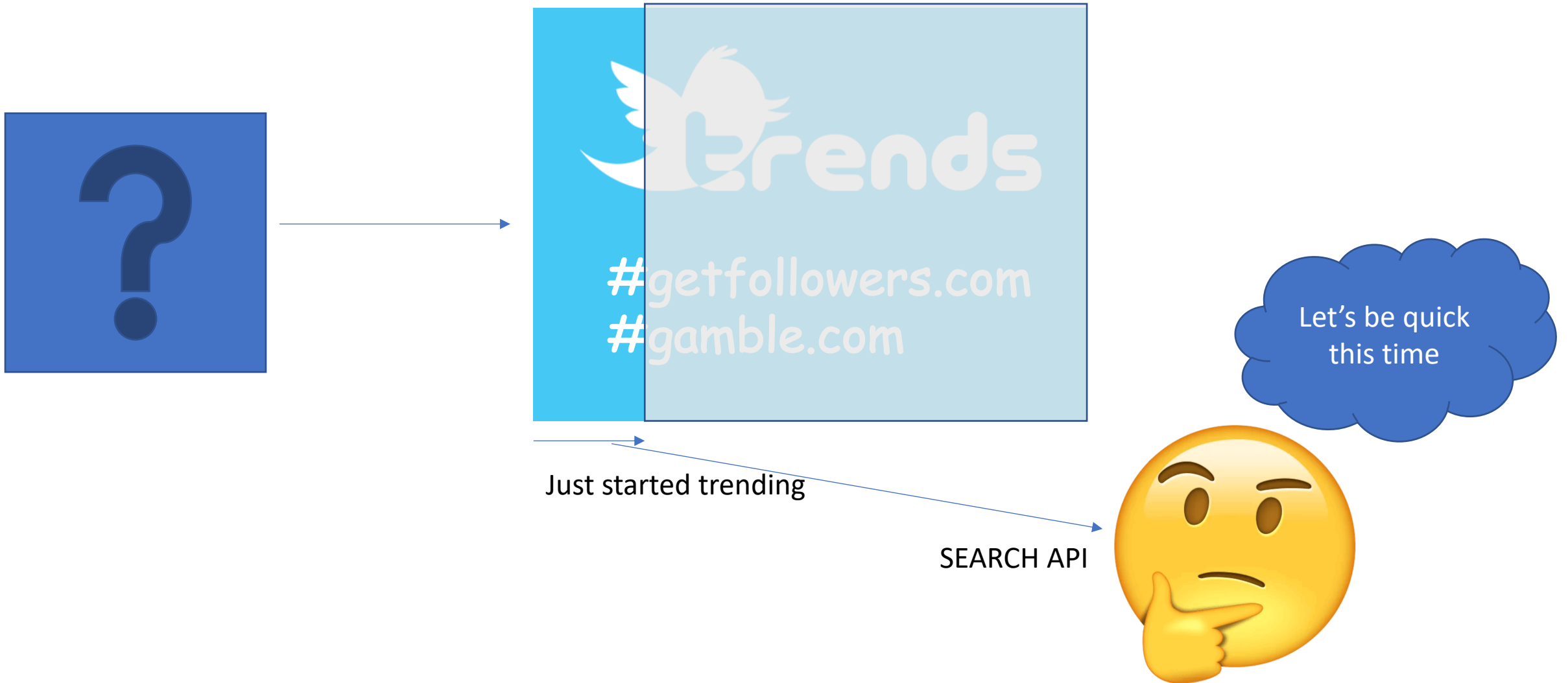


Been trending for hours...

SEARCH API



Lateral Astroturfing Attack

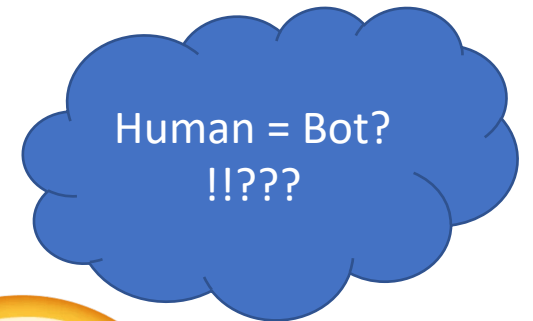
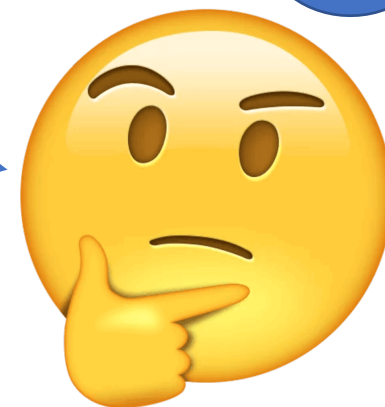


Lateral Astroturfing Attack



Bot tweets with random words, i.e.:
duck to cycle New York (city) praying mantis #getfollowerscom

SEARCH API



Lateral Astroturfing Attack

We regret
posting
these random
tweets

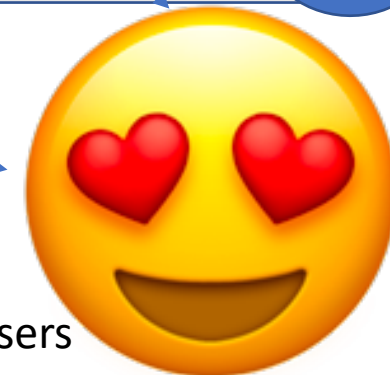


TWITTER NOTICE:
These users decided to delete these tweets
Please remove them from your system

Oh here's my
new paper

Bot tweets with random words, i.e.:
duck to cycle New York (city) praying mantis #getfollowerscom

Get realtime
tweets by these users

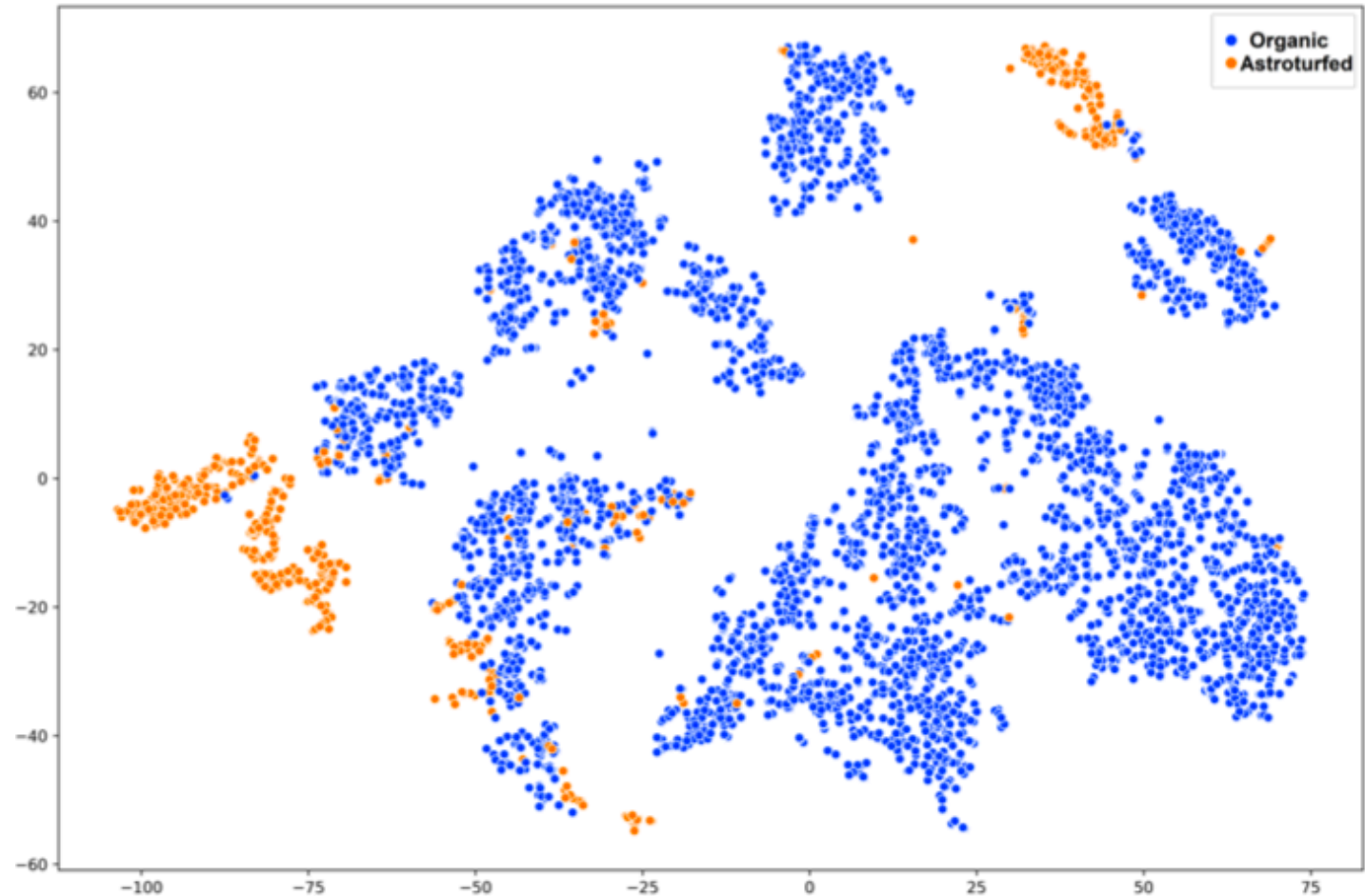


Investigation

- Went through 1% of all tweets (Freely available in archive.org)
- Only Turkish case found
 - Saudi Arabia and Korea has astroturfed trends too, but not many deletions
- Rule based classification
 - Classify tweet with random words based on rules
 - Many tweet with random words
 - Have the same keyword
 - Created in small time frame (3 minutes)
 - Deleted in small time frame (3 minutes)
 - => Astroturfing
- Problem: Adversaries can change the content and evade rules.

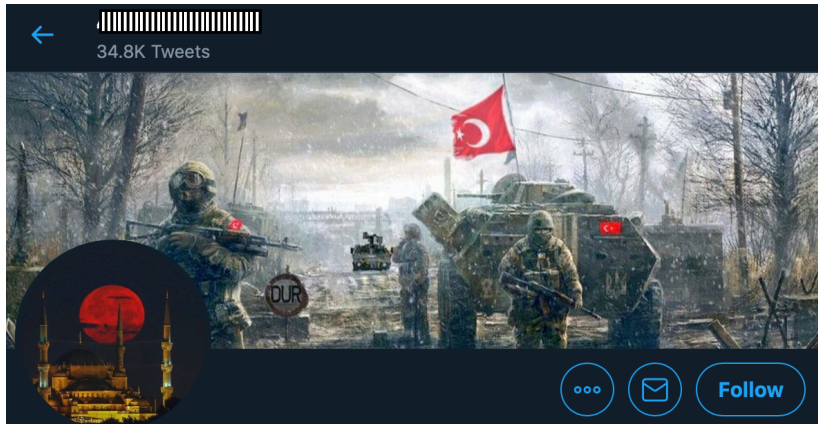
ML Based Content Free Classification

- 2 Months of trend & tweets
 - 2000 astroturfed
 - 600 astroturfed trends
 - 1% of all Twitter tweets
- ML based classification
 - Deletion count
 - Geotagged tweet count
 - Unique geolocations
 - Unique screen name
 - Retweet count
 - Url count
 - Mention count
 - Hashtag in a tweet count



TSNE plot of trends
Neural network, Prec: 95%, Rec: 91.7%

Meet The Astrobots



34.8K Tweets

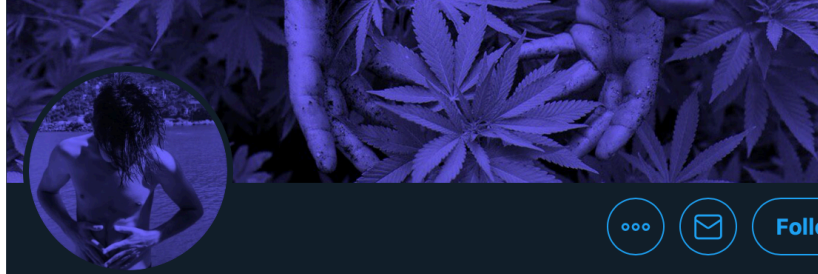
kırlırdı bir gün bütün dişliler döner şanlı şanlı çarkımız bizim. (KIZIL ELMA). 08 ekim.

Translate bio

İstanbul, Türkiye

Joined December 2017

0 Following 787 Followers



DUMLUPINAR ÜN.V./4 Yıllık bir bölüm atar mutluluklar dileriz .

Translate bio

Ordu

instagram.com/

Joined May 2014

36 Following 583 Followers

Not followed by anyone you're following



14 Tweets

Joined September 2010

105 Following 12 Followers

Not followed by anyone you're following



Doğu

Joined March 2016

0 Following 0 Followers

Joined January 2015

82 Following 13 Followers

Not followed by anyone you're following

3

Tweets Tweets & replies Media Likes

ilmiklenme Gorabette TRninYaninda sağır pencere Rh kuru kuyu yaklaşılabilir

horozlanış #BoksGecesiSekabet kazaz sahan

İşte tüm zamanların en kötü tasarıma sahip telefonlar ntv.com.tr/galeri/teknolo..

İ promise you #AlwaysWithChen

8,156 Tweets



#AlwaysWithChen

SeKai's Mom

Fan account

1

2

1

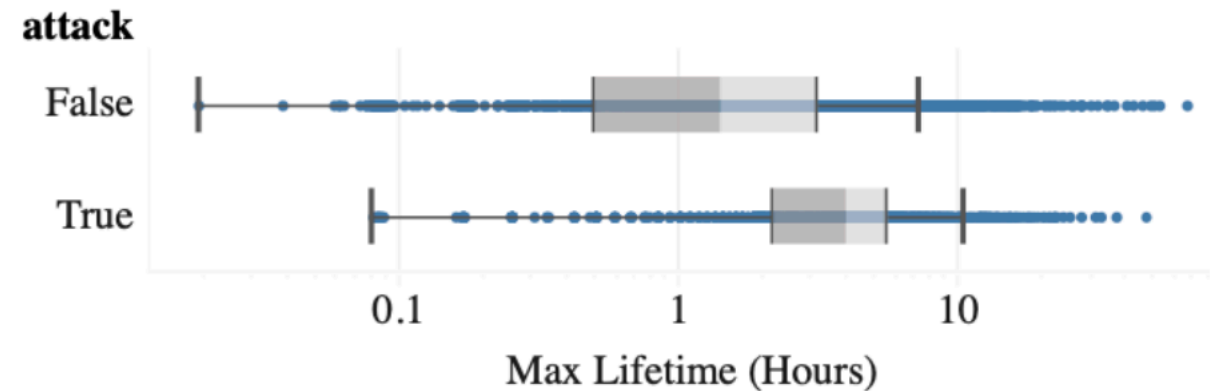
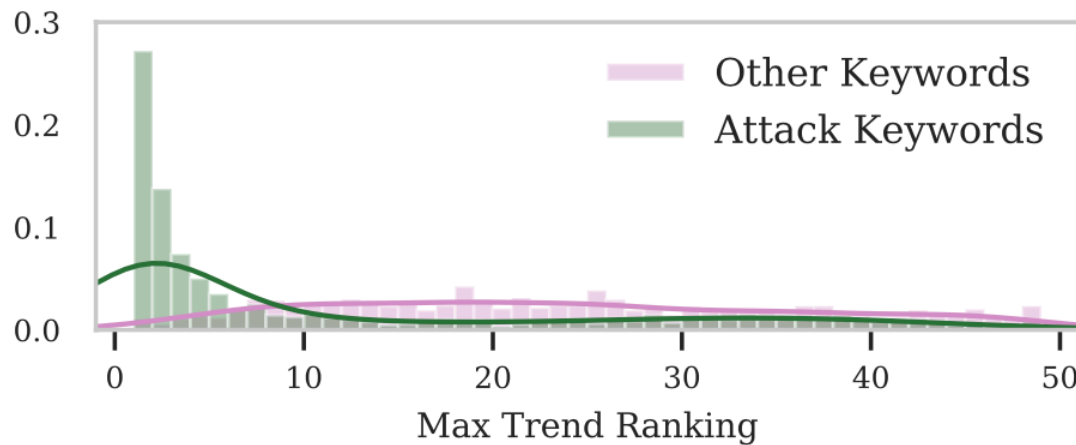
4

5

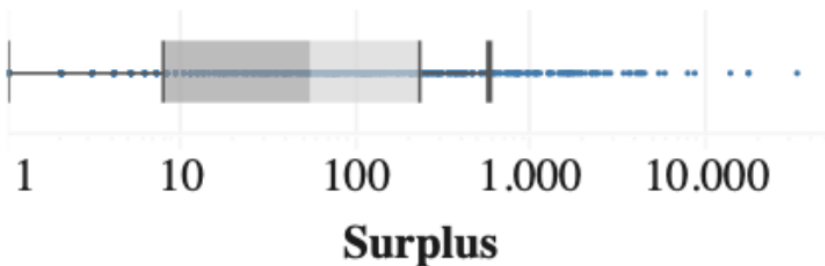
6

Quantifying Success of the Attacks

- Top 10 Guarante & High Duration



- Effect on public: Number of undeleted tweets (Surplus)



What are the Trends About?

- Ads: «Gambling, buy fake followers, tv series»
- Appeals to Government: »We want jobs»
- Election Campaign: «Vote for x»
- Fun: «Galatasaray is the best»
- Disinformation: «Erdoğan is the best.»

Conclusion

- We reported our analysis to Twitter



bugtriage-alex posted a comment.

Jul 30th (6 months ago)

Thank you for your report @tugrulz,

We are evaluating the behavior you described and will get back to you when we have more information.

Thank you for thinking of Twitter security.



acamacho changed the status to **Triaged**.

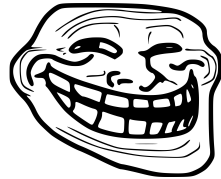
Jul 31st (6 months ago)

Thank you for your report. We believe it may be a valid security issue and will investigate it further. It could take some time to find and update the root cause for an issue, so we thank you for your patience.

Thank you for helping keep Twitter secure!

Conclusion

- They do not care



[bugtriage-alex](#) closed the report and changed the status to **Informative**.

Aug 21st (5 months ago)

Thank you for your report [@tugrulz](#),

We appreciate your suggestion and are always looking to improve our spam detection capabilities. However, after a discussion with our team, we have decided that this is a Defense-In-Depth measure rather than a specific vulnerability in our software. We are reasonably satisfied with our current spam detection security controls, and we have chosen to not to make any specific changes to our controls based on this report. Although we will be closing this report, we still want to thank you for all your effort in bringing this behavior to our attention.

Thank you for thinking of Twitter security

The attacks have started in 2015 and are still ongoing....

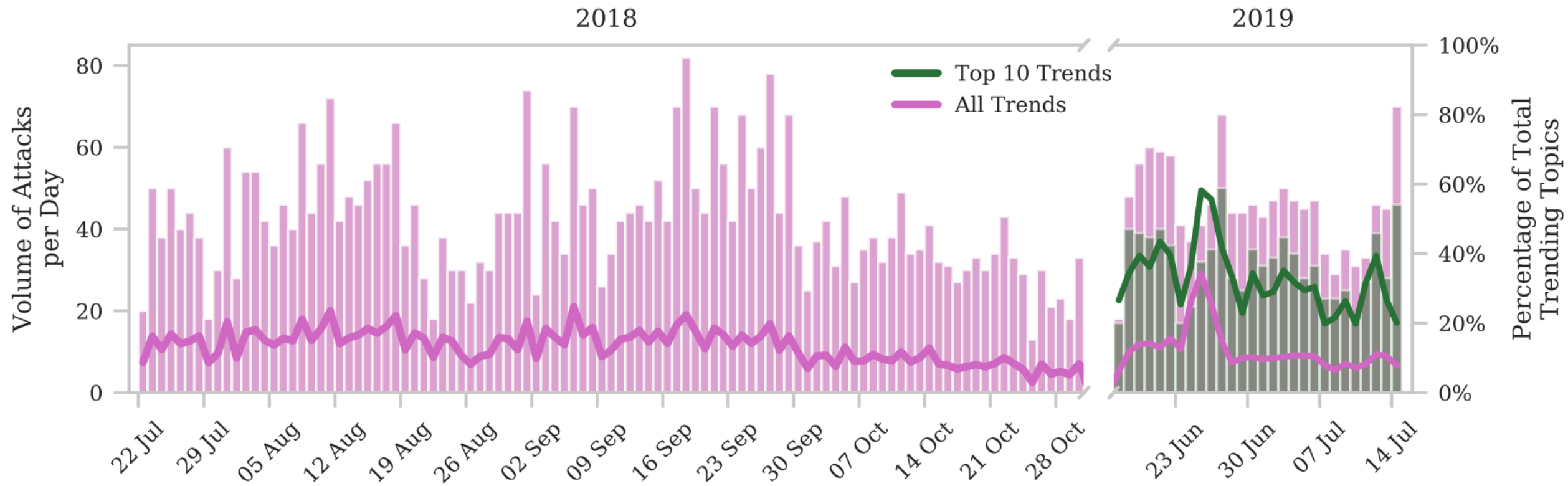
Thanks

- Questions?

Live Demo

- <https://trends24.in/turkey/>

How Many Fake Trends Per Day?



The Botnet from October 2018

