Paris Attacks

A pretext to sentiment analysis on social media

M. Gaborit – F. Blain

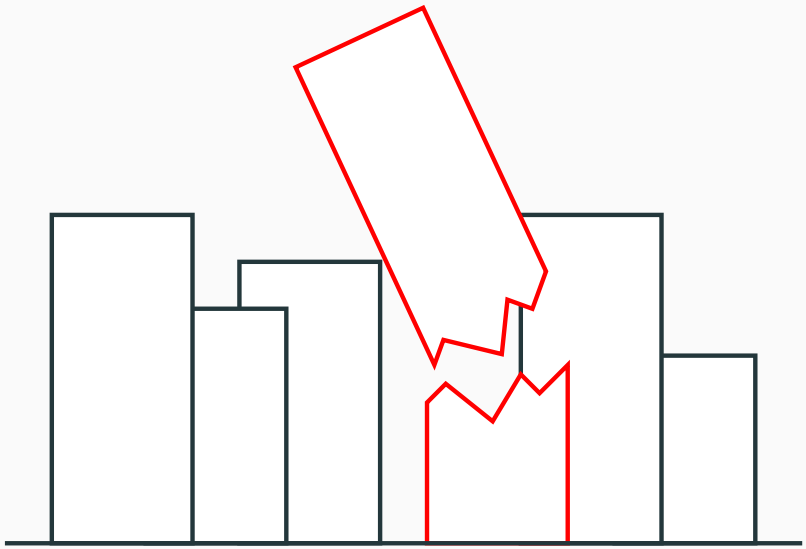
HAUM Talks – Janvier 2016

Contents

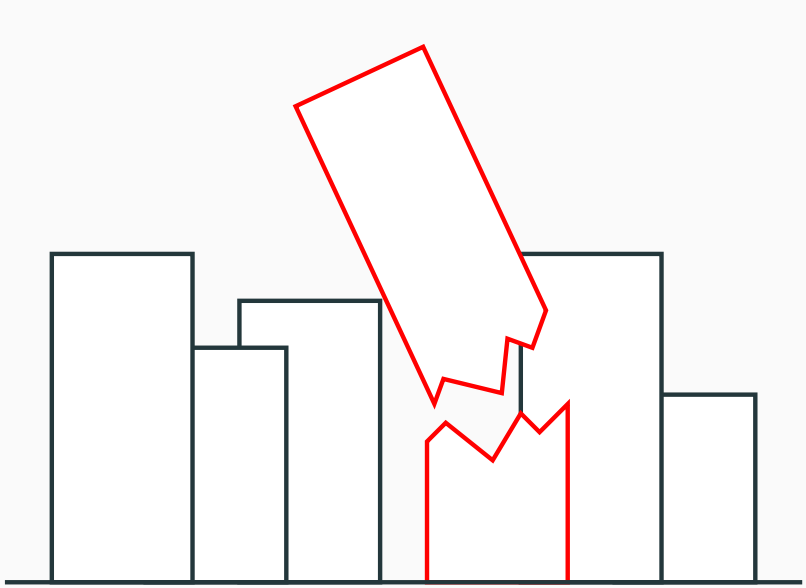
1. Idea
2. Implementation
3. Hypothesis and Follow-up
4. What others do ?

Idea

A new type of reaction to events



A new type of reaction to events

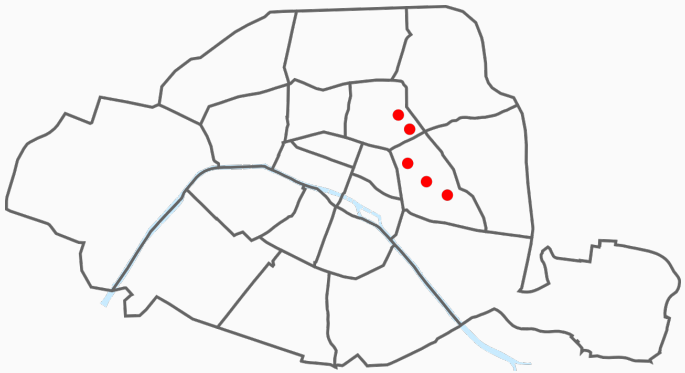


An event triggers a reaction which is public and spontaneous.

Terrorist attacks on November 13th

Special for different reasons :

- unexpected
- lots of witnesses & victim
- impacting a young population ❓
- claimed after the strike ⏩



Terrosist attacks on November 13th

Reactions as important as the violence of the strikes :

- huge amount of messages posted online (connected population)
- reaction of some people that spanned on several hours/days
- heavy cell-phone usage to post ⇒ geotagged data
- wide range of reaction types

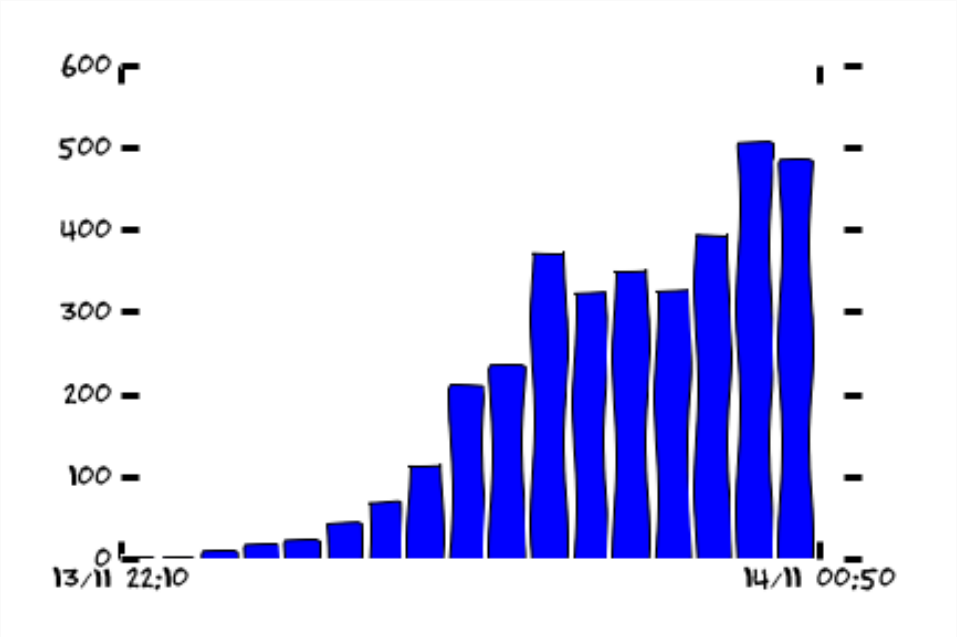


Figure 1: Evolution of the number of tweets posted during the first hours after the strikes

Proposition

Is it possible to extract social characteristic times from the tweets posted after the terrorsit attacks ?

Proposition

Is it possible to extract social characteristic times from the tweets posted after the terrorsit attacks ?

Characteristic time : duration used to discriminate different events and/or evaluate their response velocity.

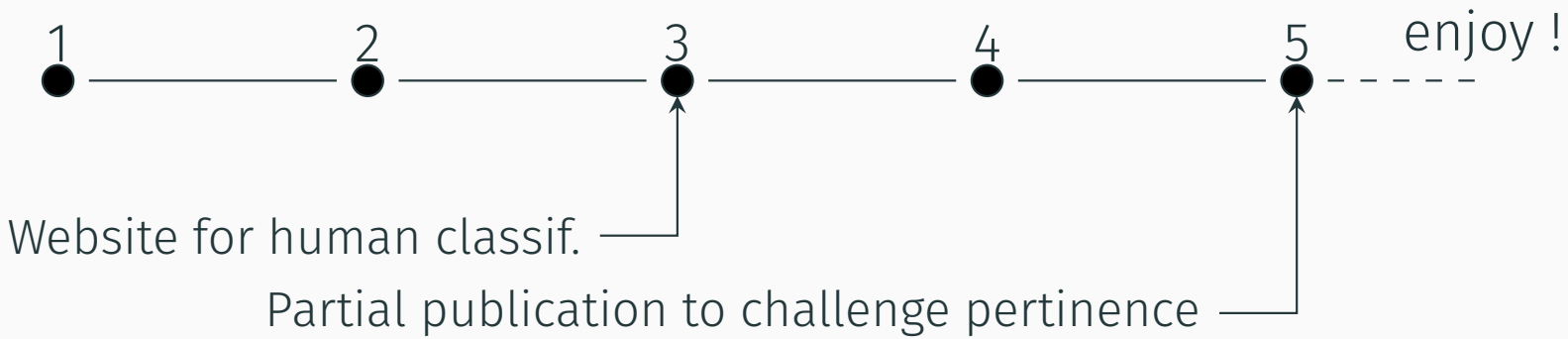
Why tweets ?

- Short format
  - Easy to process
  - Prone to clear and unambiguous reactions ❓
  - Writting with (almost) no thinking
- Easy to fetch
  - Using the native MongoDB format
  - Efficient *off-line* processing
  - Possibility for several passes
- Huge dataset to work with (several tenth of thousands...)

Implementation

Planning

1. Fetch and store tweets for off-line analysis (*ParisAttacks*)
2. Language based auto-classification
3. Manual sentiment-based classification of a subset
4. Auto-classification using the classified subset as a training set
5. Propose observables to deduce social characteristic times



Hypothesis and Follow-up

Hypothesis

Long-term Reactions

Unambiguous

After-claim turn-around

Hypothesis

Long-term Reactions

To be verified checking the ratio of accounts with long-term reactions

Unambiguous

Non dumb hypothesis in a post-traumatic context, sustained by the short format

After-claim turn-around

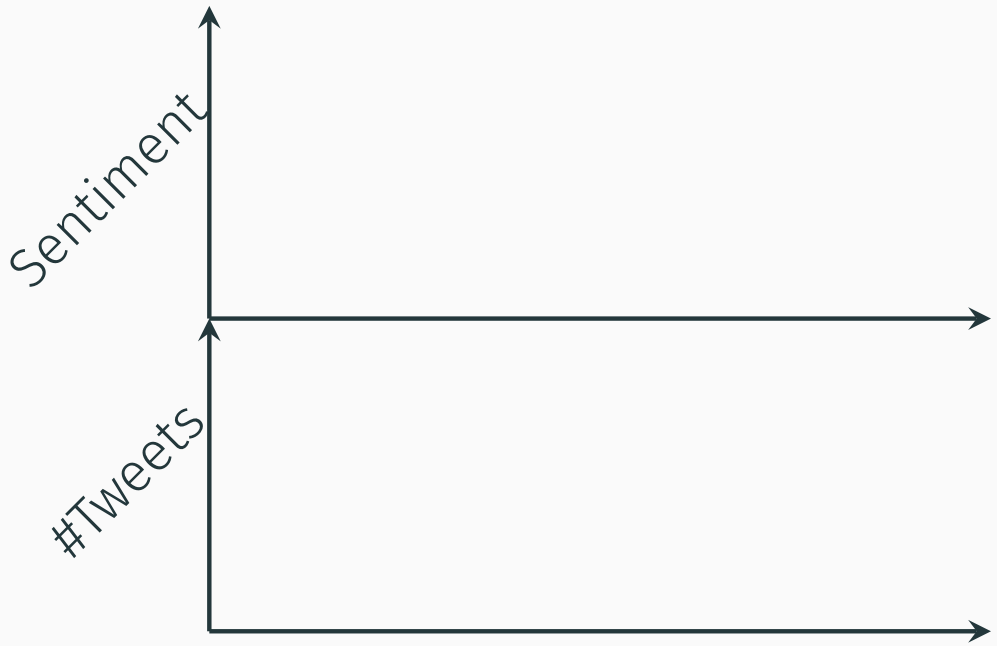
To be verified after analysis... may be we'll need to exclude Daech oriented tweets

Follow-ups

- Use geo-tagging to enhance analysis
- Correlate results with papers/news publication
- Compare with other public events (sports, political sutff, etc...)
- Per-channel analysis :
  - medias
  - politics
  - others...

Follow-ups

- Compare with CharlieHebdo events in January 2015 (coupled reaction hypothesis)
- Explore self-organization modes (PorteOuverte, I'm Safe)
- Self-emulation or response to a stimulus (tweet of a political persron, a media)
- Bi-varying analysis



About stats

- Non representative population sample
- First reactions biased by media annonces (a media always has an opinion)
- Induced bias from the manual classification
- etc...

What others do ?

What others do ?


Michal Lukasik, Trevor Cohn, and Kalina Bontcheva.
Classifying Tweet Level Judgements of Rumours in Social Media.

In *Proceedings of EMNLP*, volume 796, pages 2590–2595, 2015.


Michal Lukasik, Trevor Cohn, and Kalina Bontcheva.
Modeling Tweet Arrival Times using Log-Gaussian Cox Processes.
In *Proceedings of EMNLP*, volume 796, pages 250–255, 2015.


Laurent Luce.
Twitter sentiment analysis using Python and NLTK.


Oriane Piquer-Louis.
Documenter l'expérience habitante de la ville sur les réseaux sociaux : corpus photographiques et données numériques, le sens des collections.

Thank you !

Questions ?

fred@haum.org — mathieu@haum.org