Telefónica

TREND REPORT_

Twitter botnets detection in sports events

In collaboration with:

etisalat | Singtel

CYBER SECURED BY ElevenPaths

elevenpaths.com

Table of contents

1. Executive summary

This report is an analysis of **social botnets**. We have analyzed two different Twitter conversations and identified the behavior of several groups of profiles with automatic and non-standard activity. The first conversation focuses on the **2018 FIFA World Cup**, held between the 14th of June and the 15th of July. We worked on the communities of the entire conversation and discovered several groups of profiles behaving oddly. After applying proprietary algorithms, we confirmed that those communities were **content diffusing botnets** that were activated during the football games in order to share **illegal streaming URL's**.

In order to make a wider analysis, we investigated then a larger scope of sports, as a second conversation, in the entire world for five weeks; from the 1st of November until the 10th of December 2018. This new conversation covered the **sports streaming theme in Twitter** and included important events like the Premier League, UEFA Champions League or the weekly NFL games in the US. Not only we discovered that some of the botnets identified in the FIFA World Cup were still active and diffusing current content, but **new botnets** were found focused in different sports.

Our goals for this report were:

- Find botnets associated to the illegal broadcasting of sport events.
- Observe their evolution over time.
- Identify new botnets from the previously found ones.
- Understand the behavior of Twitter botnets.
- Prevent and control de diffusion of content via Twitter botnets.

As a result, we have learnt how to identify this kind of networks, and we have started to apply these insights in several uses cases, such as in **reputational crisis**.

When comparing the analyzed conversations, we did observe that **25 bot profiles** coincide and have been active and sharing streaming content since June 2018.

Some of the conclusions were:

- Twitter botnets are **not usually focused** on a particular sport; they consider any major relevant sport event to promote their illegal broadcasting.
- Tweets related to illegal American Football broadcasting represents up to **0.5%** of the tweets related to this sport.
- Some countries have only a **few servers** broadcasting illegal sport links.
- The **three top countries** broadcasting illegal sport content are United States, Spain and Germany.

2. Context

It is well known how the industry of Broadcasting & Media Rights in sport[1] is concerned about illegal broadcasting, affecting directly their expected revenue. The most known ways are via **foreign satellite systems** (someone trying to sell a foreign satellite system) or through **unauthorized websites**.

As an example, the Premier League in the UK has a website to report such cases: https://www.idinquiries.com/premier-league. In Spain, there is an organization called IPRORED that provides protection of rights on the Internet (http://www.iproded.com/english/index.html). Furthermore, LaLiga has developed its own technology to overcome illegal broadcasting.[2]

First analyzed conversation: During the last 2018 FIFA World Football Cup celebrated in Russia, we did observe how different Twitter botnets were illegally broadcasting the 64 football games thanks to the **tweets sharing the streaming URLs**.

Second analyzed conversation: We have now extended this observation to more sport events and covering more countries. This time we have focused on other main sports such as rugby, cricket, baseball, basketball and of course football.

2.1. What is a Twitter botnet?

We all know that a **botnet** is a number of Internet-connected devices, each of which is running one or more bots. Botnets can be used to perform DDoS attacks, steal data, send spam, and allow the attacker to access the device and its connection. The owner can control the botnet using command and control (C&C) software.

Many Twitter users are bots, which are accounts controlled and sometimes created by computers. **Twitter bots** can send spam tweets, manipulate public opinion, be used for online fraud, and in our case, for **promoting URL access to illegal sport content**. The botnet contains a **single type of bot**, showing exactly the same properties throughout the botnet. Bots are sold for money as fake followers.

There are Twitter botnets for everything, such as pushing anti-political behaviors, hijacking of Cryptocurrency-Related Accounts[3], etc. One of the biggest Twitter botnets had 350.000 accounts[4].

[1] https://www.wipo.int/ip-sport/en/broadcasting.html

[2] https://elpais.com/deportes/2018/10/30/es_laliga/1540911095_179420.html

[3] https://nulltx.com/twitter-botnet-is-responsible-for-the-hijacking-of-cryptocurrency-related-accounts/

[4] https://www.newscientist.com/article/2117811-army-of-350000-star-wars-bots-found-lurking-on-twitter/

3. Methodology: Aldara

Aldara is an own developed tool to perform Social Network Analysis. It applies different proprietary algorithms and Machine Learning in order to provide the insights the analysts require. Aldara works in several phases:

1. **Aggregating Data Sources**: Directly connected to Twitter and integrated with other partners to provide access to different data sources.

2. **Applying AI and Machine Learning**: Using Machine Learning and Natural Language Processing for processing both real time and batch information. Furthermore, Aldara uses proprietary techniques to better understand how people interact.

3. **Empower the enterprise**: It provides dashboards and integrations with third-party enterprise tools to help analysts to better understand what is actually happening in the Social Networks.

4. **Providing Actionable Intelligence**: the brain of Aldara has the capability to provide relevant insights from the beginning, providing insights that help to contextualize social profiles and conversations.

As a result, Aldara is used by analysts to address different challenges, such as:

- **Profiles contextualization**: contextualizing anonymous profiles and allowing to identify the groups they belong to and even who they are in real life.
- **Deep conversation analysis**: understanding how people interact and isolating the core of the actual conversation from the resilient buzz.
- **Influencers identification and analysis**: discovering new influencers by topic and evaluating the previously identified ones.
- **Market research**: analyzing what competitors are doing, which are the most relevant markets and who are the most important players among them.
- **Risk evaluation**: identifying the groups and actors behind a campaign and analyzing their past behavior in order to evaluate the actual risk of a potential attack.
- **Bot identification**: detecting bot-led disinformation campaigns against your brand.

4. 1st Analysed Conversation: Russian World Cup 2018

During the 2018 FIFA World Cup (14th of June – 15th of July) we monitored the online conversation in Twitter **to identify illegal content diffusers**. In order to do so, we combined several components in the query using Aldara to get the clearest results to form an insightful graph. The query was built in English and contained a combination of the World Cup and streaming related terms.

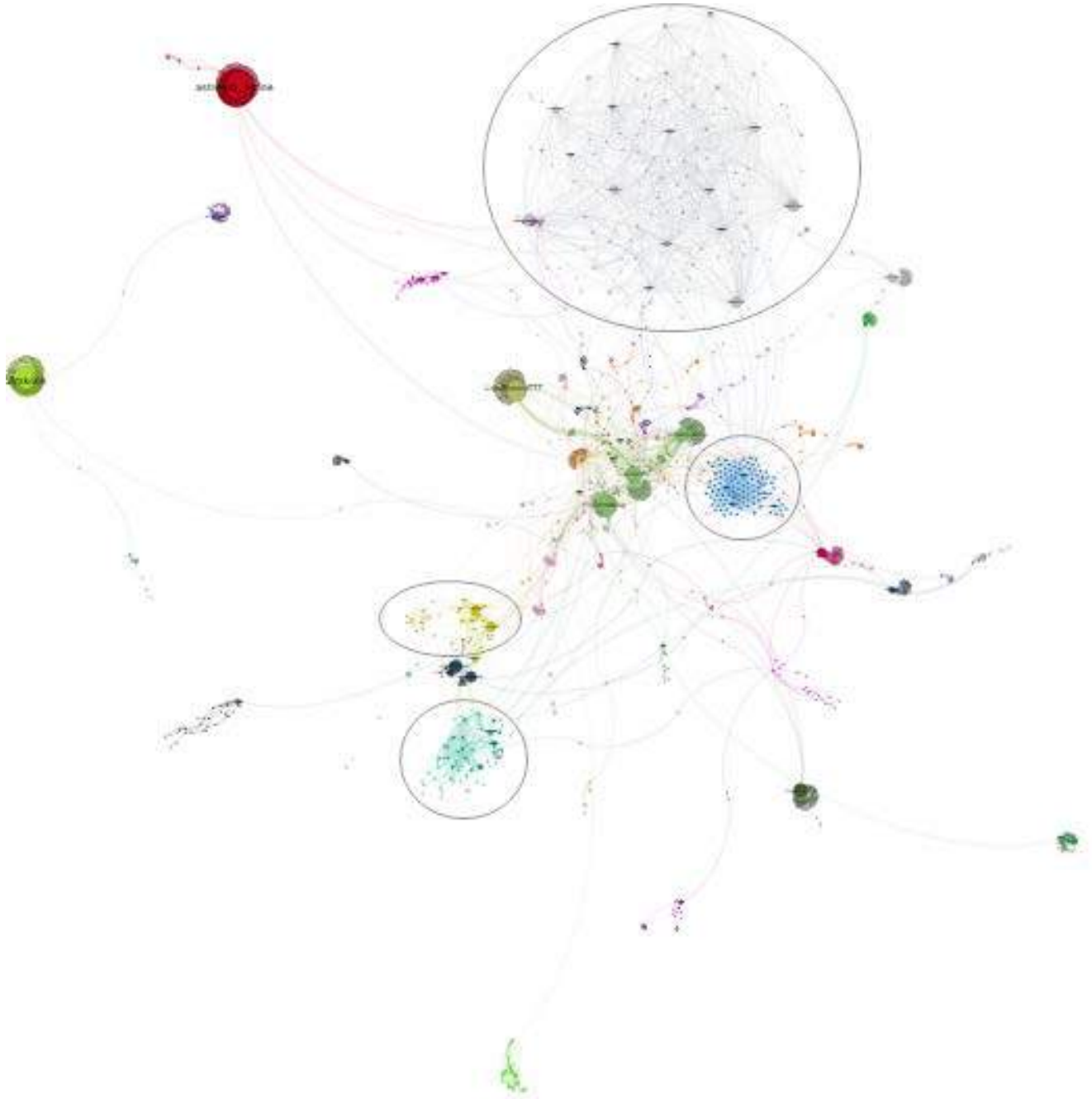
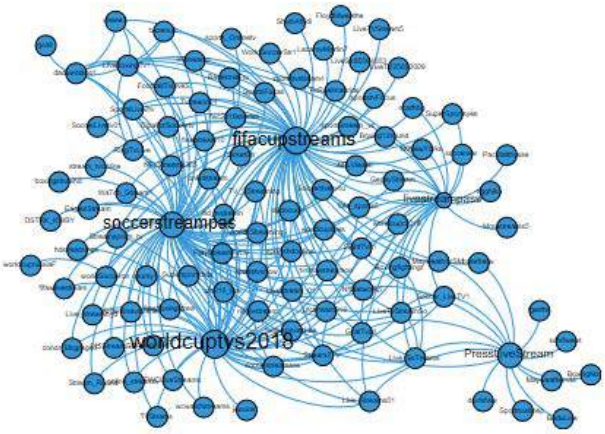
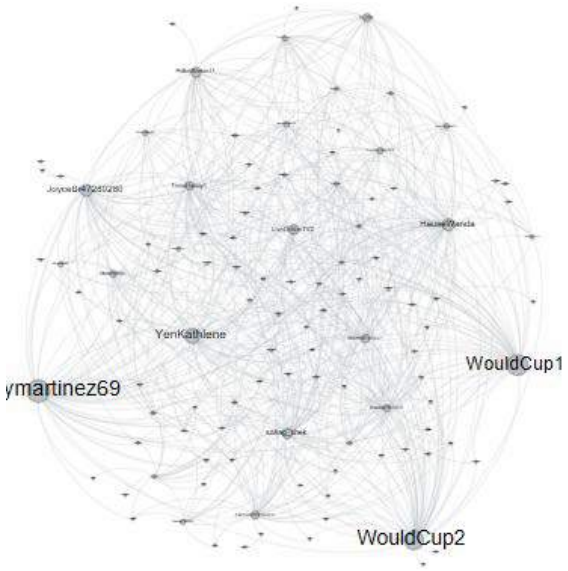


Figure 1: Twitter conversation about the streaming of football games

In the graph we can distinguish up to 4 Twitter botnets:

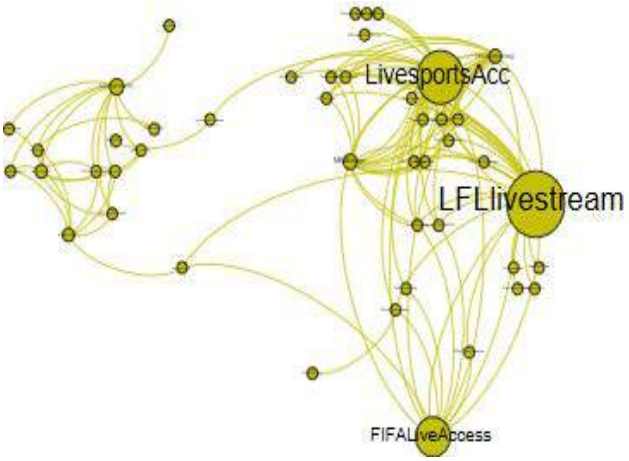


Botnet 1: It covers European football. The botnet was formed by 106 profiles, from which 102 are still active



Botnet 2: It covers European leagues mostly. From 111 initial profiles, two have been deleted or suspended

Botnet 3: It covers the NFL, AFL, UK (Premier) and boxing streaming. From 67 initial profiles, 59 are still active



Botnet 4: It covers boxing, NFL and MLB streaming. It was formed by 46 profiles and 44 are still active

Figure 2: Twitter botnets

The way to identify these botnets is by **analyzing the communicative behavior of the profiles**. As it can be seen in the examples, the accounts in each botnet are all interconnected, whereas in a normal conversation, there are some dominant profiles with a lot of diffuser or engaging accounts around, as the figure below.

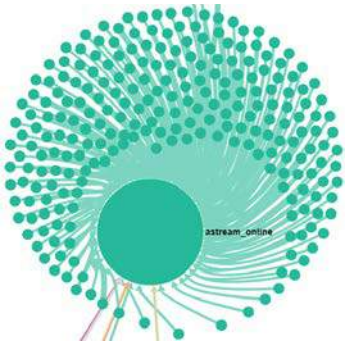


Figure 3: Twitter normal conversation

From the analysis of this graph, we have been able to identify the sports scope of these botnets (apart from the Russian World Cup) and we have checked if they were still active or they have been deleted or suspended in Twitter.

From this analysis we decided to go further in the analysis of content diffusing botnets and we developed a methodology to identify new botnets around the world related to a wider spectrum of sports.

5. 2nd Analyzed Conversation: Sport events worldwide

We decided to extend the FIFA Football World Cup 2018 conversation to more sport events and covering more countries. This time we focused the analysis on other main sports such as rugby, cricket, American football, baseball, basketball, and of course football.

For **5 weeks**, we have monitored in Aldara all the contents related to sport events streaming.

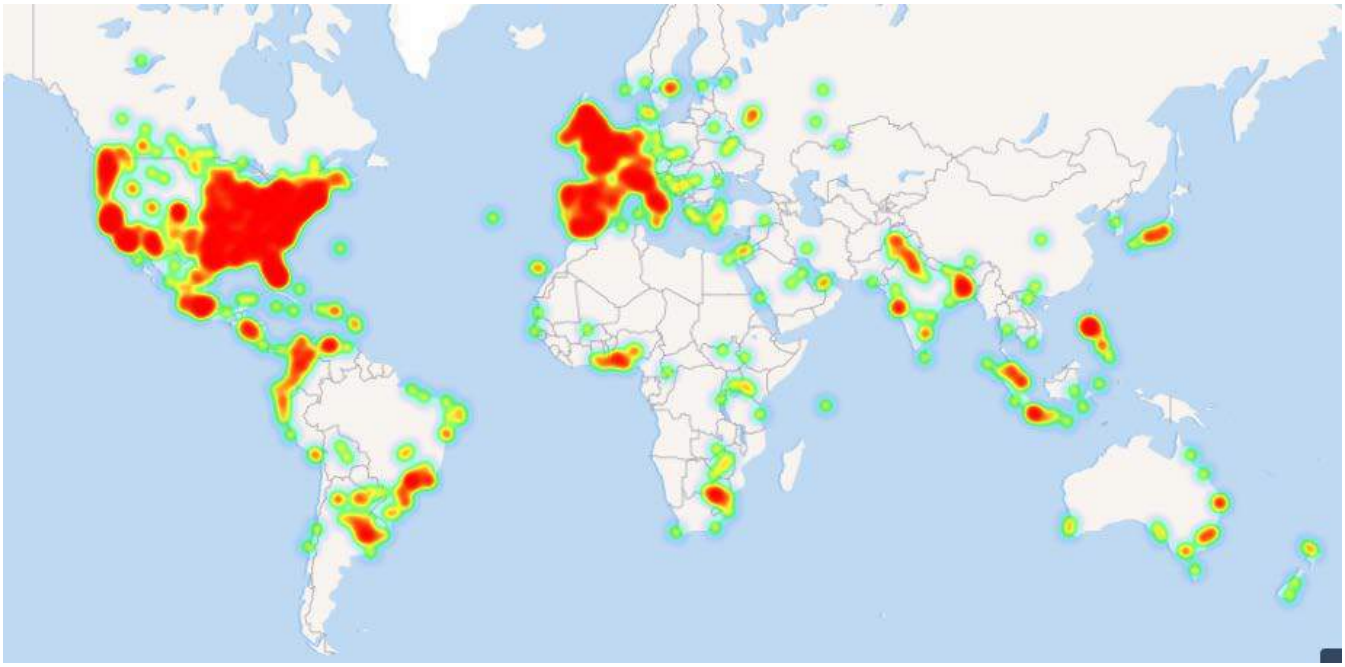


Figure 4: Tweets heat map related to the 5 weeks sport events monitoring

In order to have a clear view of the sport's streaming diffusion in Twitter, we created a query formed by the **major sports and leagues of the world** and streaming related terminology. As stated above, this analysis covers 5 weeks; from the 1st of November until the 10th of December of 2018.

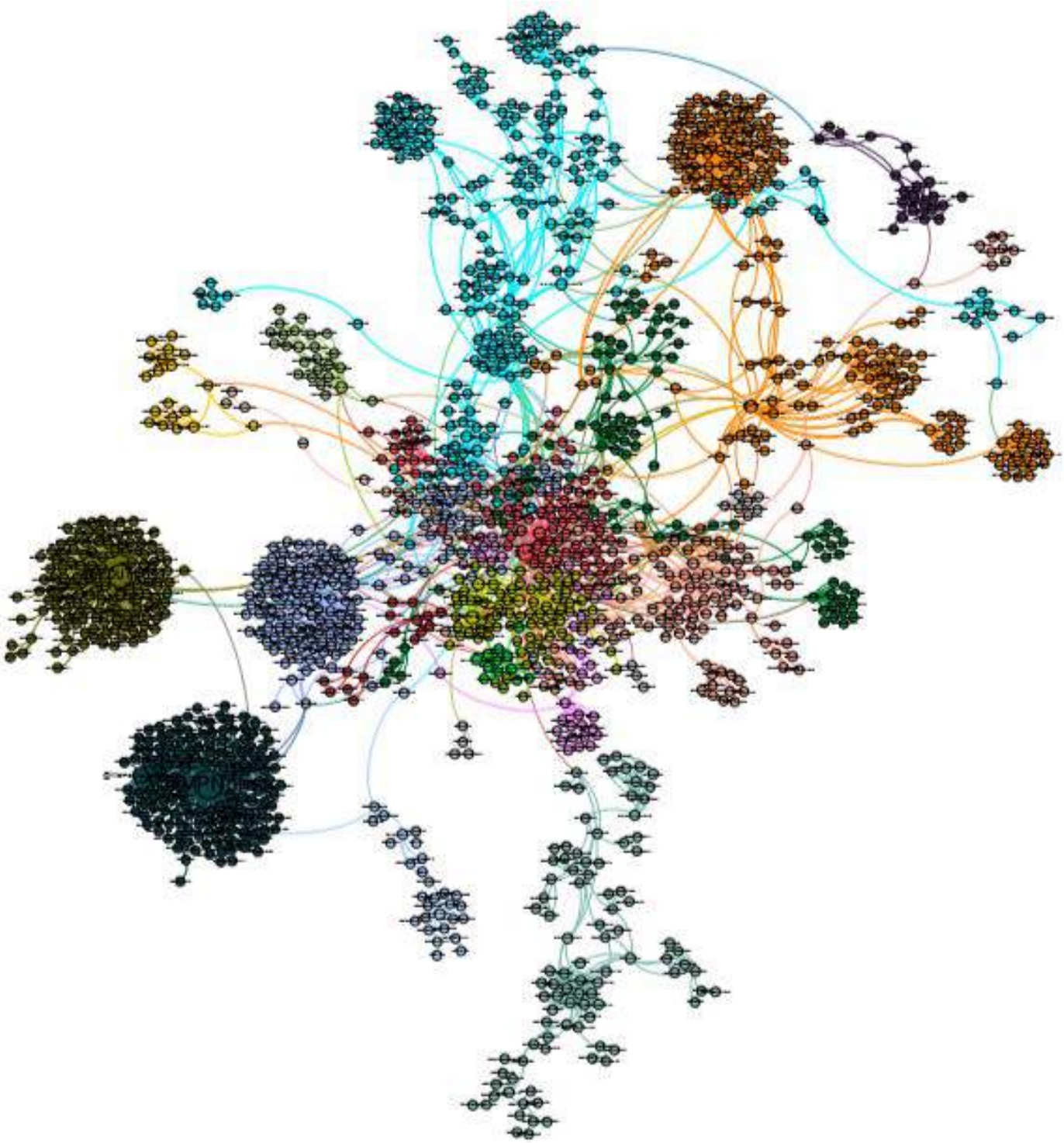
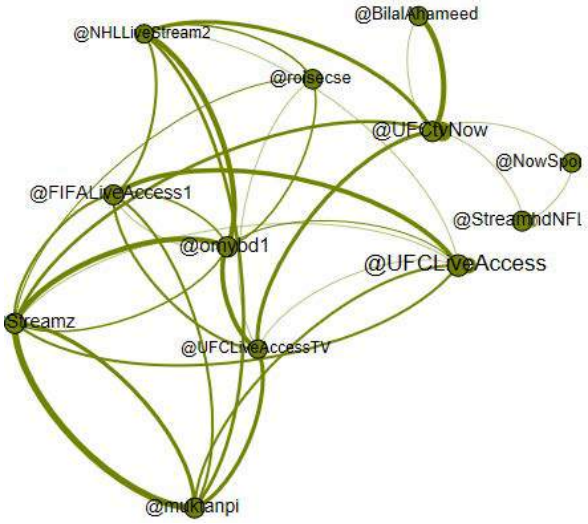
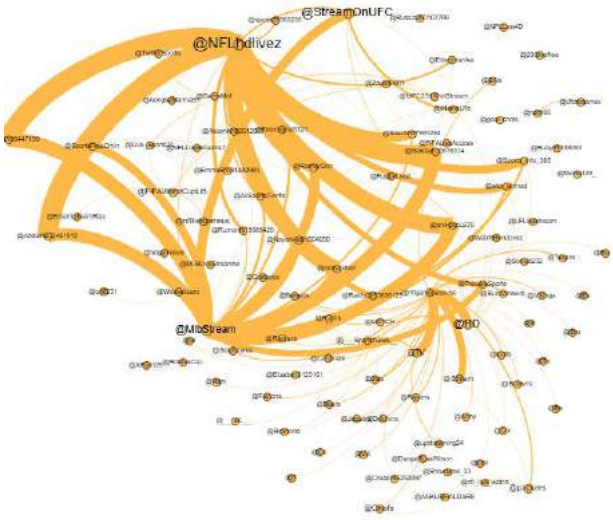


Figure 5: Twitter conversation about the streaming of several sports.

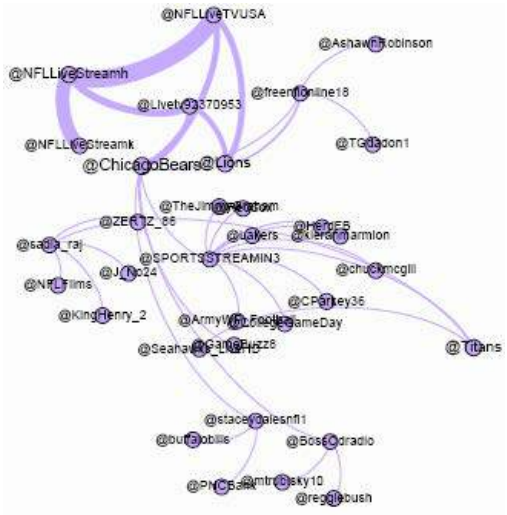
In this conversation, we have identified 5 content diffusing botnets:



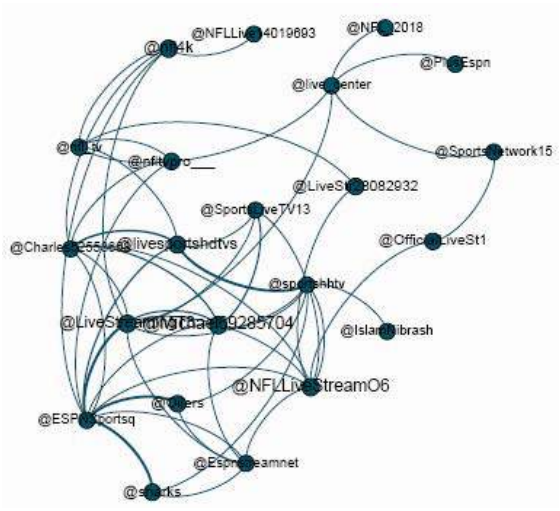
Botnet 1: It covers UFC, hockey and baseball streaming. The botnet was formed by 12 profiles and all of them are still active



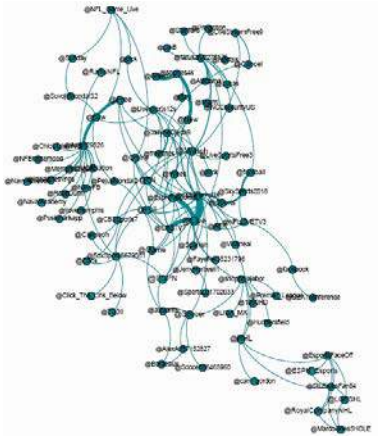
Botnet 2: It covers European leagues mostly. From 96 initial profiles, 4 have been deleted or suspended



Botnet 3: It covers the NFL and AFL streaming. From 22 initial profiles, 21 are still active



Botnet 4: It covers boxing, NFL and MLB streaming. It was formed by 46 profiles and 45 are still active



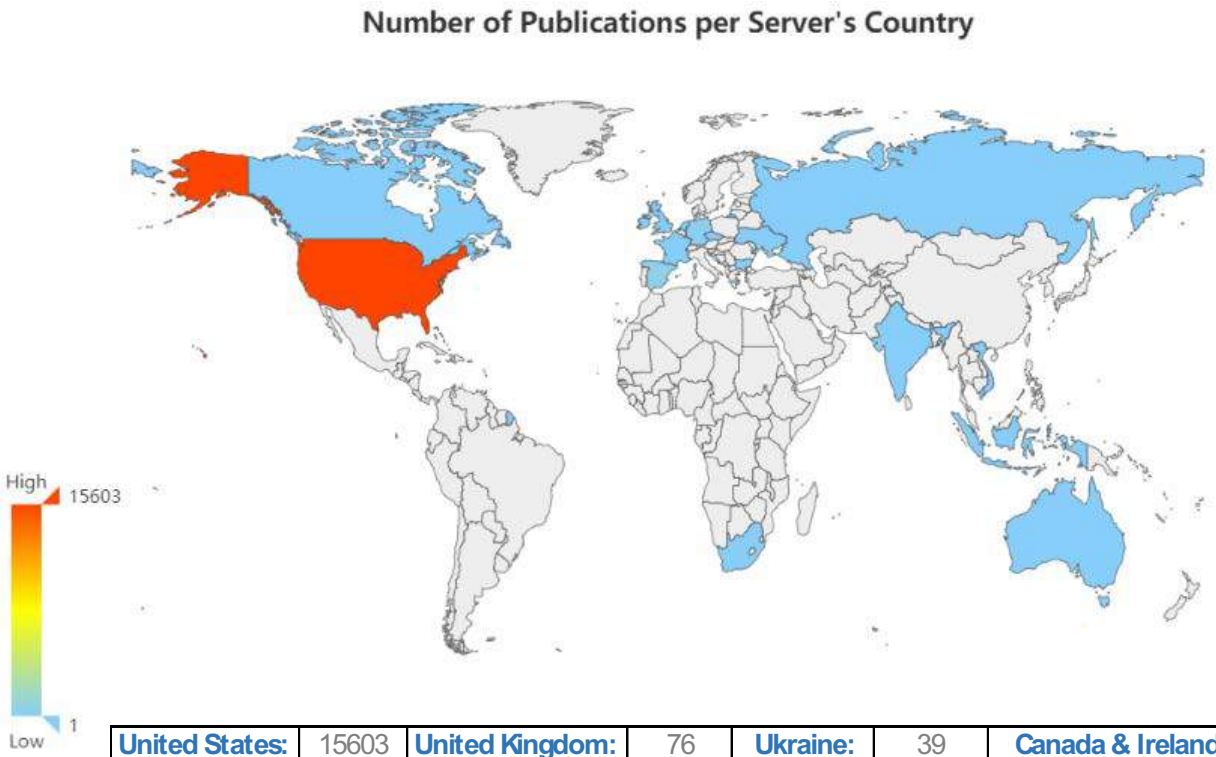
Botnet 5: It covers American football and European leagues. It was formed by 187 profiles and 183 are still active

Figure 6: Twitter botnets

When comparing the analyzed World Cup's profiles and the recent analysis, we can see that **25 bot profiles** coincide and have been active and sharing streaming content since June 2018. Specifically, the botnet number 2 from the latest analysis' has most of FIFA World Cup's botnet number 4 profiles.

In addition to the analysis of the botnets of the conversation, we have analyzed the **IP addresses of the streaming websites** distributed in the tweets.

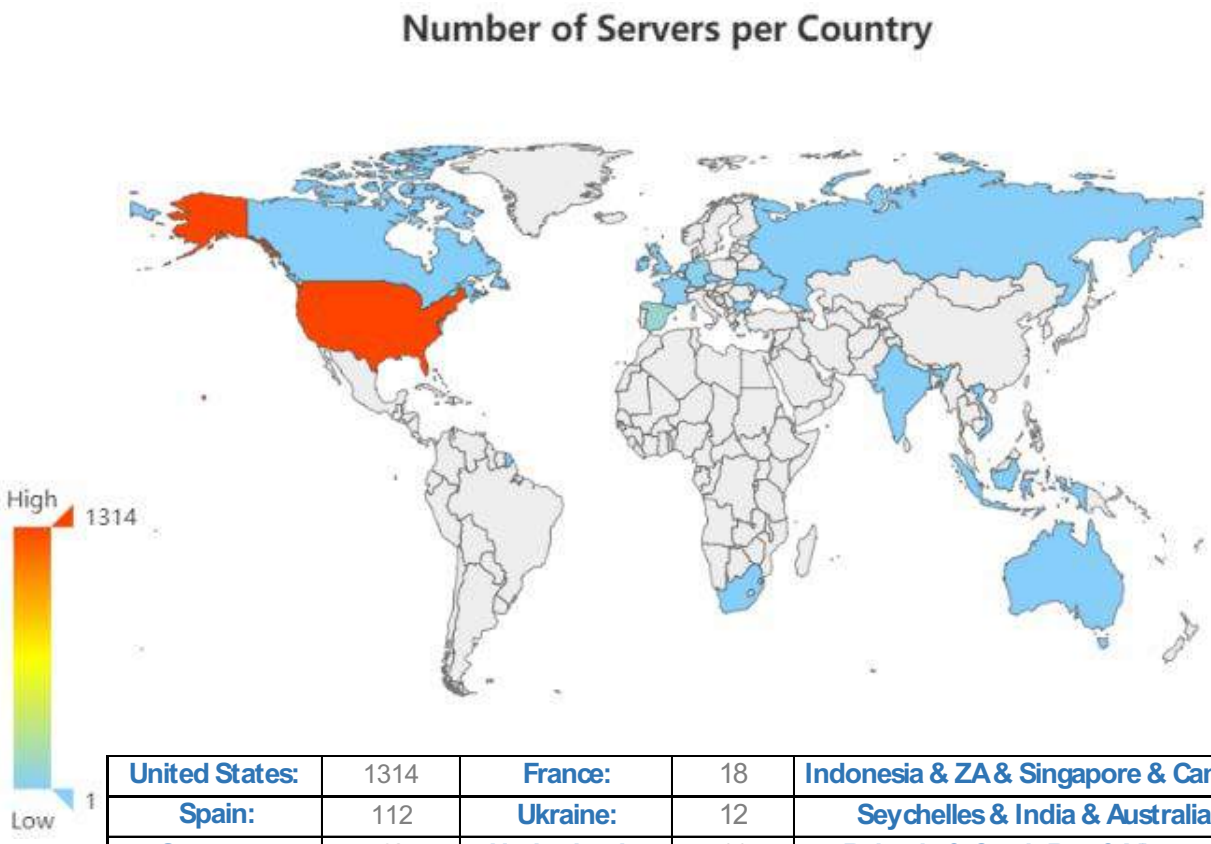
First, we have analyzed the number of publications (links) per server's country;



United States:	15603	**United Kingdom:**	76	**Ukraine:**	39	**Canada & Ireland:**	4
Spain:	525	**South Africa:**	63	**Russia:**	15	**Vietnam:**	3
Germany:	359	**France:**	57	**Indonesia:**	7	**Australia:**	2
Seychelles:	80	**Netherlands:**	40	**Singapore:**	5	**Bulgaria & Czech Rep**	1

Figure 7: Number of publications per server's country

In addition, we have also identified the number of Servers per country;



United States:	1314	France:	18	Indonesia & ZA & Singapore & Canada:	3
Spain:	112	Ukraine:	12	Seychelles & India & Australia:	2
Germany:	40	Netherlands:	11	Bulgaria & Czech Rep & Vietnam:	1
United Kingdom:	34	Russia & Ireland:	4		

Figure 8: Number of Servers per country

When you compare the last two graphs, we can observe statistics such as how 2 servers in Seychelles are illegally broadcasting up to 80 different links or how in the Netherlands, each server broadcasts on average 4 links of illegal content.

5.1. Examples of Tweets

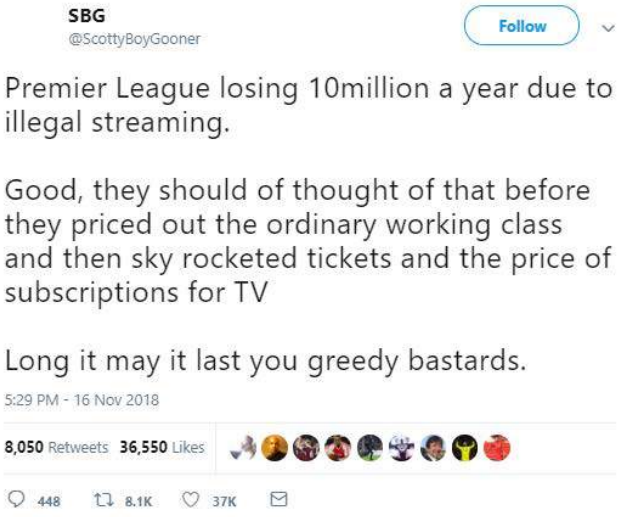


Figure 9: Tweet complaining about high prices related to legal broadcasting

Figure 10: Tweet with a link to illegal Copa Libertadores broadcasting

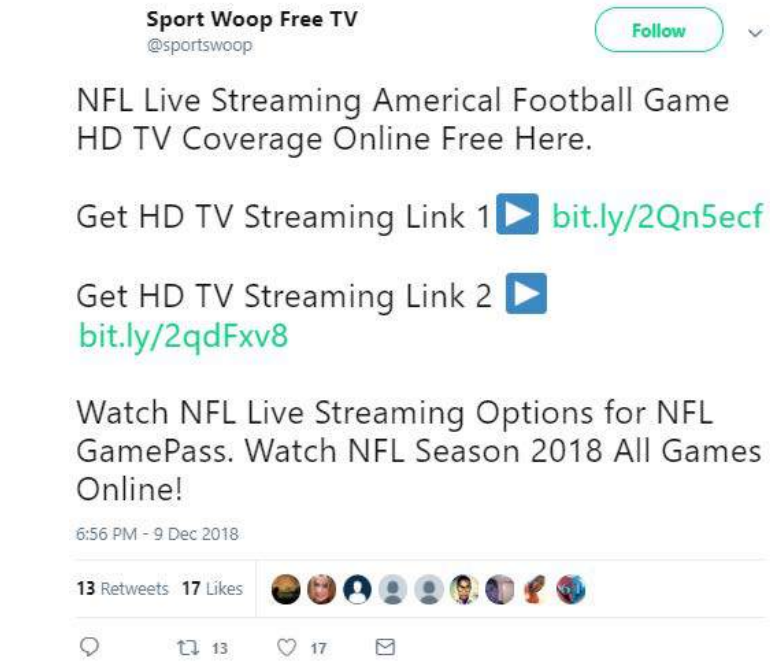


Figure 11: Tweet promoting NFL illegal broadcasting

Figure 12: Tweet promoting several illegal sports broadcasting

6. Conclusions

- With our tool Aldara we were able to identify **Twitter botnets** based on the "particular shape" of the composition and based of several algorithms.

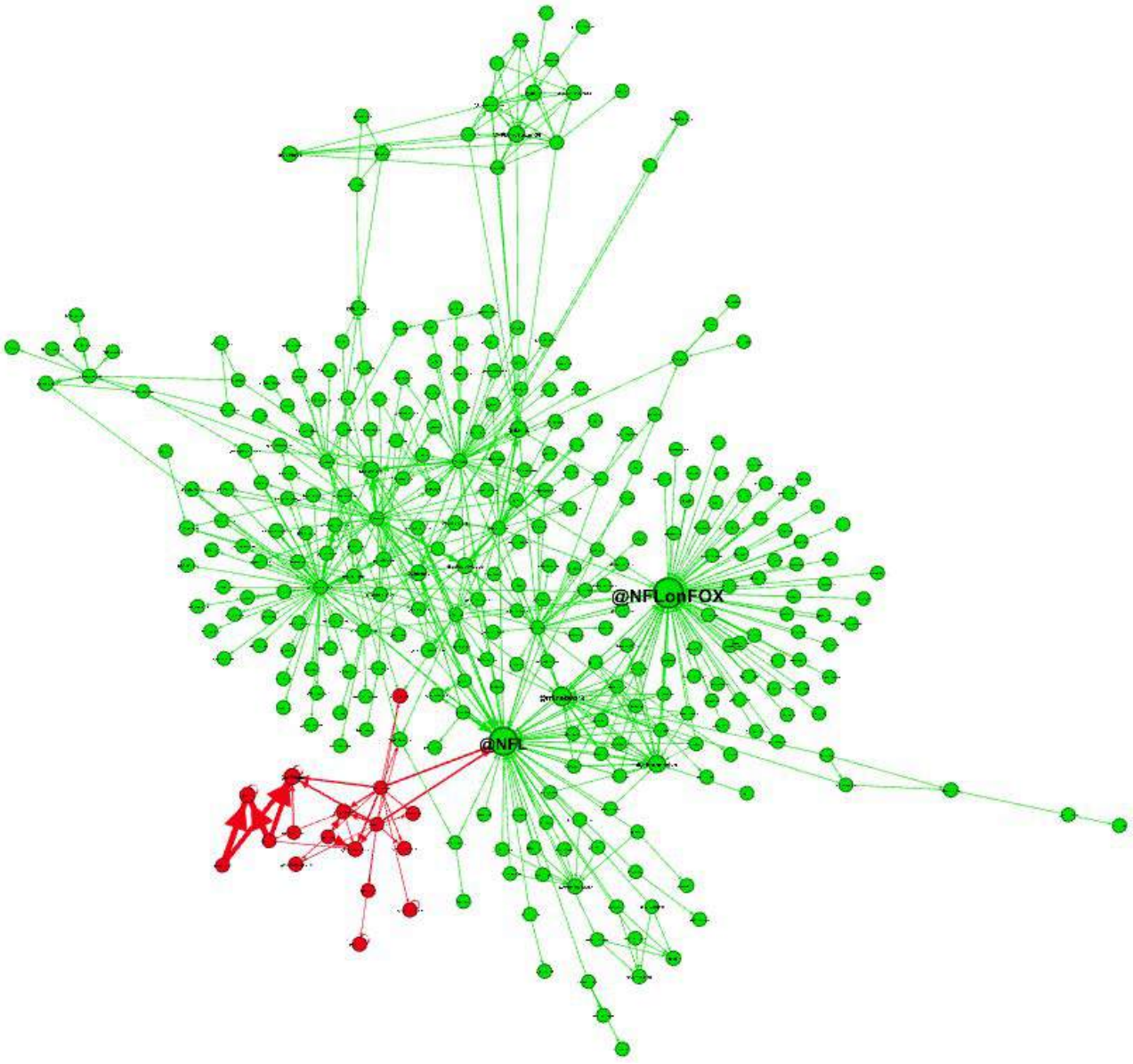


Figure 13: Twitter graphs (Red colour is a twitter botnet)

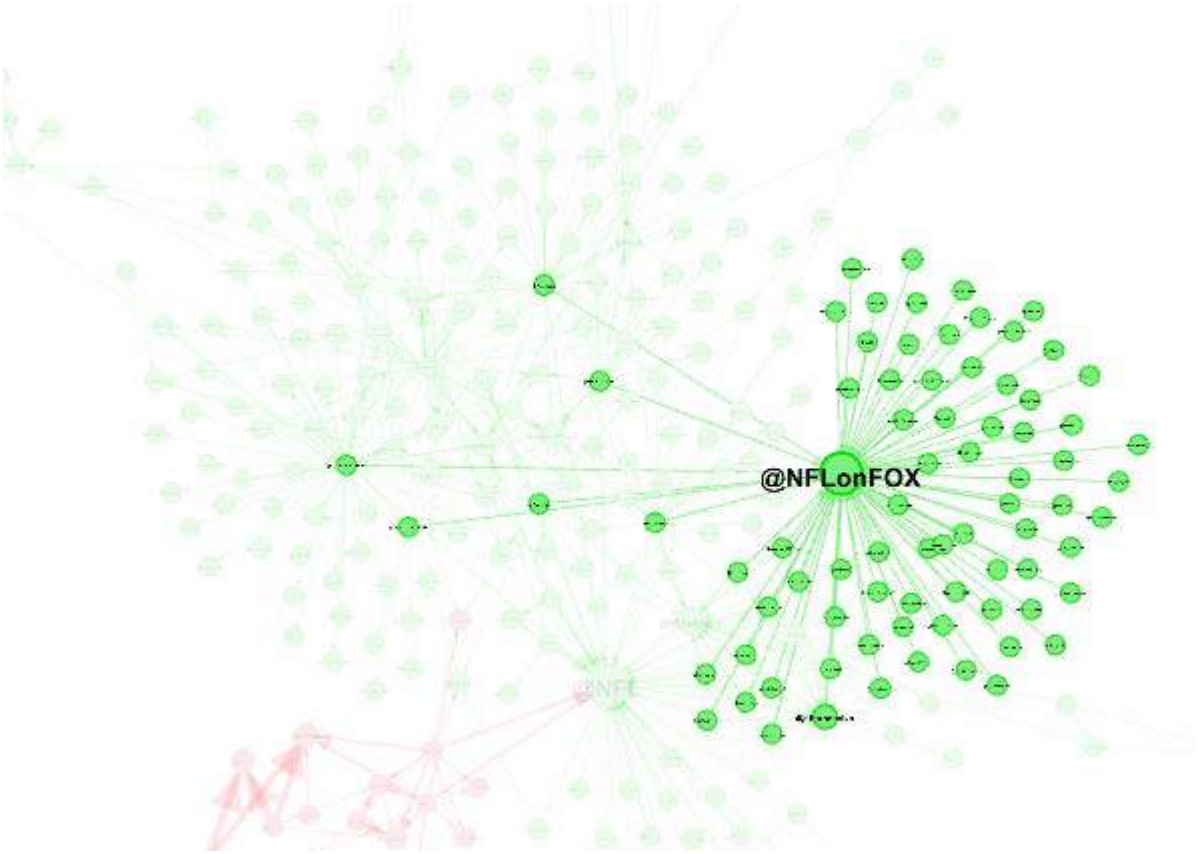


Figure 14: Expected behavior of a legal broadcasting (unidirectional behavior)

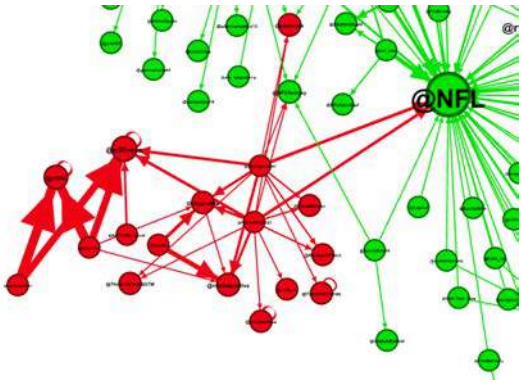


Figure 15: Twitter botnet (chaotic and multidirectional)

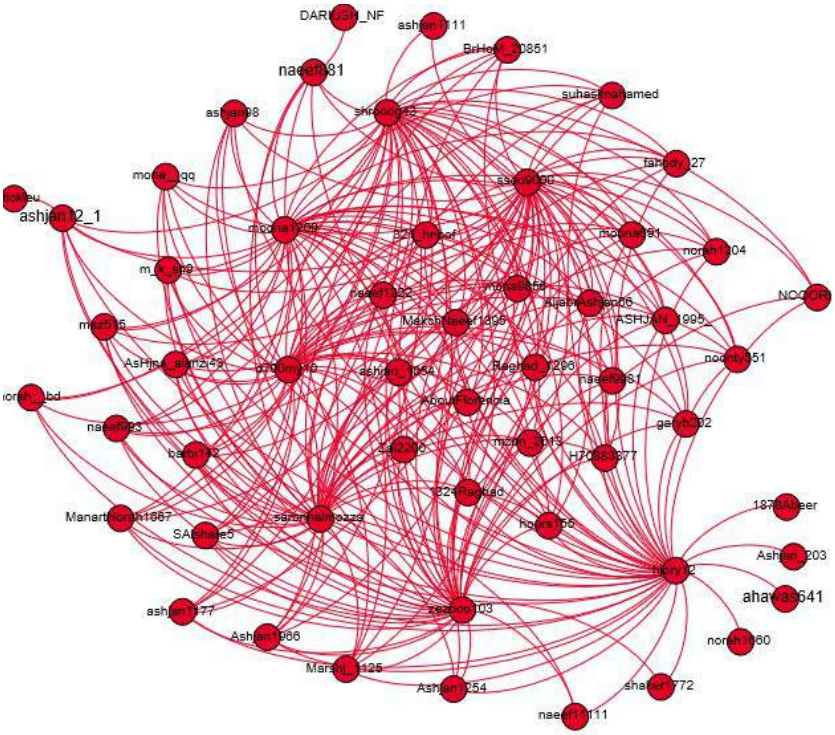


Figure 16: Another "typical" Twitter botnet image

- Twitter botnets are **not usually** focused on one specific sport; they consider any major sport event to promote their illegal broadcasting.

- The size of the botnets varies; the biggest botnet was composed by 183 profiles and the smallest one by 12 profiles.

- During the **5** monitored **weeks**, the **average percentage** of illegal broadcasting related tweets per sport are:

    - American Football: **0.53%** (45.756 tweets)
    - Rugby: **0.15%** (1.200 tweets)
    - Football: **0.14%** (53.304 tweets)
    - Basketball: **0.07%** (9.132 tweets)
    - Baseball: **0.06%** (2.381 tweets)
    - Cricket: **0.03%** (333 tweets)

- And it seems the three top countries broadcasting illegal sport content are United States, Spain and Germany.

- As expected, tweets **related to football with embedded links** (to illegal streaming websites) are more common than the other sports.

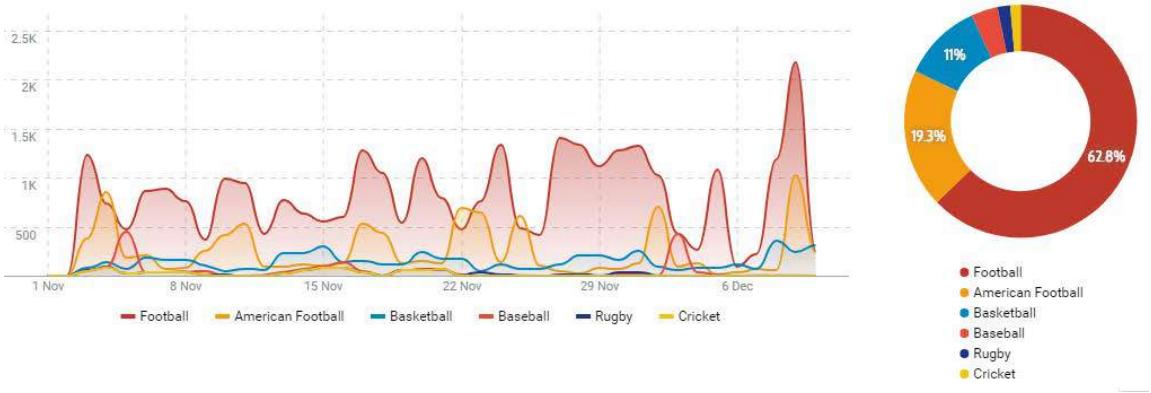


Figure 17: Tweets with embedded link are more commonly related to Football

7. Recommendations

- To prevent this type of illegal broadcasting, the TSA (Telco Security Alliance) has different methods to fight against this particular cybercrime which goes from specific procedures with the main ISPs in the world to perform the necessary takedowns up to investing in start-ups such as Smart Protection (http://smartprotection.com/) which works to protect intellectual property and improper use of your digital contents.

- We have also used Aldara's algorithms in reputational crisis and it has helped us to blow away the artificial "noise" (botnets or automatic accounts). This type of **cleaning leads** to the analysis of the real conversation and **identify the real scope of a crisis**.

8. Glossary

- **Attacker**: actor whose activity includes cyberattacks against the assets of a company or government.
- **Behavioral Flamingo Score (BFS)**: Influence metric developed by ElevenPaths to measure the influence of any social profile in a social network. It is based on how other profiles interact with the evaluated profile.
- **Broadcaster**: actor whose activity is mainly focused on spread messages instead of performing any kind of attack.
- **Conversation**: group of tweets associated to the same topic.
- **Conversation Graph**: graphical representation of one conversation (see Conversation).
- **Influencer**: Individuals who have the power to affect decisions of others because of their (real or perceived) authority, knowledge, position, or relationship.
- **Profile**: concept associated to each user of one social network.
- **Users Graphs**: graphical representation of the social relationships (see Social Relationship) starting from one or more profiles and taking 2 or more steps. That is, from user A, 1 step would return every relationship from and to the profile B. A second step would return every relationship from and to every profile obtained in the step 1 and so on.

About the Telco Security Alliance

The alliance, integrated by Telefónica, Etisalat, SoftBank and Singtel, is one of the world's biggest cyber security providers, with more than 1.2 billion customers in over 60 countries across Asia Pacific, Europe, the Middle East and the Americas. Through their combined resources and capabilities, the group can protect enterprises against the rising cyber security risks as the information security environment becomes increasingly complex.

Through the alliance, members can achieve operational synergies and economies of scale that will eventually help lower costs for their customers. The group's members operate 22 world-class Security Operation Centres (SOCs) and employ more than 6,000 cyber security experts. To expand their global footprint, the alliance is open to bringing in new members over time.

Under the agreement, the group will share network intelligence on cyber threats and leverage their joint global reach, assets and cyber security capabilities to serve customers worldwide. Leveraging each member's respective geographic footprint and expertise, the alliance is able to support each other's customers anywhere and anytime, allowing them to respond rapidly to any cyber security threats.

To enhance their cyber security portfolio, the members will also look into the possibility of developing new technologies such as predictive analytics using machine learning and advanced cyber security for the Internet of Things. The alliance will also consider developing a joint roadmap for the evolution of their security portfolios and explore joint investments in security products and services, SOCs, platforms, start-ups and R&D.