

An empirical approach towards a whom-to-mention Twitter app

Soumajit Pramanik¹, **Maximilien Danisch**², Qinna Wang²
Jean-Loup Guillaume³, Bivas Mitra¹

1- CNeRG, IIT Kharagpur, India

2- Complex Networks team, LIP6, UPMC-CNRS, Paris, France

3- L3I, University of La Rochelle

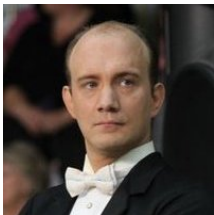
Supported by:

The French National Research Agency - contract CODDDE
ANR-13-CORD-0017-01 and ECIS - Cefipra Indo-French Targeted
Program

This is joint work with:



Soumajit Pramanik (IIT), Maximilien Danisch (LIP6), Dr. Qinna Wang (LIP6)



Prof. Jean-Loup Guillaume (ULR)



Prof. Bivas Mitra (IIT)

What is Twitter?

“Twitter is a social microblogging system which has become one of the most important ways for sharing information online.”

What kind of information?

- **very local:** I'll have a coffee at our bakery in 10min
- **local:** IITkqp party on Friday night, all invited
- **global:** Nicolas Sarkozy is back to politics
- **very global:** There has been a terrorist attack in Paris

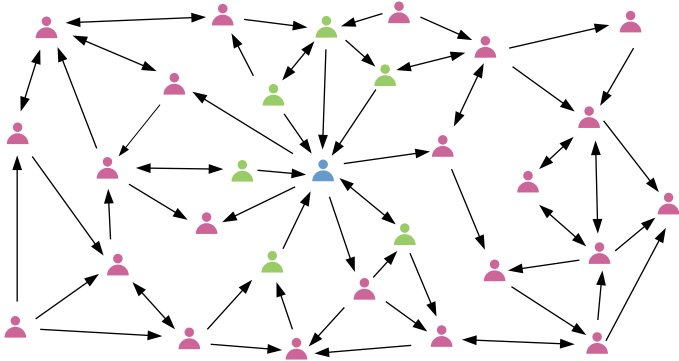
Retweets = propagation

Where do retweets come from?

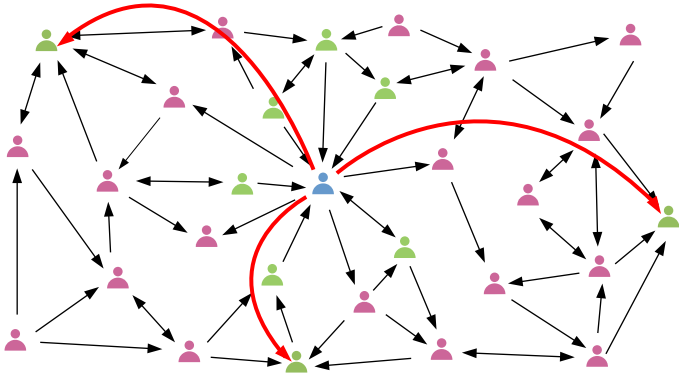
Someone has to see your tweet:

- One of your followers
- One of the followers of someone that retweeted your tweet
- **Someone you mentioned @**
- Someone who saw your tweet searching for keywords or #
- Someone who saw it somewhere else on the Internet

Follow network



You can do something with @



Our goal

How to share an information/opinion with the greatest number of interested people?

→ use Twitter + use mention = use our App

We target:

- ① (very) global tweets and
- ② (very) local tweets

Outline

- ① **Related work on mention recommendation**
- ② Data study to get insights
- ③ Online application
- ④ Experiments “faky app”
- ⑤ Conclusion & perspectives

Three related papers

1- Whom to mention: expand the diffusion of tweets by @ recommendation on micro-blogging systems.

Wang et al. WWW2013

PROS

- First in-depth study of mention in microblogs
- Extract features (interest match, user relationship and user influence)
- A ranking function is trained
- Address recommendation overload problem

CONS

- No real-time experiments
- **No usable app provided**

Three related papers

2- Locating targets through mention in Twitter.

Tang et al. World Wide Web 2014

PROS

- Extract four categories of features (content, social, location and time)

CONS

- The goal is to mention someone interested and trigger interactions (mentions, replies, retweets, friendship) but not to maximize the spreading
- No real-time experiments
- **No usable app provided**

Three related papers

3- Who will retweet this? Automatically identifying and engaging strangers on Twitter to spread information. Lee et al. IUI2014

PROS

- More features extracted for machine learning methods and relevance is evaluated
- Waiting-time model
- Real-time experiments

CONS

- Goal is to mention someone so that he retweets, not to maximize the spreading
- Not real users (“Public Safety News” and “Bird Flu News”)
- **No usable app provided**

What we suggest

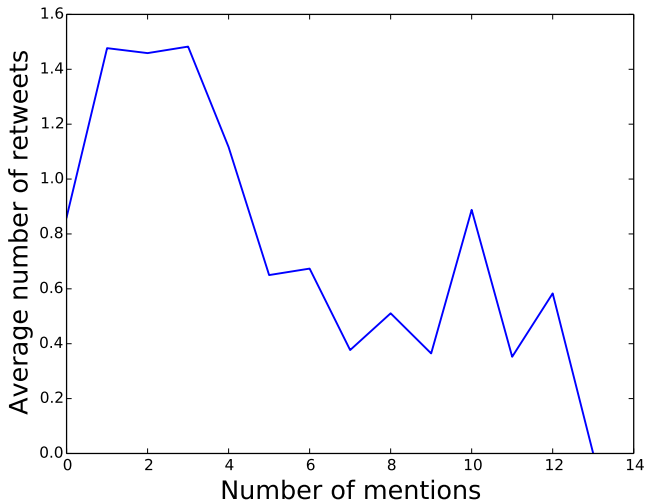
Make a whom-to-mention recommendation system

- to maximize the spread of a tweet,
- working only on real time experiments,
- with a usable app,
- and improve the app every time it is used.

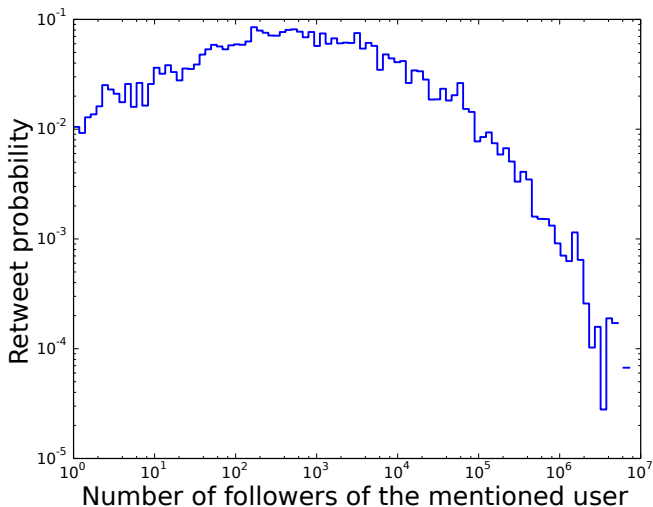
Outline

- ① Related work on mention recommendation
- ② **Data study to get insights**
- ③ Online application
- ④ Experiments “faky app”
- ⑤ Conclusion & perspectives

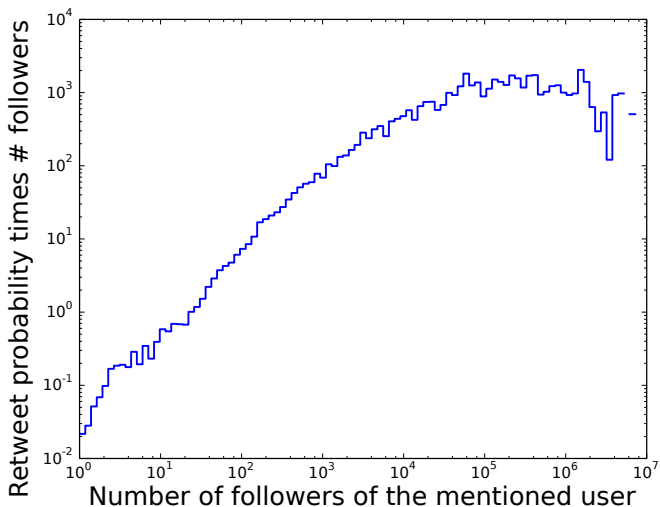
Data study



Data study



Data study



Outline

- ① Related work on mention recommendation
- ② Data study to get insights
- ③ **Online application**
- ④ Experiments “faky app”
- ⑤ Conclusion & perspectives

Online application

Users' database ?

- (very) global tweets → use a database of “random” users
- (very) local tweets → target non reciprocal followees

Online application

Target users that

- are popular: $f_P(u)$
- are susceptible to retweet the tweet: $f_{RT}(u, t)$

→ i.e. have a high $S(u, t) = f_P(u) \cdot f_{RT}(u, t)$

Approximation

Neglect overlap (as we mention several users) and maximize the number of people that will see the tweet at hop one.

- $\sum_u S(u, t)$
- $f_P(u)$ = number of followers of u
- Learn $f_{RT}(u, t)$ using logistic regression

Logistic regression

Principle:


Regression (or binary classification of vectors) by fitting parameters of a logistic function: $V_{\Theta}(X) = \frac{1}{1+\exp(\Theta^T X)}$ to optimize a convex quality function.

Advantages:

- Returns the probability for a user to retweet;
- A few parameters to store: convenient for an online app.

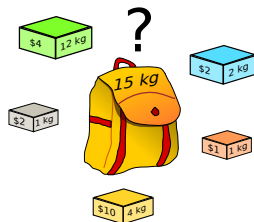
Map into the knapsack problem

The goal is to use the remaining place as efficiently as possible:

 **Maximilien Danisch** @maxdan94 · now
I'll present a work on mention recommendation at the Twitter conference in Lyon #twlyon2015 @StuartWShulman @stanleybaudu @AlainCouttolenc

 **Maximilien Danisch** @maxdan94 · 22s
I'll present a work on mention recommendation at the Twitter conference in Lyon #twlyon2015 @EMLYON @kolb @kerzol @ndriaa @dgosset @twytof

- Solve the knapsack problem with:
 - total budget: $140 - L_{tweet}$
 - user weight: $2 + L_{screenname}$
 - user value: $S(u, t)$



Improving the app

Every time the app is used we

- ① collect users to increase the database
- ② check which mentioned user retweets and improve $f_{RT}(u, t)$

Live Demo

Mention Recommendation System

Username

Tweet

Keywords

Node	Rank
deepikapadukone	9.90256613445
iamsrk	9.31518241019
SrBachchan	9.08844864743
arrahman	8.49729229928
sonunigam	8.18229932923

Outline

- ① Related work on mention recommendation
- ② Data study to get insights
- ③ Online application
- ④ **Experiments “faky app”**
- ⑤ Conclusion & perspectives

Current work: faky app

- Select tweets containing mention(s) made recently on Twitter by random users and retweet them from our fake account replacing the mention(s). → Try to have more retweets!
- We chose the tweets based on keywords and a manual selection in order to have only global tweets.
- 50 tweets lead to:
 - 4 retweets,
 - 2 followers,
 - 5 favourites,
 - 6 mentions,
 - 1 mentioned user blocked us.

Outline

- ① Related work on mention recommendation
- ② Data study to get insights
- ③ Online application
- ④ Experiments “faky app”
- ⑤ **Conclusion & perspectives**

Discussion

Impact of number of mentions

Is knapsack worth it?

Tweeting several times the same tweets mentioning different users

Extreme case

Do not use friendship... use mention! → Twitter2.0

Should not target “bad” users

Sybils (users creating many fake accounts)

Social capitalists (FMFY/IFYFM) <http://www.bit.ly/DDPapp>