Information Reliability
with Marked Temporal Point Processes

HUMAN-CENTERED MACHINE LEARNING
http://courses.mpi-sws.org/hcml-ws18/
Information gathering is an online activity

People can learn about a wide variety of topics in Wikipedia

People can get answers to their questions in Q&A sites

People can attend MOOCs to learn about a subject

People can be up to date with latest “news”
Opinionated, inaccurate, **false** facts

**Wikipedia: List of hoaxes on Wikipedia**

From Wikipedia, the free encyclopedia

⚠️ Please do not attempt to create new hoaxes on Wikipedia; [here is why.](https://en.wikipedia.org/wiki/Wikipedia/List_of_hoaxes_on_Wikipedia)

---

**The New York Times**

*As Fake News Spreads Lies, More Readers Shrug at the Truth*

*BBC*

'Fake news': What's the best way to tame the beast?

---

**Breitbart**

12 Fake News Stories From The Mainstream Media

---

03:21, 20 September 2016

Barack Hussein Obama II is a Kenyan politician

(Barack Obama’s Wikipedia article)
Solution: Resort to the crowd

Allow users to verify and/or refute information

Challenges

1. Users may be untrustworthy
2. Information may be disputed
3. Users need to verify/refute
Reliability & trustworthiness

Can we quantify info reliability and source trustworthiness?

Why this goal?

Showcase reliable information, fix unreliable information

Identify (un)trustworthy information sources

Increase information quality
Information reliability: key, simple idea

A source is trustworthy if:

- Its contributions are verified more frequently
- Its contributions are refuted more rarely

Challenge

At a time $t$, a document may be disputed

Verifications: rarer
Refutations: more frequent

[Tabibian et al., WWW 2017]
Information reliability: key, simple idea

A source is trustworthy if:

Its contributions are verified more frequently

and/or

Its contributions are refuted more rarely

Over time, each document has a different level of inherent unrealibility

Challenge

At a time t, a document may be disputed

Verifications: rarer

Refutations: more frequent

[Tabibian et al., WWW 2017]
Statement additions
(one process per document)

\[ N_A(t) \]

Statement refutations
(one process per statement)

\[ N_{R1}(t) \quad N_{R2}(t) \quad N_{R3}(t) \]

Barack Obama

Barack Obama: Revision history

05:41, 28 November 2016  Ranze (talk | contribs)  . . (301,105 bytes) (+18) . .
03:32, 28 November 2016  Xin Deui (talk | contribs)  . . (301,087 bytes) (+128) . .
00:57, 28 November 2016  SporkBot (talk | contribs)  m . . (301,155 bytes) (-37)
07:03, 27 November 2016  Saiph121 (talk | contribs)  . . (301,192 bytes) (+25).

Statement:
\[ e = (s, t, \tau) \]

[Tabibian et al., WWW 2017]
Intensity of statement additions

Statement additions (one process per article)

\[ \lambda_d(t) = \sum_{j} \phi_{d,j} k(t - t_j) + \sum_{e_i \in \mathcal{H}_d(t)} w_d^{\top} \gamma_{s_i} g(t - \tau_i) \]

- **Intensity or rate** (Statements per time unit)
- **Article unreliability** (Mixture of Gaussians)
  - Temporal evolution of the\textit{ intrinsic} reliability of the article
- **Effect of past refutations** (topic dependent; topic weight $w_d$)
  - Refuted statements trigger the arrival of new statements to replace them

[Tabibian et al., WWW 2017]
Intensity of statement refutations

Statement additions (one process per article)

Statement refutations (one process per statement)

\[ \mu_i(t) = (1 - N_i(t)) \left[ \sum_j \beta_{d,j} k(t + t_i - t_j) + \mathbf{W}_d^T \alpha_{s_i} \right] \]

- **Refutations happen only once**
- **Article unreliability** (Mixture of Gaussians)
- **Source trustworthiness** (topic dependent; topic weight \( w_d \))

Intensity or rate (Statements per time unit)

- Shared across statements of an article!

Refutations happen only once

The higher the parameter \( \alpha_{s_i} \), the quickest an article gets refuted

[Tabibian et al., WWW 2017]
Model inference from event data

Conditional intensities

\[ \{ \lambda_d(t) \} \quad \{ \mu_i(t) \} \]

Events likelihood

\[
\sum_{d=1}^{|D|} \sum_{i:e_i \in \mathcal{H}_d(T)} \log p(t_i | \mathcal{H}_d(t_i), \phi_d, \{ \gamma_s \}, w_d) + \sum_{d=1}^{|D|} \sum_{i:e_i \in \mathcal{H}_d(T)} \log p(\Delta_i | t_i, \beta_d, \{ \alpha_s \}, w_d)
\]

Theorem. The maximum likelihood problem is convex in the model parameters.

[Tabibian et al., WWW 2017]
Wikipedia dataset

- Complete edit history of Wikipedia up to July, 2014
- 50k web sources (with more than 10 additions)
- who were used in 100k articles (with more than 20 additions)
  - by means of 10.4 million statement additions
  - 9 million statement refutations
- 10 topics

[Tabibian et al., WWW 2017]
What can the model tell us about the article unreliability?
Barack Obama’s biography

Peaks match noteworthy events
Difference between arrival and removal indicates controversy

Democratic nomination
US elections

Inferred intrinsic unreliability

[Tabibian et al., WWW 2017]
What about the trustworthiness of the sources?
Source trustworthiness

Probability of refutation within 6 months in a stable Wikipedia article

[Tabibian et al., WWW 2017]
Demo for other pages and sources