

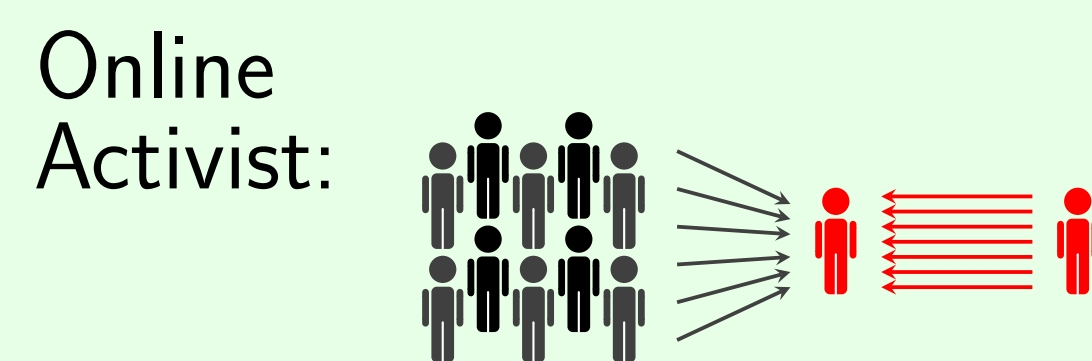
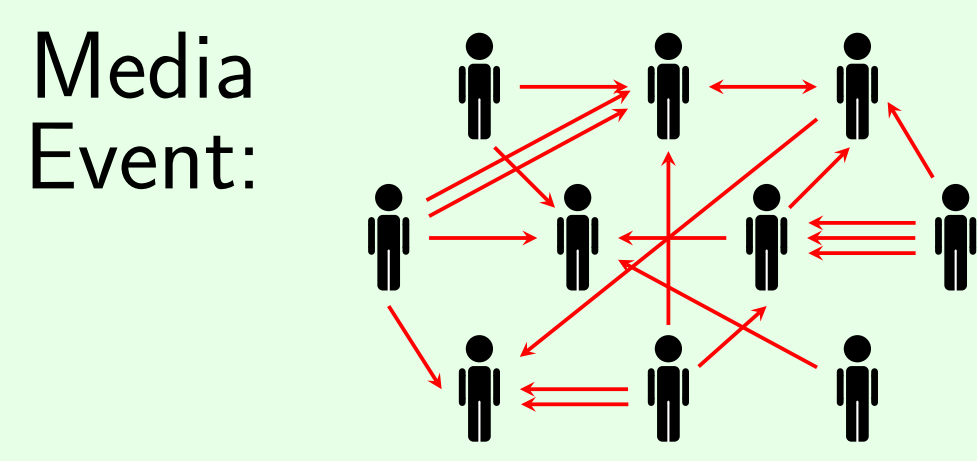
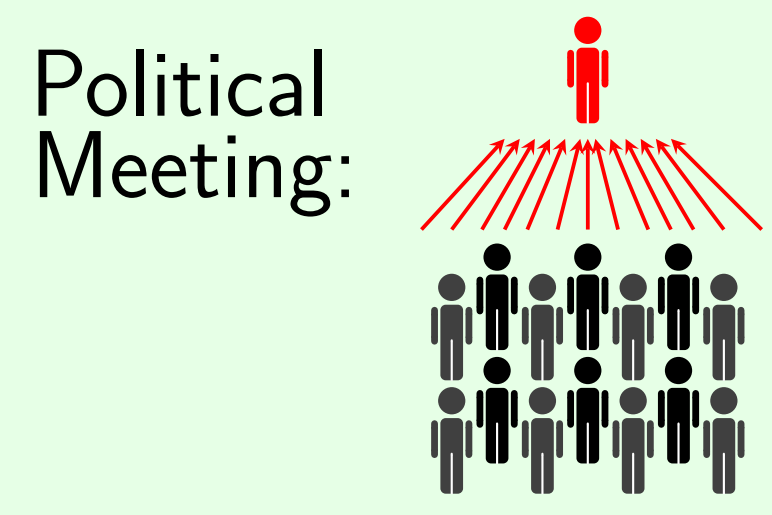
A MULTIDIMENSIONAL METHOD TO FIND IRREGULAR PATTERNS IN INTERACTION DATA

Audrey Wilmet, Robin Lamarche-Perrin - first.lastname@lip6.fr

LIP6 - Sorbonne Université - Complex Network Team, France

OBJECTIVES AND APPROACH

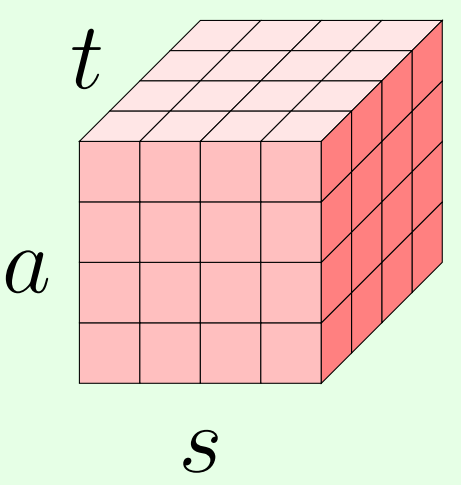
Goal: find unexpected behaviors related to political communication via Twitter.



An Interaction on Twitter = a Retweet

→ a **spreader** s retweets an **author** a at time t

Set of retweets = $\{(s, a, t)\} \in V \times V \times T =$ **Data Cube**

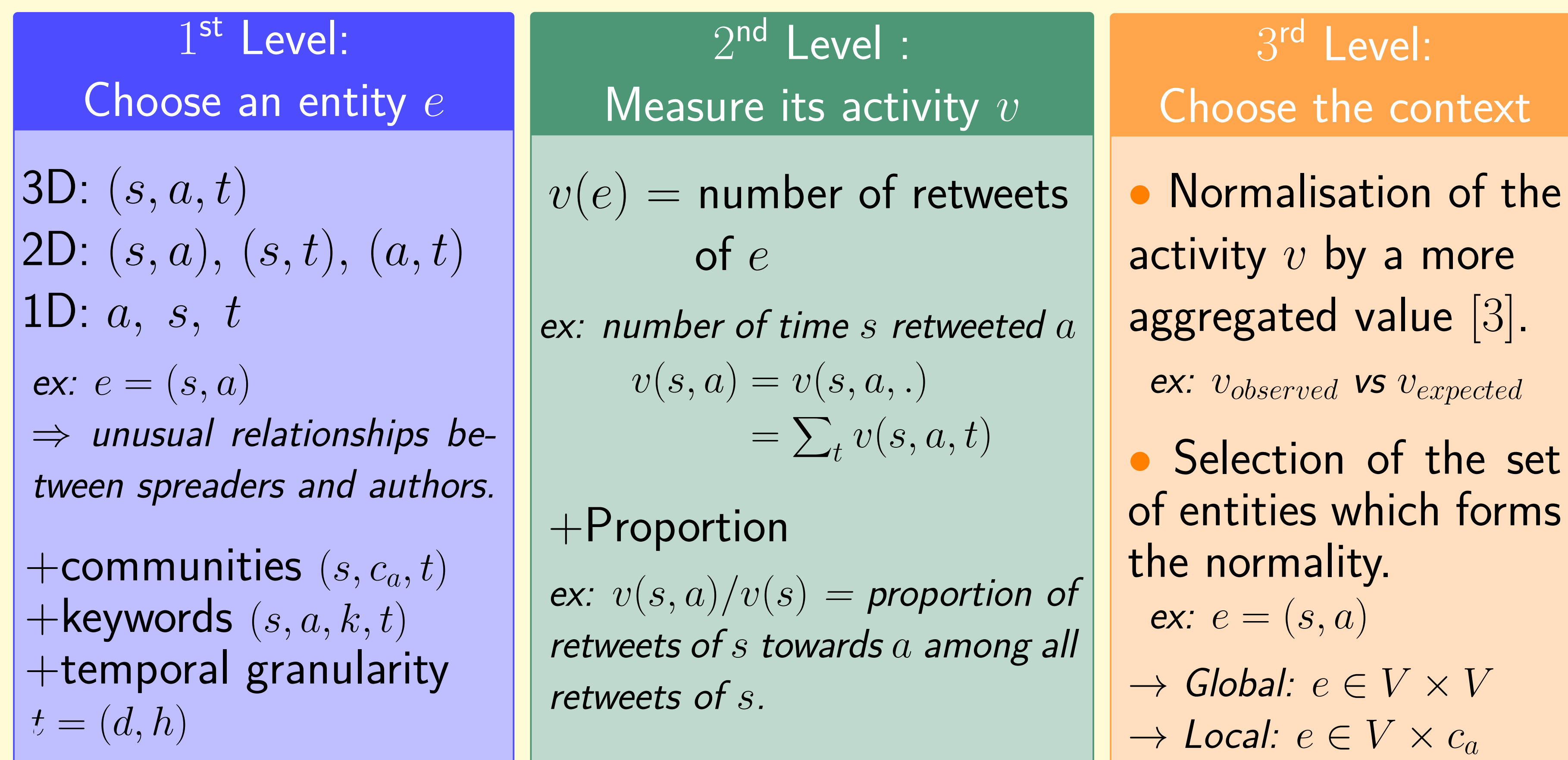


Politoscope Dataset: retweets sampled from political accounts and specific keywords during August 2016 [2].

DEFINE WHAT AN IRREGULAR PATTERN IS

In temporal interaction data, irregular patterns can be defined in an almost unlimited number of ways by differently **combining time and structure at different levels**.

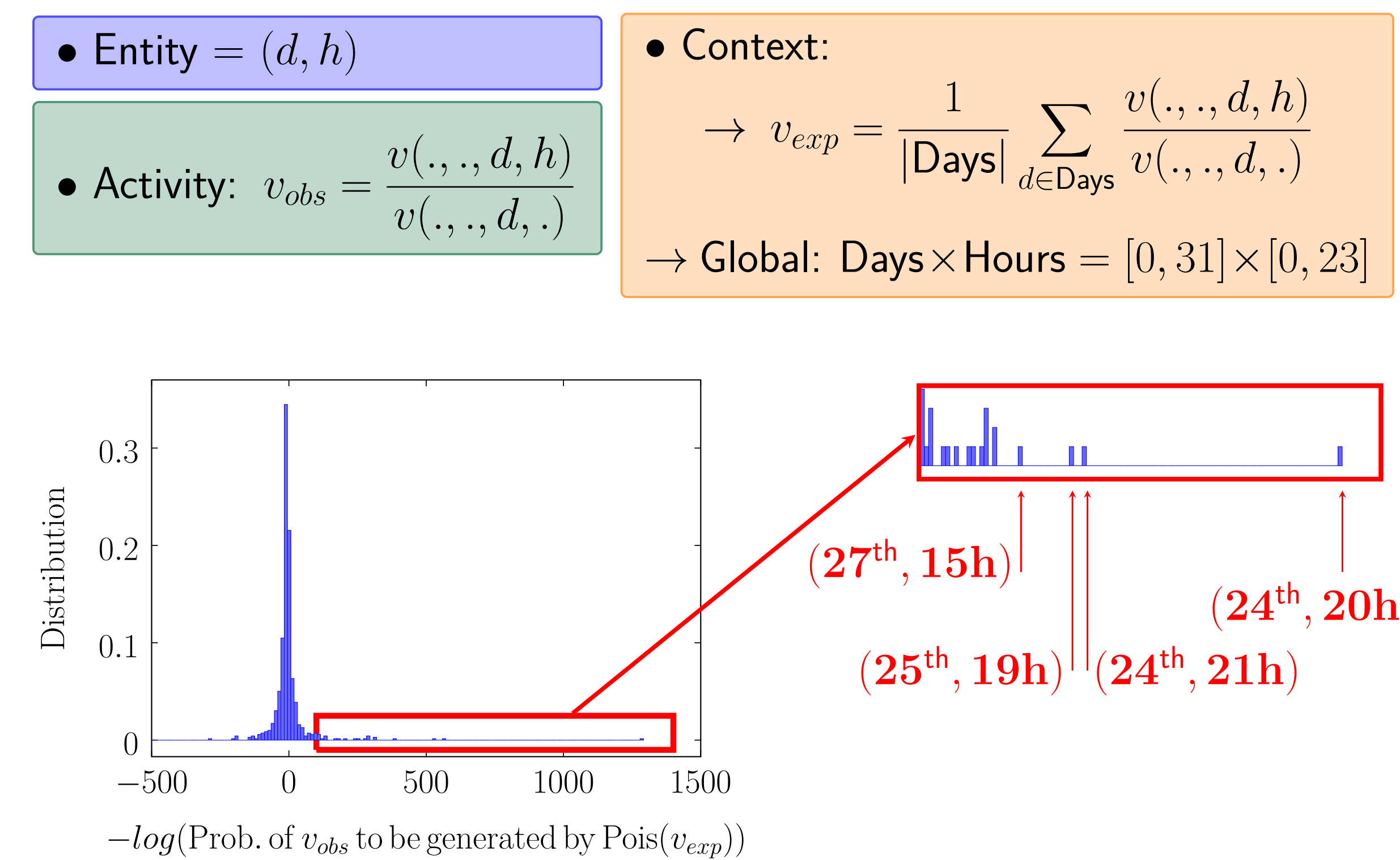
An Irregular Pattern = an **entity** which statistically deviates from others [1]



CASE STUDY

1ST STEP: FIND ABNORMAL HOURS

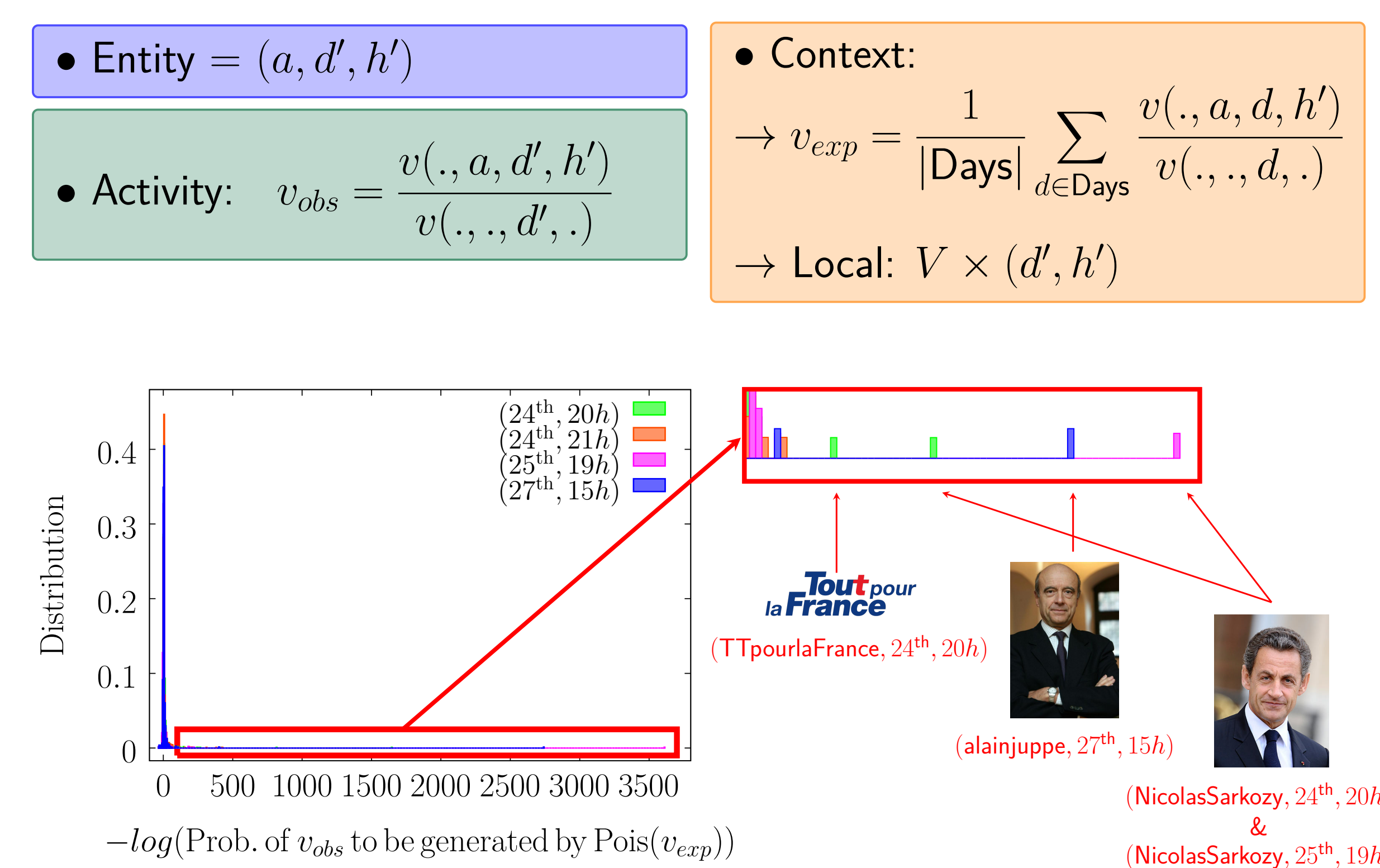
⇒ **couples** $t = (d, h)$ during which the proportion of retweets observed during hour h of day d among all retweets of day d is higher than the expected proportion during hour h on all other days, compared to all other couples (d, h) .



ex: On the 24th of August, there has been much more retweets from 20:00 to 21:00 than there usually are during this hour when looking at other days.

2ND STEP: FIND WHO IS RESPONSABLE FOR IT

⇒ **triplets** (a, d', h') for which the proportion of retweets directed towards a during hour h' of day d' among all retweets of day d' is higher than the expected proportion for this author during this hour on all other days, for all authors during a previously found abnormal hour (d', h') .



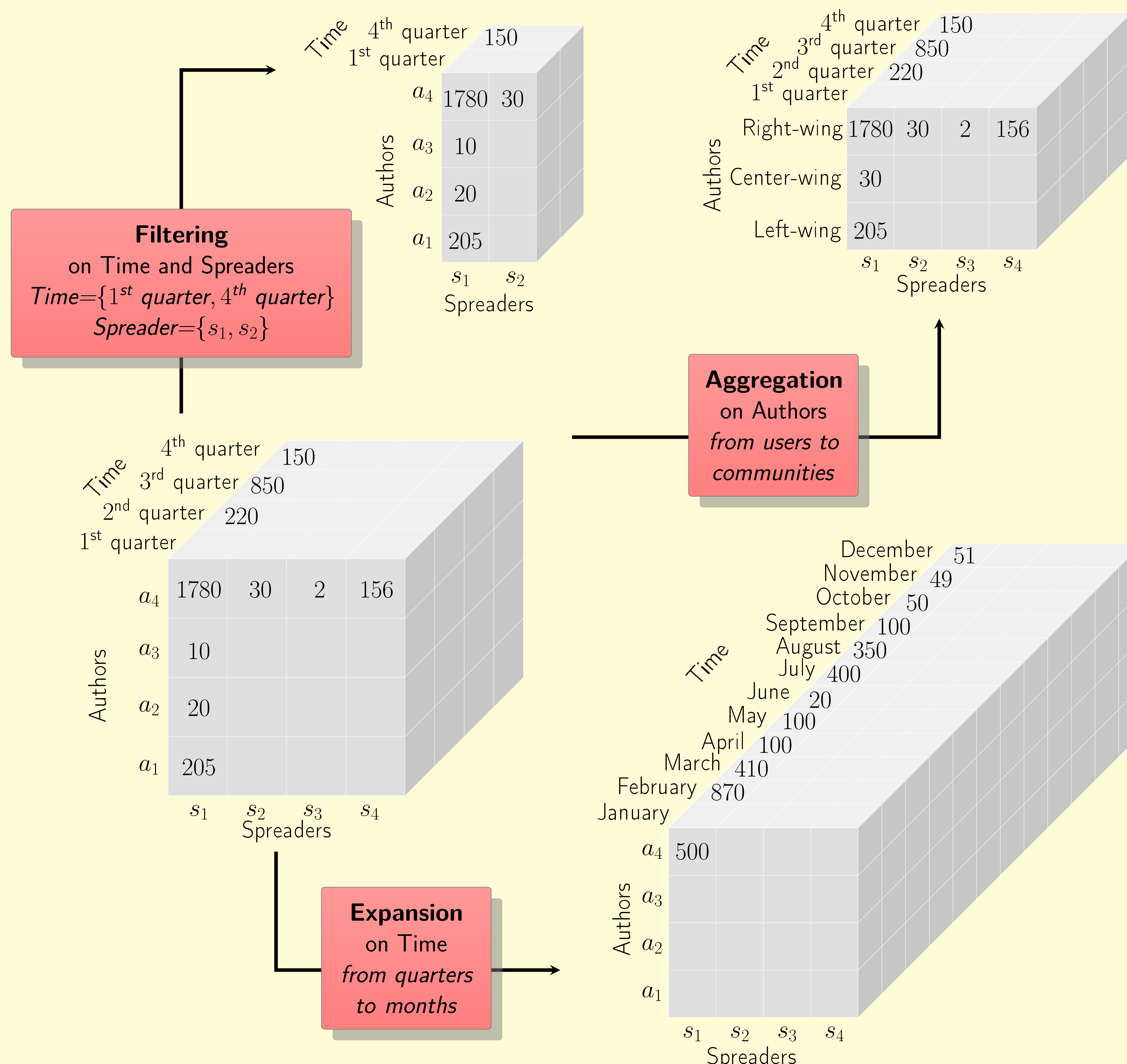
ex: On the 25th of August from 19:00 to 20:00, Nicolas Sarkozy has been much more retweeted than he is usually during this hour when looking at other days.

RESULTS COMPARISON WITH NEWS RECORDS

Leading authors : – (25th, 19h): Political Meeting of Nicolas Sarkozy.
– (27th, 15h): Political Meeting of Alain Juppé.
– (24th, 20h): Interview of Nicolas Sarkozy on television news. His campaign slogan is "Tout pout la France".

No leading authors: – (24th, 21h): Public reaction to Nicolas Sarkozy's interview that took place one hour before.

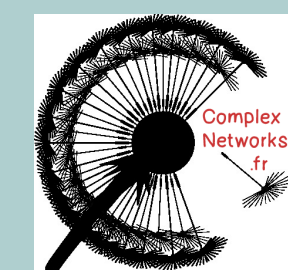
PERFORM OPERATIONS ON THE DATA CUBE



See Refs [3, 4].

REFERENCES

- [1] C. Aggarwal. *Outlier Analysis*. Springer International Publishing, 2016.
- [2] N. Gaumont, M. Panahi, and D. Chavalarias. Methods for the reconstruction of the socio-semantic dynamics of political activist Twitter networks: Application to the 2017 French Presidential elections. [hal-01575456v2](https://arxiv.org/abs/1615.05662), 2017.
- [3] C. Grasland, R. Lamarche-Perrin, B. Loveluck, and H. Pecout. International agenda-setting, the media and geography: A multi-dimensional analysis of news flows. *Espace géographique (English Edition)*, 45(1):1–18, 2016.
- [4] J. Han, J. Pei, and M. Kamber. *Data mining: concepts and techniques*. Elsevier, 2011.



ACKNOWLEDGEMENTS

This work is funded in part by the European Commission H2020 FETPROACT 2016-2017 program under grant 732942 (ODYCCEUS), by the ANR (French National Agency of Research) under grants ANR-15-CE38-0001 (AlgoDiv).

CONTACT US

E-mail: audrey.wilmet@lip6.fr
Web: www.complexnetworks.fr