

# Twitter's Trending Topics exploitation patterns

Despoina Antonakaki

Paraskevi Fragopoulou, Sotiris Ioannidis

iSocial Meeting, February 4-5th, 2014

# Online Users

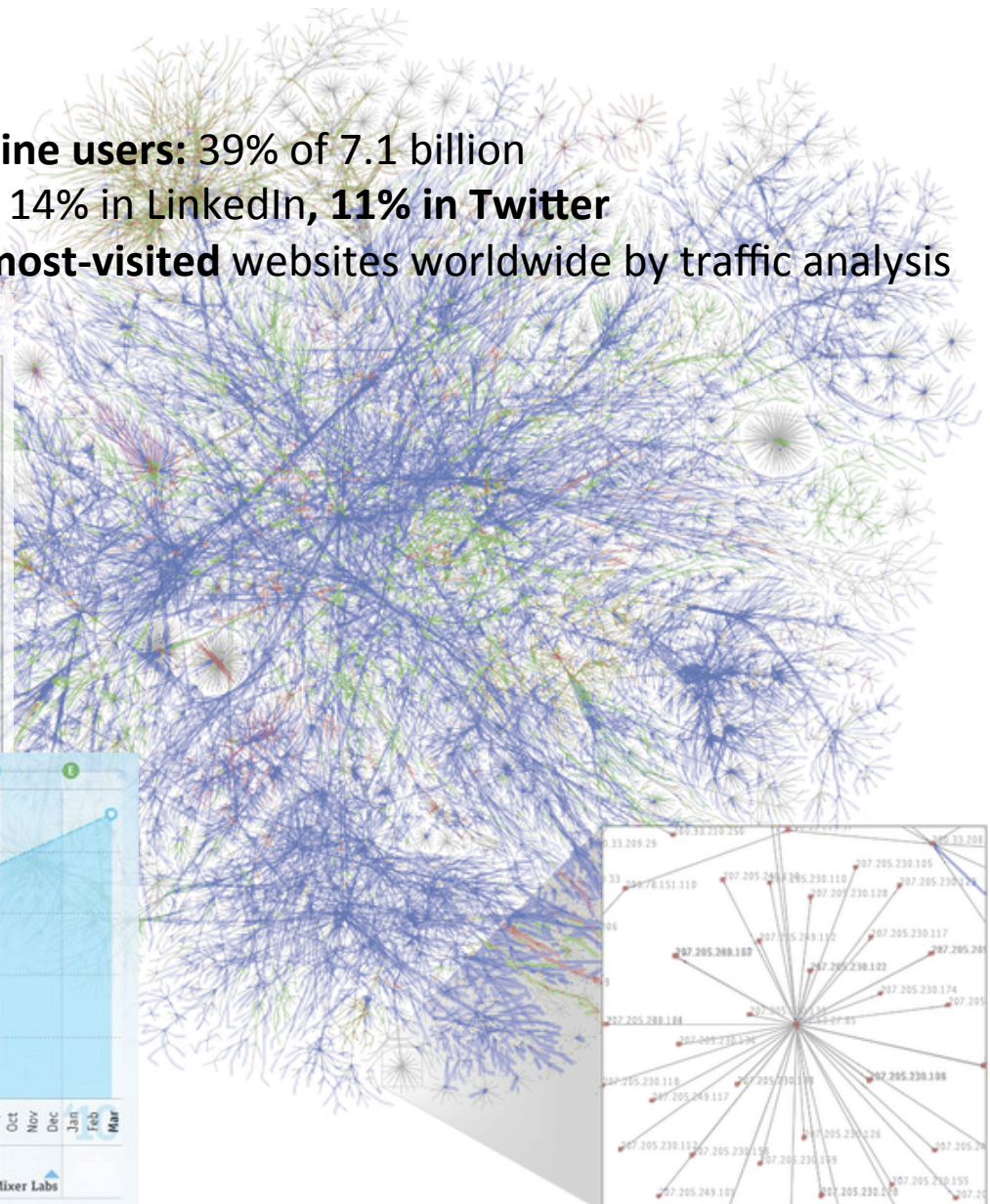
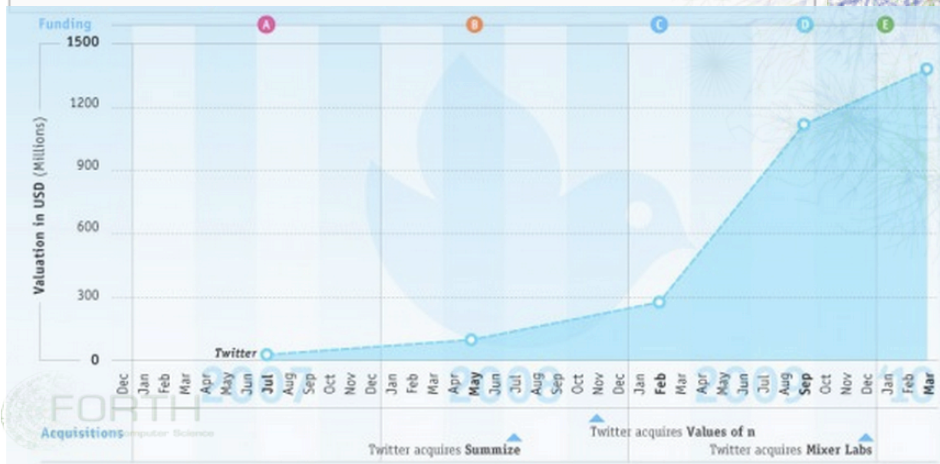
- World populations percentage of **online users**: 39% of 7.1 billion
- 58% are **in any SN**, 56% in Facebook, 14% in LinkedIn, **11% in Twitter**
- Alexa's web: Twitter one of the **ten-most-visited** websites worldwide by traffic analysis

## Worldwide Internet users

	2005	2010	2013 <sup>a</sup>
World population <sup>[3]</sup>	6.5 billion	6.9 billion	7.1 billion
Not using the Internet	84%	70%	61%
Using the Internet	16%	30%	39%
Users in the developing world	8%	21%	31%
Users in the developed world	51%	67%	77%

<sup>a</sup> Estimate.

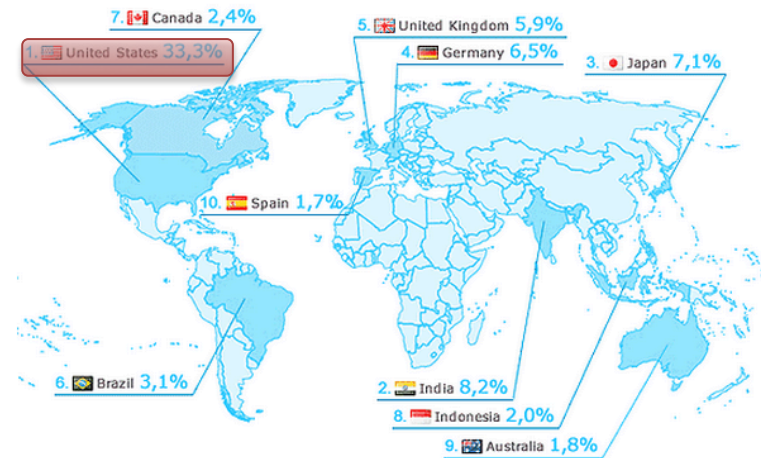
Source: International Telecommunications Union.<sup>[4]</sup>



# Twitter in numbers

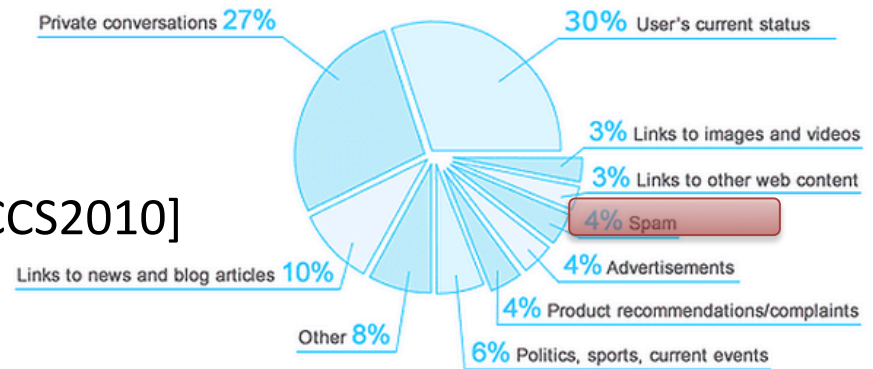
- Fast growing OSN with 650 million **registered** users
- 232 million monthly **active** users - daily tweets: March 2011 140 million, now 500 million tweets per day in the world
- Twitter **queries per day**: 2.1 billion

Top 10 countries (percent of site traffic)



In our experiments we apply to US data (1/3 of worldwide traffic)

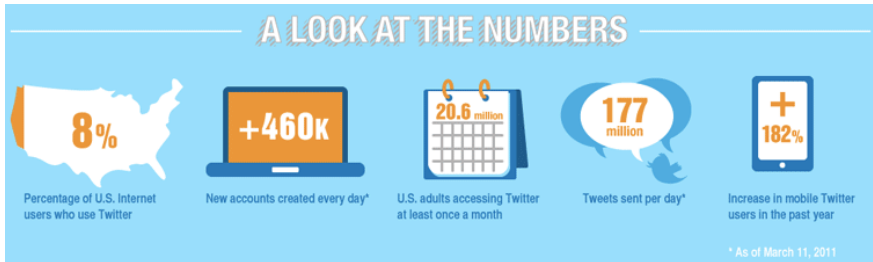
## Twitter message types



## Spam: 8% of 25 million URLs [Chris Grier, CCS2010]

- 2011: 3.5 billion malicious tweets

The highest spam level - almost **11%** of Tweets was noted in August 2009.



# The problem of Spam in Twitter

Twitter is **ideal** place for spammers:

- **Public** content
- **Easy** creation of network (without approval of users)
- **Missing filtering** mechanism

## Twitter Features used by spammers:

- **Trend hijacking:** Ride on success of trends
- **Trend setting:** Create trending topic
- **Callouts@:** Mentions used to personalize
- **Tweet hijacking:** RT prepending the tweet with spam URLs
- **Retweets:** RT of blacklisted URLs

[Chris Grier, CCS2010]

# Feature frequency in spam by blacklists

- 8% of 25 million URLs posted point to **phishing, malware, and scams** listed on popular blacklists.
- **Blacklists lag behind Twitter:** most spam messages appear 4–20 days before the URLs embedded in messages become flagged

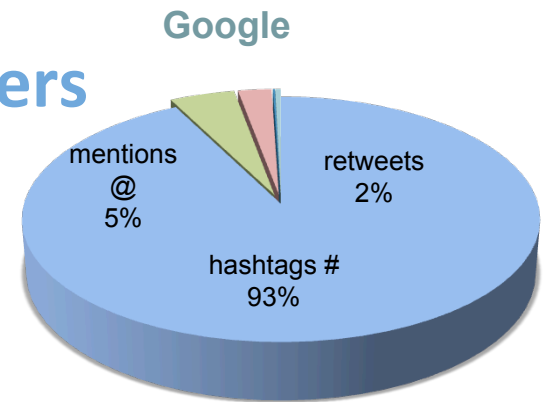
[Chris Grier, CCS2010]

**hence, we are not counting (entirely) on blacklists!**

## Twitter - A marketplace for spammers

- **Twitter account sellers:** accounts composing & sending tweets, Paid Owner
- **Sponsored Tweets:** pay users to host ads in their tweets
- **Spam affiliate programs:** bots that create accounts and post spam links (exploiting trends) .

[Thomas SIGCOMM '11 ]



hashtags #	70.10%
mentions @	3.50%
retweets	1.80%
random sample of tweets	0.10%
a random sample of tweets containing URLs.	0.30%

**This work deals with Spam in Twitter**

# Background work

## General about Spam:

- **Malicious URL detection** in “*Design and Evaluation of a Real-Time URL Spam Filtering Service*” [**Security and Privacy’11**]
- Measuring the **conversion rate** of spam in “*Spamalytics*” [**CACMSpam09**]
- Distribution **infiltration** for measuring spam campaigns from the inside in “*On the Spam Campaign Trail*” [**Usenix 2008**]

## Privacy in SNs:

- Building working prototype of a variant of Twitter protecting tweet contents, hash tags, interests in “*Hummingbird: Privacy at the time of Twitter*” [**Security and Privacy ’12**]
- Inferring private from public attributes in SNs in “*You are what you like information leakage through users' interest*” [**NDSS ’12**]
- Decoupling the retrieval of user-specific content from the loading of a social plugin in “*Privacy-Preserving Social Plugins*” [**Usenix ’12**]

# Background work

## Spam in Twitter:

- Reconstruct spam into campaigns for clustering: six features used by spammers & provide **online spam filtering** in OSN server in *“Towards Online Spam Filtering in Social Networks”* [NDSS '12]
- Detecting if account is automated from tweet timestamp in *“Detecting and Analyzing **Automated Activity** on Twitter”* [PAM '11]
- A complete analysis of spam in twitter: analyzing compromised accounts, analyzing features spammer use, group spam to campaigns, **identify trends that distinguish phishing malware** and test blacklists in *“@spam, Detecting and Analyzing Automated Activity on Twitter”* [CCS 2010]
- **Identify spam accounts**, campaigns, identify marketplace/spam affiliate programs in *“Suspended Accounts in Retrospect: An Analysis of Twitter Spam”*[IMC '11]

## A new approach

**Given that:** spammers exploit trends

**We need to:** investigate the features of:

- Spam tweets (#trends, RTs, @, #followers, ...)
- Spam campaigns (graph structure)

**In order to:**

- Understand spam
- Identify spammers vs. legit



## Our Dataset

- Top ten trending topics every ten minutes for 10 days
- Ten twitter accounts downloading in parallel
- In average: 1.500.000 tweets / day
- **Data from US**
- 1.181 USA trending topics
- In average: 400.000 different users / day
- Use blacklists and heuristics to identify spammers

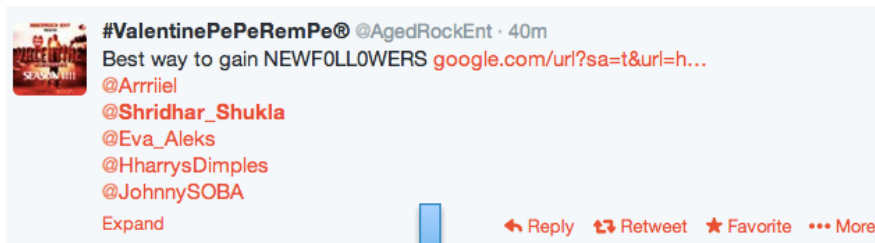
# Preprocessing of data

For every **user** we download & save :

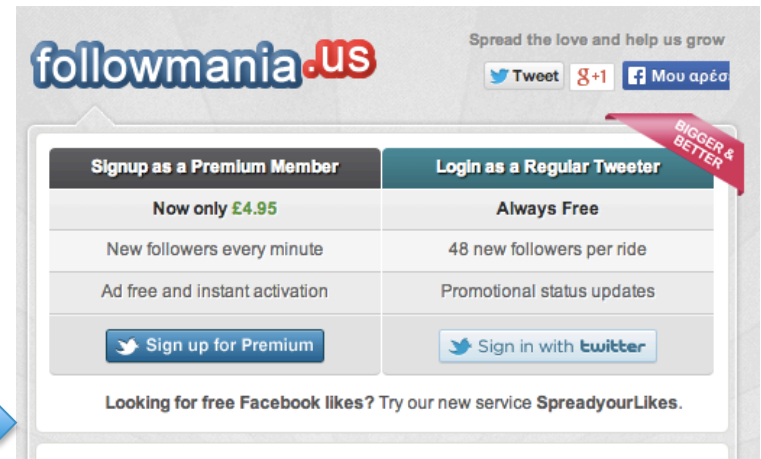
- #tweets
- #trends
- #different trends
- #number of tweets for each trend
- #spam\_urls (according to blacklists)
- #legit\_urls
- #percentage of tweets with trends
- #percentage of tweets with urls
- #RTs (not yet included)
- #User Mentions (not yet included)
- #followings (not yet included)
- #followers (not yet included)

# Google heuristic

Spam URLs are hidden in a Google search results link



[www.google.com.tr/url?sa=t&rct=j&q=&esrc=s&frm=1&source=web&cd=4&cad=rja&sqj=2&ved=0CEAQFjAD&url=http%3A%2F%2Fwww.followmania.us%2F&ei=FveHUoDBHITdswa2q4HYCw&usg=AFQjCNG1fWSqWsWxdI2QCMSeu3WGXHETaw&sig2=fEuJc66pEvdq9vLp2XyGdg&bvm=bv.56643336,d.Yms...](http://www.google.com.tr/url?sa=t&rct=j&q=&esrc=s&frm=1&source=web&cd=4&cad=rja&sqj=2&ved=0CEAQFjAD&url=http%3A%2F%2Fwww.followmania.us%2F&ei=FveHUoDBHITdswa2q4HYCw&usg=AFQjCNG1fWSqWsWxdI2QCMSeu3WGXHETaw&sig2=fEuJc66pEvdq9vLp2XyGdg&bvm=bv.56643336,d.Yms...)

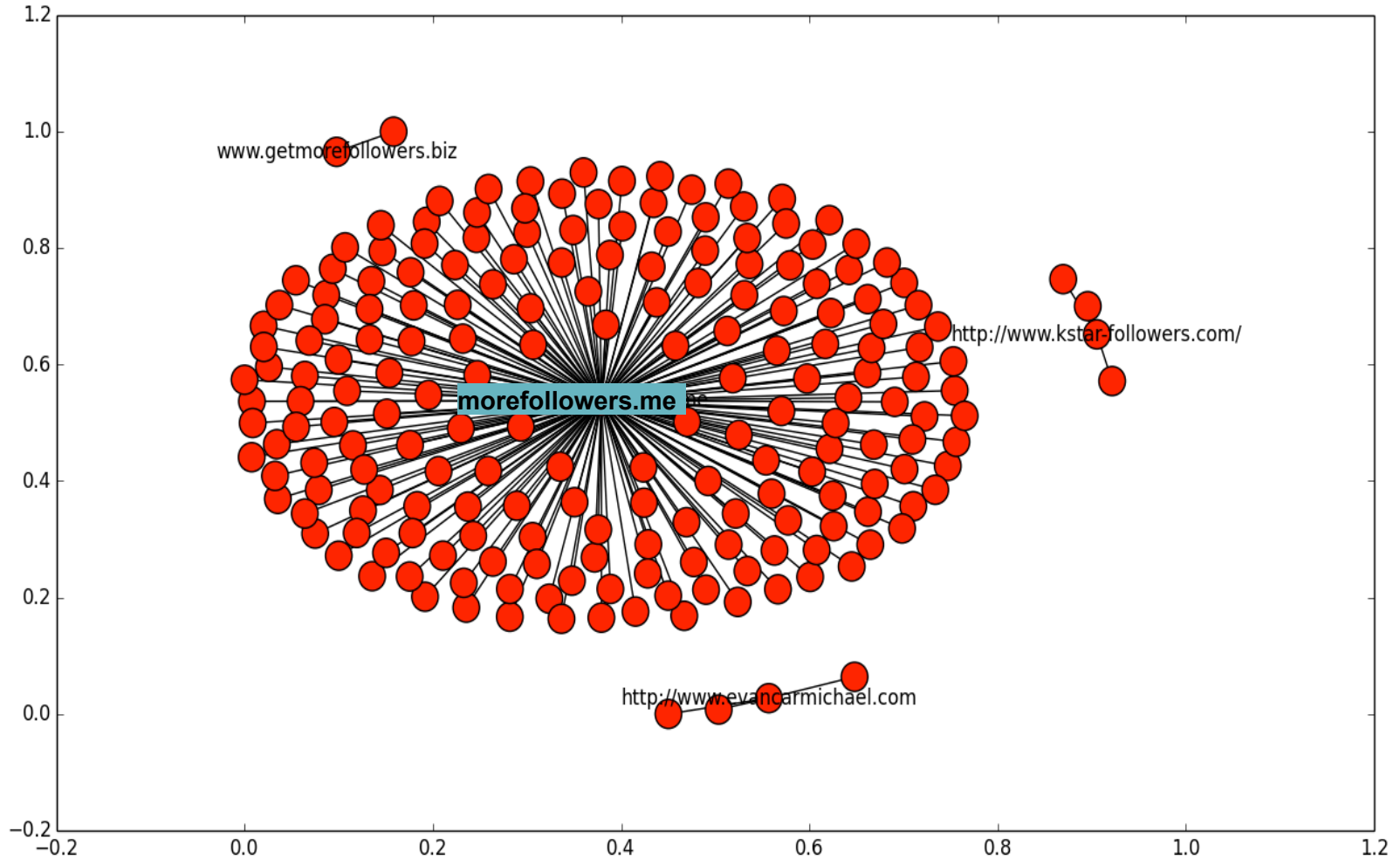


- Fast identification
- Reliable
- **Use as a golden standard**
- **3.500 spammers ( ~1%)**

Growing industry:

- **“Twitter follower markets”**
- 1-10 millions of victims
- 40\$-200\$ per victim

# How does a “get more followers” campaign looks like?



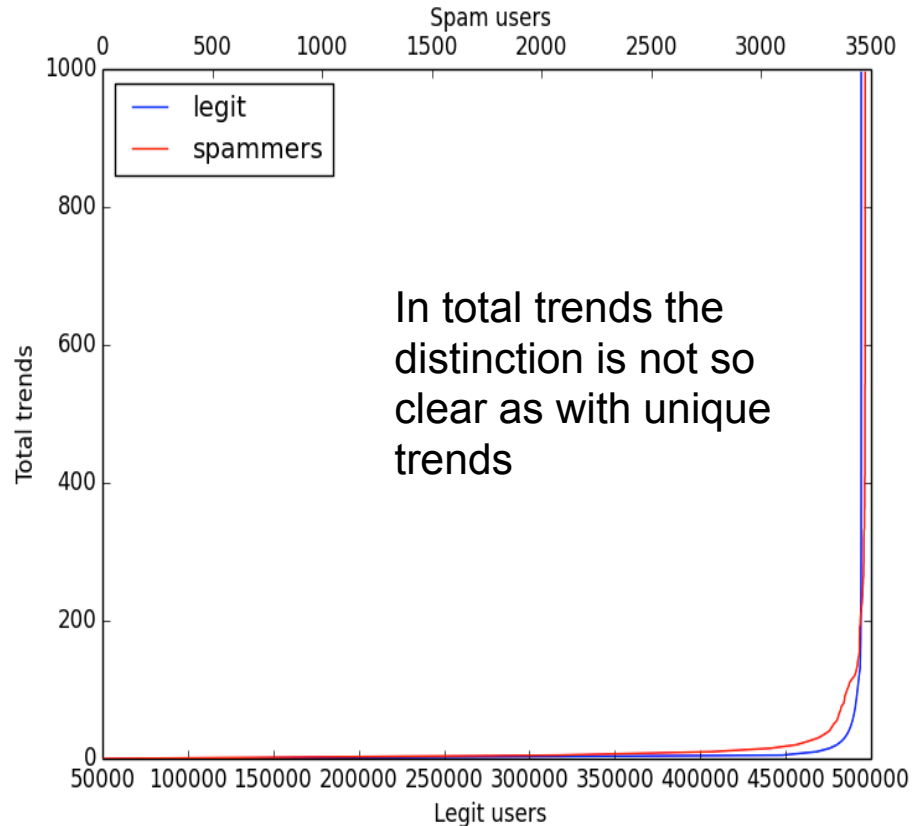
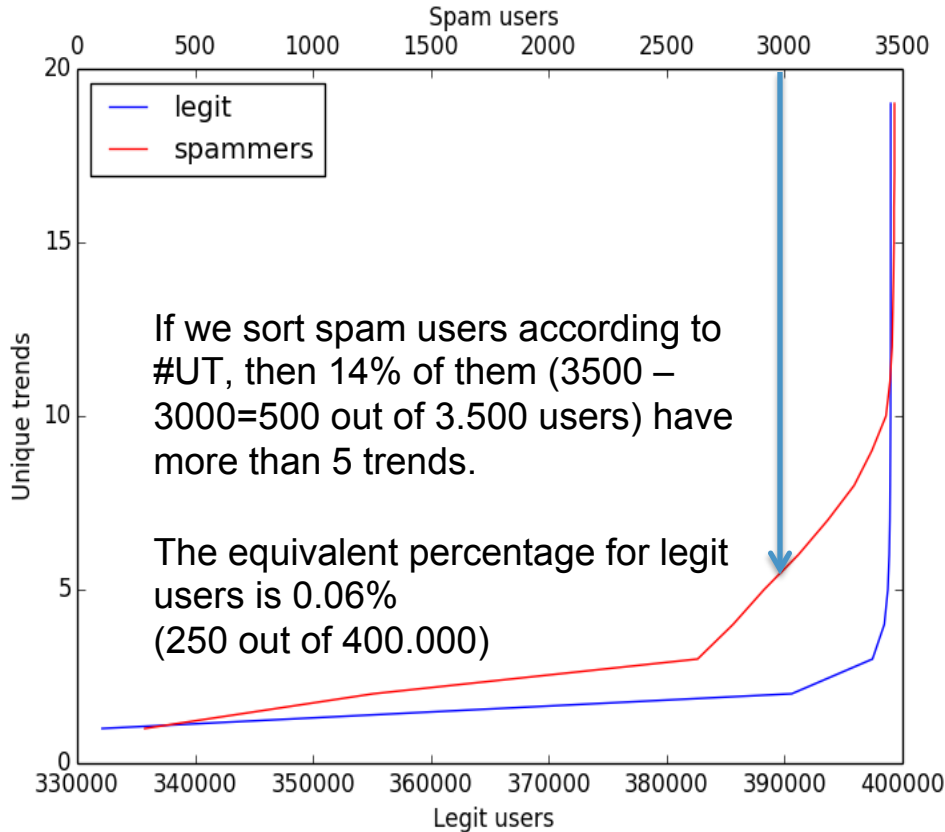
**Central nodes are the campaign links**

## Why not blacklists ?

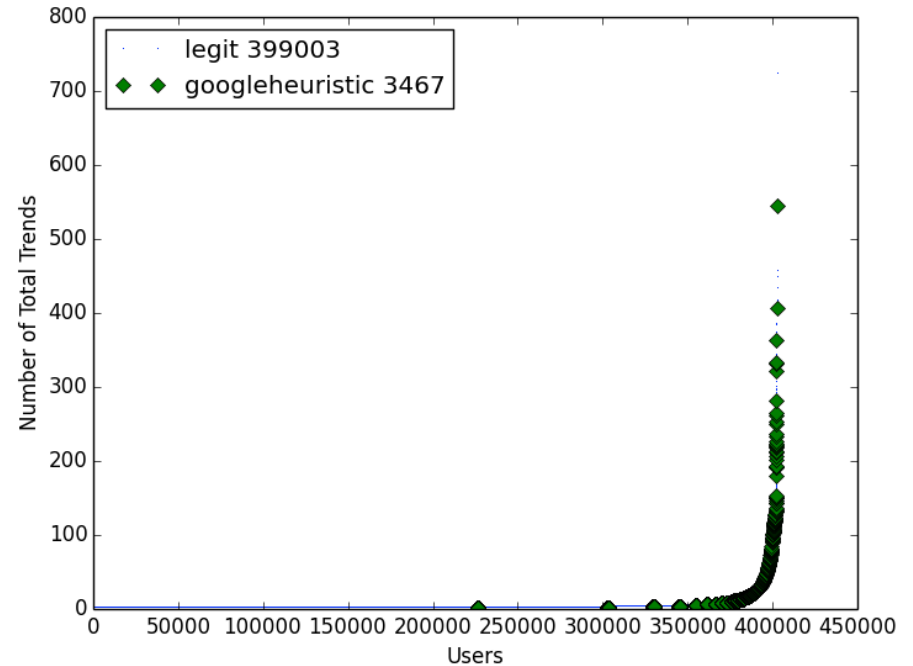
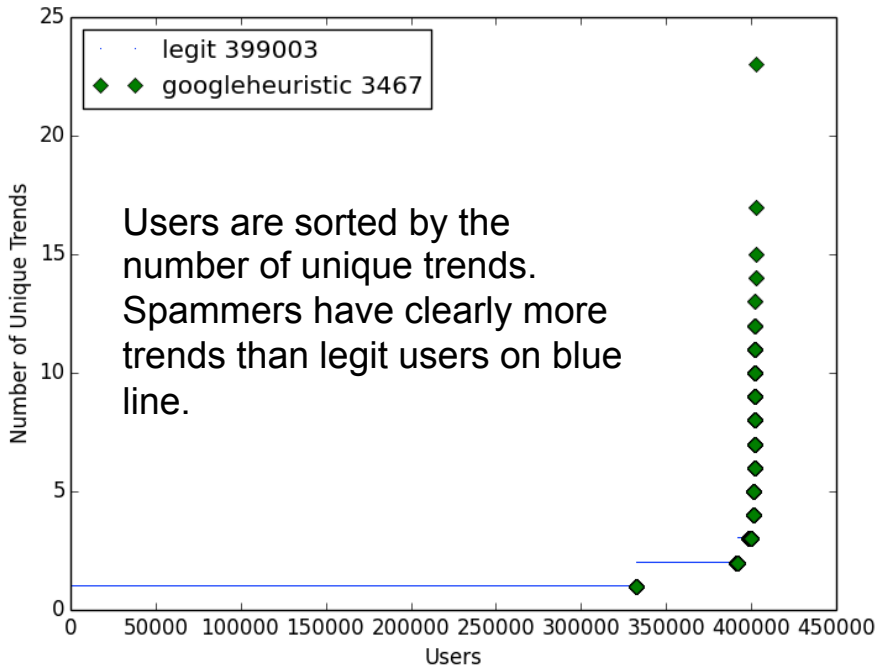
- SpamHaus
  - Many False Positives (i.e. instagram.com)
- Surbl – not so strict filters
  - High specificity (Few False Positives)
  - Low sensitivity (Many False Negatives)
- Google Safe Browsing API
  - Only for phishing and malware
  - 10.000 links lookup / 24hrs

According to these data what are the patterns that distinguish spammers?

# CDF of unique & total trends



# Plot of unique & total trends per user





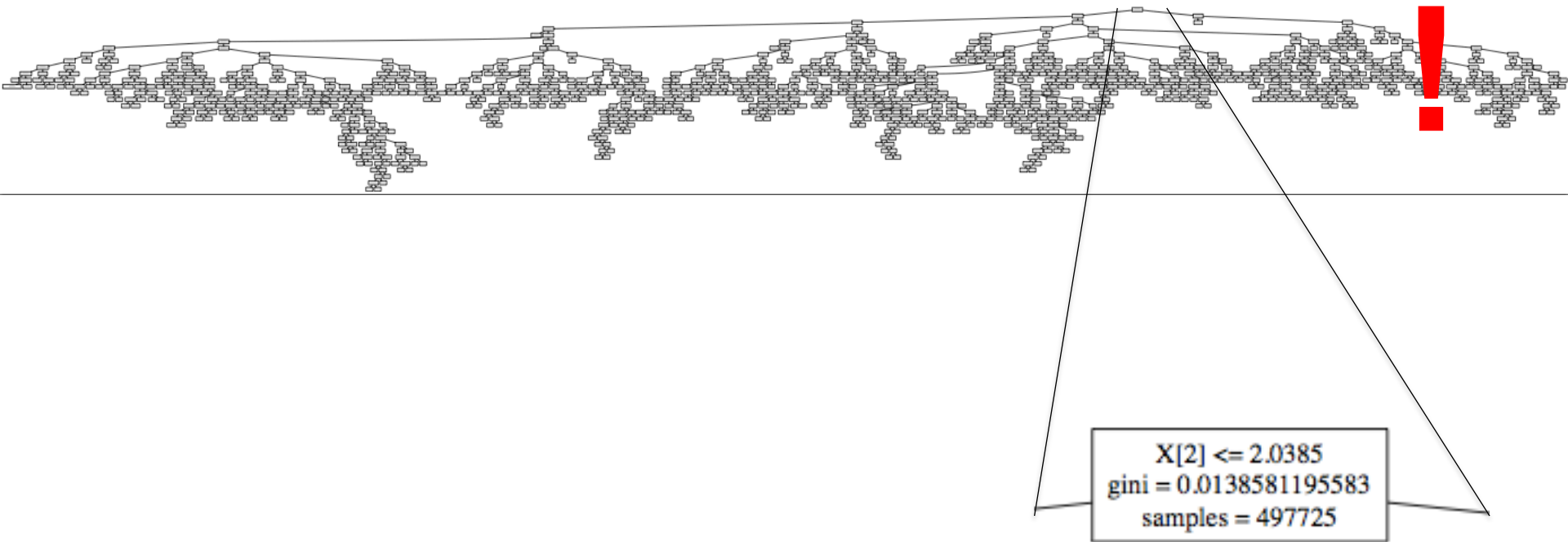
# Classification

- 490.000 Legit users, 3.500 Spam users
- 10-fold Cross Validation

Method	True Positive Ratio (Spam identified as Spam)	True Negative Ratio (Legit identified as Legit)
Logistic Regression	55.1%	99.9%
SVM	55.0%	99.8%
Decision Trees	59.2%	99.9%
Random Forest	59.3%	99.9%

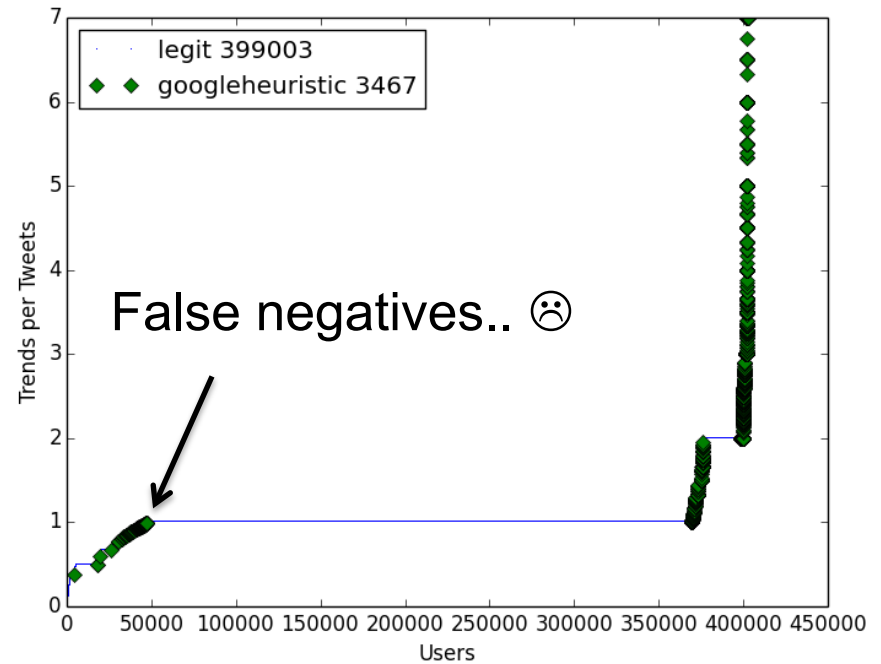
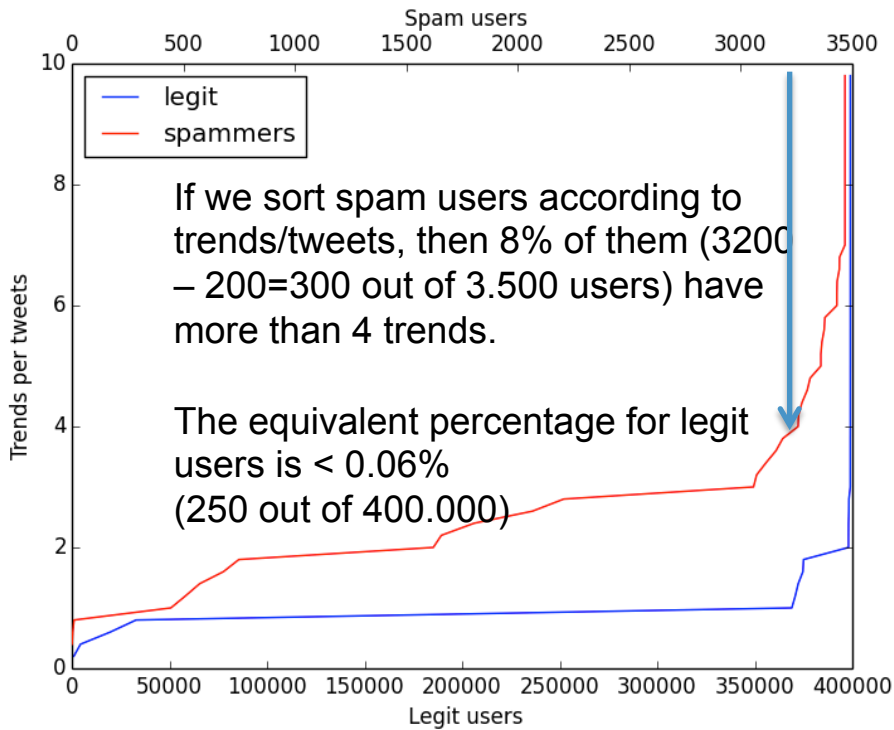
## Decision Tree Visualization

Complex and overfitted but.. the first rule is always interesting..



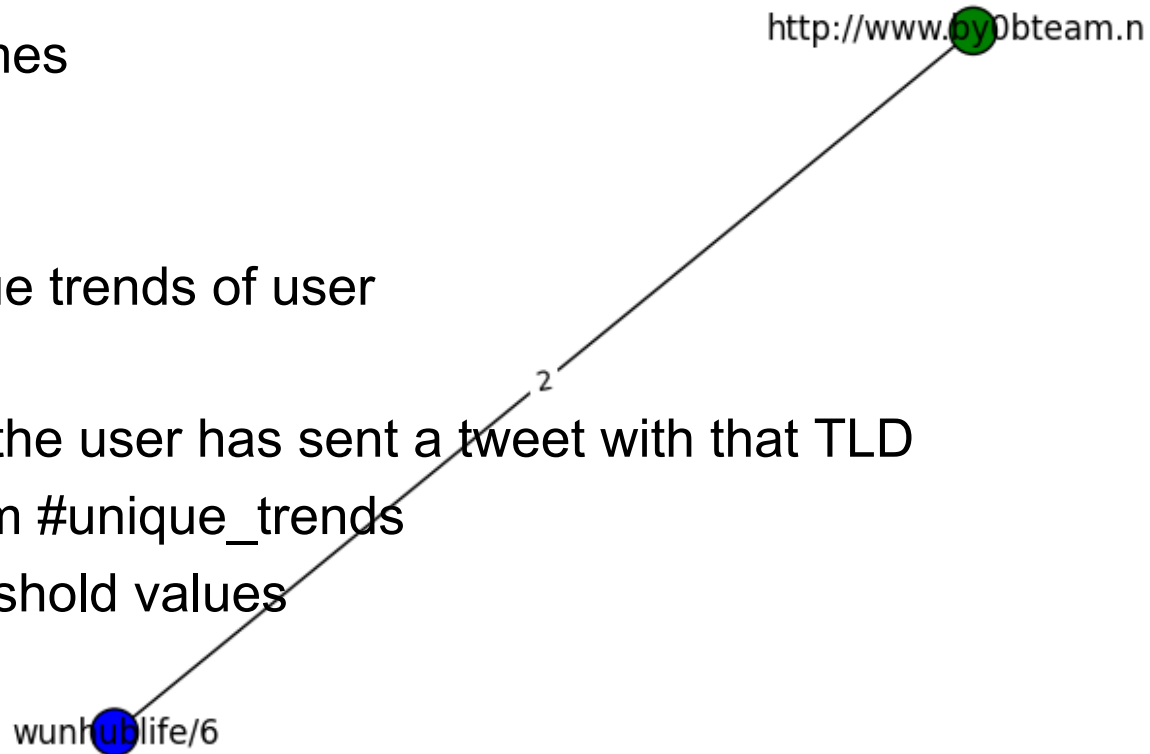
$X[2]$  is #trends / #tweets

# Plot: #Trends / #Tweets



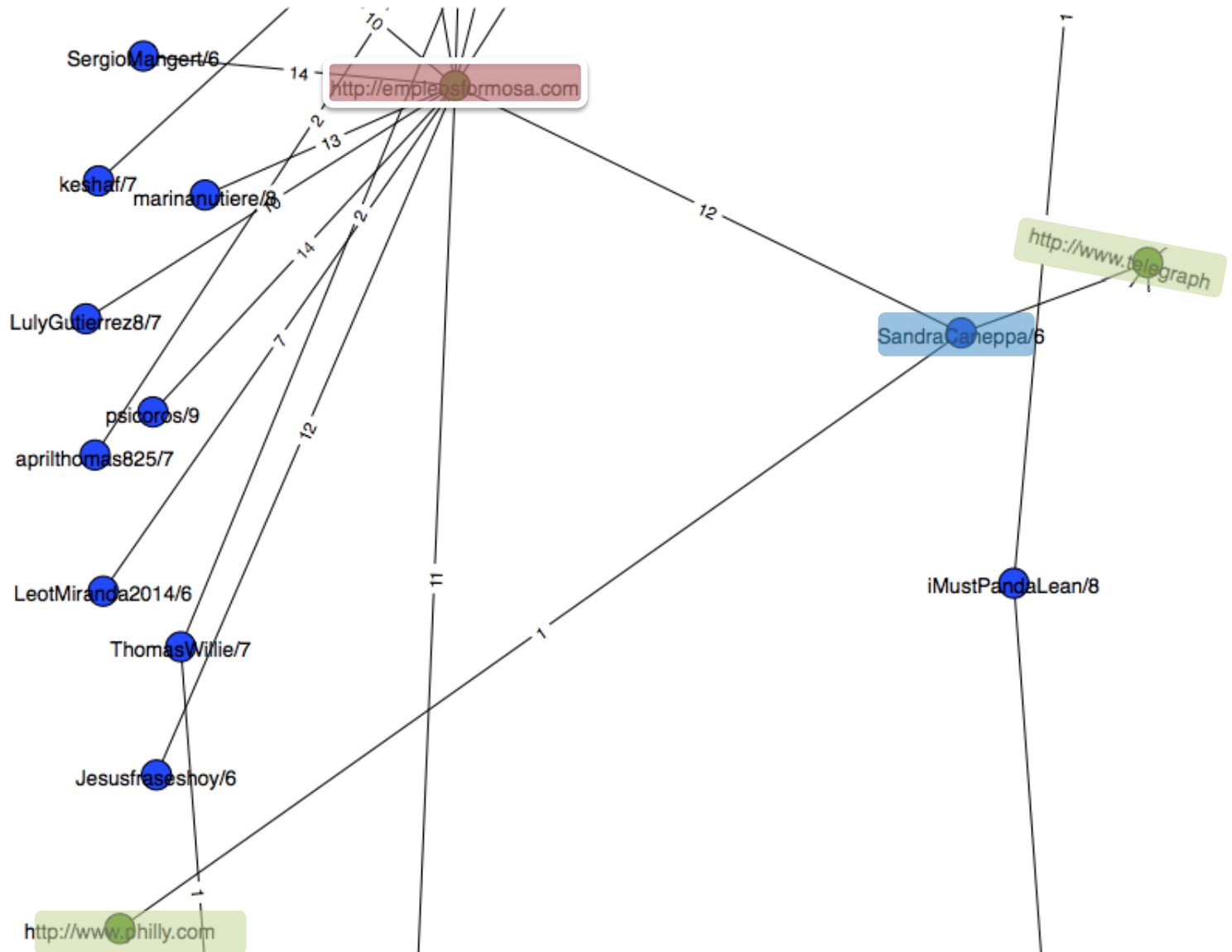
# Discover spam campaigns

- Plot TLDs linked by top trend users:
  - For top trend users plot the TLDs of links contained in tweets
  - Manually check the TLDs
  - **Blue nodes**: usernames
  - **Green nodes**: TLDs
  - Labels on users:
    - username/#unique trends of user
  - Labels on edges:
    - number of times the user has sent a tweet with that TLD
  - Threshold = Minimum #unique\_trends
  - Plot for different threshold values









## Your Story Net

Home



**Online Share Trading System - Proven For Success, Learn Why People are Necessary Trading about the stockmarket is designed with a great deal of likelihood of a lot of dough. It is exciting idea and a lot people just rush out and start currency trading**

Writing & Speaking > Book Reviews | By: Lesley Cunningham (01/21/14)  
 Trading to the stock the gives the substantial amount of likelihood a small fortune. This is an exciting idea and the majority people just gives out as well as begin buying and selling whatever stocks they've experienced the Some people try to be diligent ourselves and wind up depreciating to complex financial "instruments". Investors just like that didn't have a good quality stock market the system on hand. Without really knowing what you are doing, it is possible to will lose a lot of dough. Even those gives with stock market the lose money. I might come across having systems it is in place that you'll follow which keeps the sentiments off the beaten track. Yes, you heard that right, emotions are the biggest reasons people throw money away on stocks. If you have an established system to adhere to you may take off the emotional aspects away from business, all things considered it really is business. As soon as you clear away the emotions, you happen to be more objective and do business with your mouth and figures presented

Already a Member?

Username:

Password:

Remember Me:

[Signup to submit articles >](#)

Search:

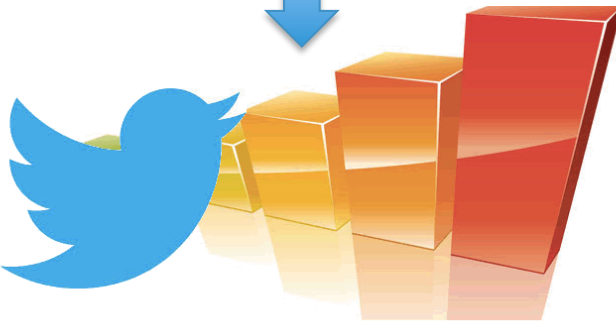
Categories

- Arts & Entertainment
- Business
- Communications
- Computers
- Environment
- Fashion
- Finance

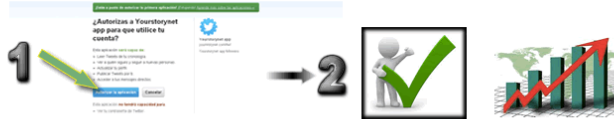
Pages

Recent Articles

- Core Elements For guilt-free - Updated
- Enter The Fun With Fellow Backpackers And Have The Excitement Of A Lifetime



**Start now FREE**



## Authorize Follow Untofollow auto 24 to use your account?

This application **will be able to:**

- Read Tweets from your timeline.
- See who you follow, and follow new people.
- Update your profile.
- Post Tweets for you.
- Access your direct messages.

This application **will not be able to:**

- See your Twitter password.

You can revoke access to any app at any time from the Applications tab of your Settings page. By authorizing an application you consent to operate under Twitter's Terms of Service. In particular, some



Follow Unfollow auto 24  
[follow.yourstorynet.com/](http://follow.yourstorynet.com/)

Follow Unfollow auto 24



Online Share Trading System - Proven For Success, Learn Why People are Necessary Trading about the stockmarket is designed with a great deal of likelihood of a lot of dough. It is exciting idea and a lot people just rush out and start currency trading

Videos Relacionados:  
**Empleos trabajos articulos casas autos motos electrodomesticos**





*High payout rates, good monthly income, fast & customized url shortener, and non-intrusive ads to my blogs. Also I've got a lot of visitors with AdF.ly as a publisher. Keep your awesome service going on!*

## Earn fast easy money schemes:

- URL shrinking service that pays the users depending on the #clicks &
- Ponzi scheme !

http://thefanmarket.com

4

TheFanMarket/6

13

/hinc

7

http://www.takipci.com

RiseOfStades/7

3

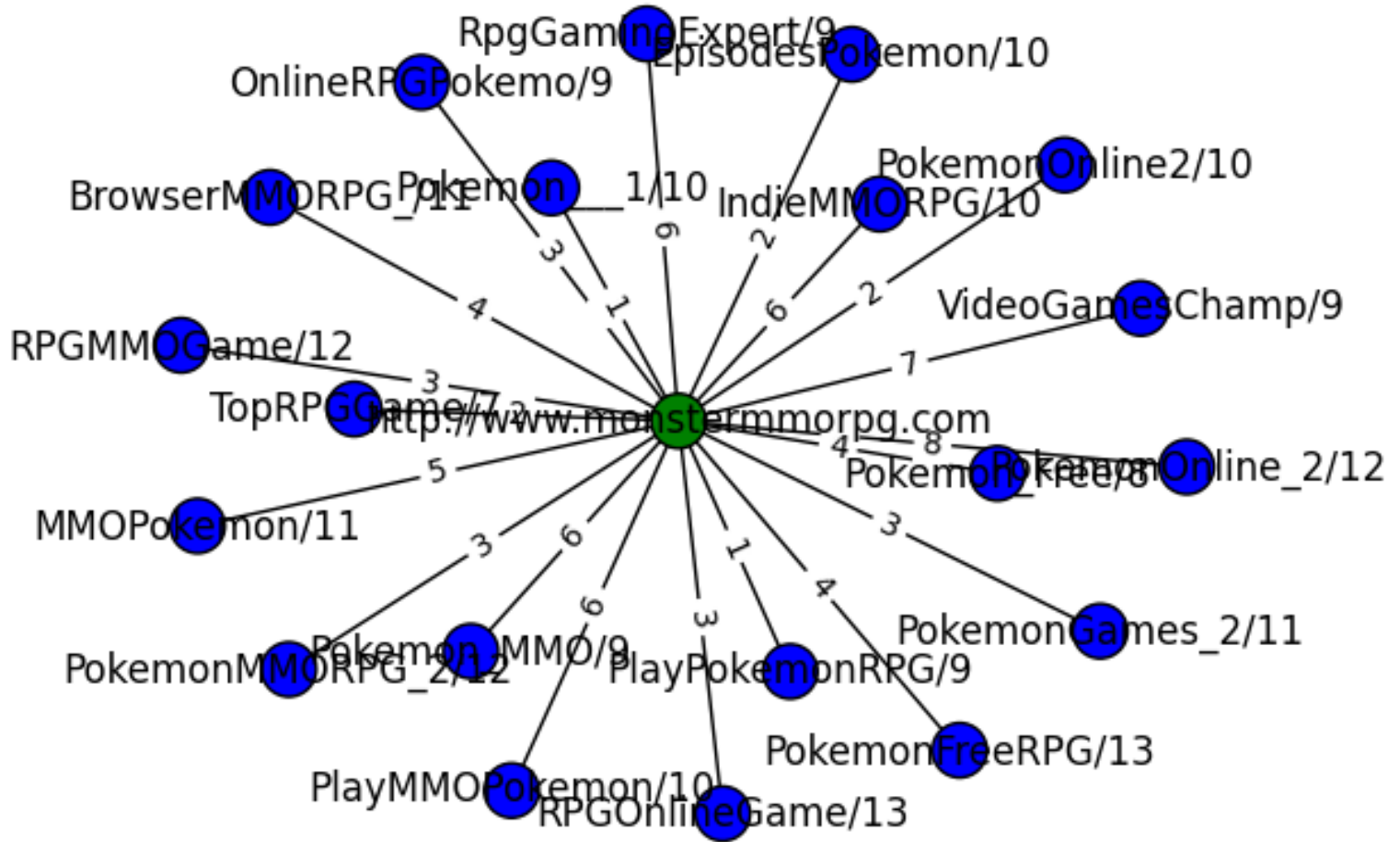
http://www.moiegogo.com

- Social media followers market

- Social media followers contests

- Social media followers market

- “An International Crowd funding Platform to Raise Money”



Monster MMORPG game online

## Discussion

- Metrics like trends/tweets, unique\_trends can distinguish more than half of spammers
- We shouldn't rely on blacklists to identify spam
- Campaign plots show that there are more complex "spammish" techniques.
- The complete picture is not yet revealed..



## Future work:

- Run for more days / weeks / months
- Assess more heuristics / blacklists
- Build simple classification mechanisms
- What are the graph properties of spam campaigns?



# Collaboration: Distributed Spam Filtering

Adjust the procedure in a gossip algorithm:

- Trend and Tweet collection is already distributed among multiple accounts
- Store in MongoDB
- Run ensemble methods for classification:
  - Adaboost
  - Random Forest Classifier
- Merge campaign graphs in order to identify larger spam campaigns,
- “get the complete picture”