Multidimensional and Multilevel Method to Find Irregular Patterns in Interaction Data

Audrey Wilmet[1], Robin Lamarche-Perrin[2]

[1]Laboratoire d'informatique de Paris 6 (LIP6) - Complex Networks Team http://www.complexnetworks.fr/
[2]Institut des Systèmes Complexes Paris Île de France (ISC-PIF)

EUSN'17, Mainz, September 28, 2017

Introduction

Interaction Data

• **Set of interactions:** set of triplets (u, v, t).
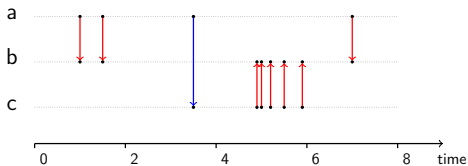
*a interacted with c
at time $t = 3.5$.*

• Goal: find irregular patterns in interactions,
$\rightarrow$ global high activity of a node, **u**,
$\rightarrow$ sharp variation in a node's activity, **(u,t)**,
$\rightarrow$ high activity between two nodes, **(u,v)**,
$\rightarrow$ high activity at a particular moment, **t**.

• Need of a framework to study interactions:
$\rightarrow$ modelling of interaction data as a cube,
$\rightarrow$ definition of 3 multilevel operations to study the cube.

Political communication on Twitter

- Interactions on twitter $=$ Retweet
- $\rightarrow$ *spreader* has retweeted *author* at hour h of day d
- $\rightarrow$ (s, a, d, h).

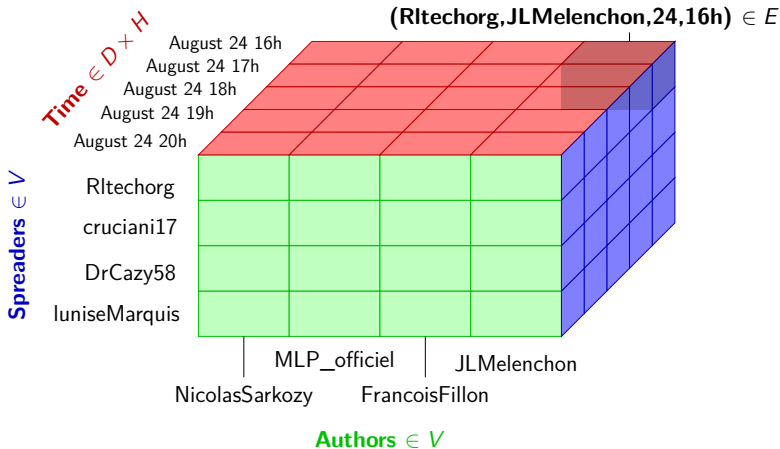- Retweets : sampled from political accounts and keywords.

- *Example*

$(s, a, d, h) = $ (EmmanuelMacron, Elysee, August 29, 10h)

Cube representation of (s, a, d, h)

$\rightarrow$ August '16: $|V| = 211, 155$, $|E| = 990, 005$, $D = [1, 31]$, $H = [0, 23]$.

$\rightarrow$ **irregular patterns**: **political leaders, media events, etc.**
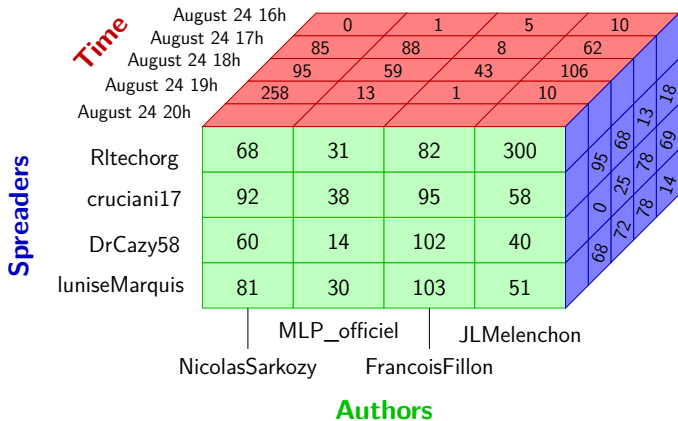
Information

Local Information

Entity: (s, a, d, h)

- Activity function $v(s, a, d, h)$
$\Rightarrow$ number of retweets of a by s during hour h of day d.

Global Information: $v(., a, d, h) = \sum_{s \in V} v(s, a, d, h)$

Entity: (a, d, h)

Global Information:

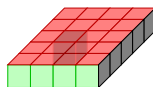$$v(s, ., ., .) = \sum_{a \in V} \sum_{d \in D} \sum_{h \in H} v(s, a, d, h)$$
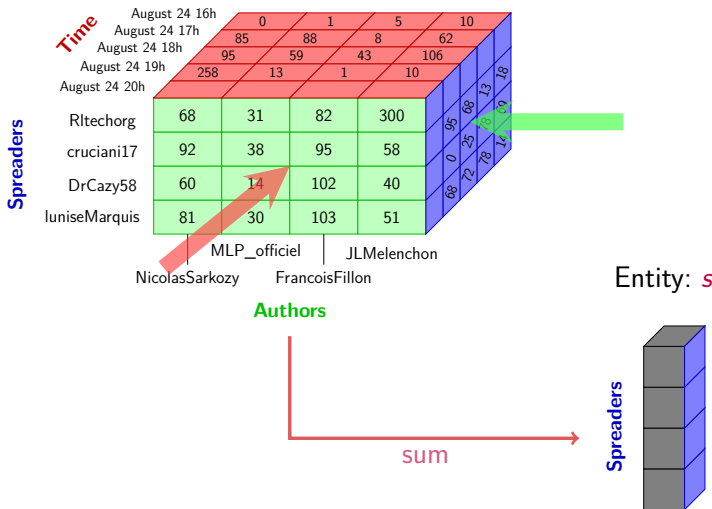
Entity: s

Normalisation

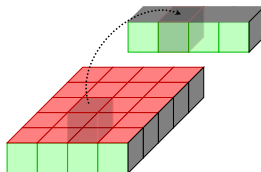## Normalisation

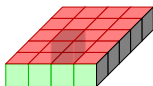[Grasland *et al.*, 2014], [Geomedia]

Evaluate the quantity of interaction of an entity with respect to another more global value.

$\rightarrow$ An aggregated value: $v_{obs} = v(., a, d, h)/v(., a, ., .)$,
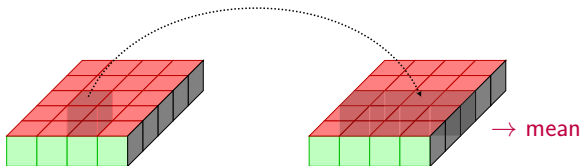
Normalisation

[Grasland *et al.*, 2014], [Geomedia]

Evaluate the quantity of interaction of an entity with respect to another more global value.

$\rightarrow$ An aggregated value,

$\rightarrow$ An expected value:

$v(., a, d, h)$ vs v_{exp} = mean of $v(., a, d, h)$ on a,

$\rightarrow$ mean

Normalisation

[Grasland *et al.*, 2014], [Geomedia]

Evaluate the quantity of interaction of an entity with respect to another more global value.
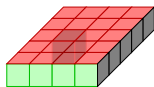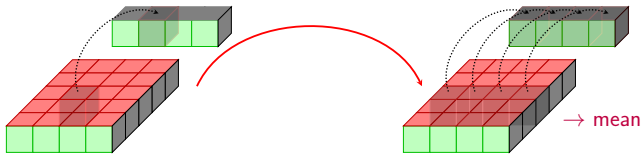
$\rightarrow$ An aggregated value: $v_{obs} = v(., a, d, h)/v(., a, d, .)$,

$\rightarrow$ Its expected value: $v(., a, d, h)$ vs v_{exp},

$\rightarrow$ Both: $v_{obs} = v(., a, d, h)/v(., a, d, .)$ vs v_{exp}.

$\rightarrow$ mean

Audrey Wilmet, Robin Lamarche-Perrin

Comparison

Global Comparison
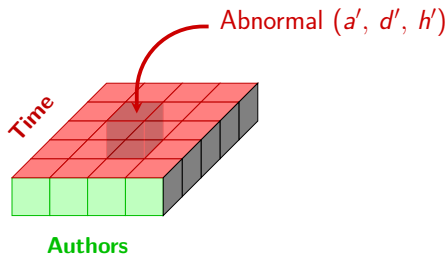
• Principle: comparing the quantity of interactions of an entity to the quantities of interactions obtained by all other entities of the same type.

Example:
Activity function: $v(., a, d, h)$
Entity: (a, d, h)
Compared Entities: (a, d, h) $\quad \forall a \in V, d \in D$ and $h \in H$

Abnormal (a', d', h')

Time

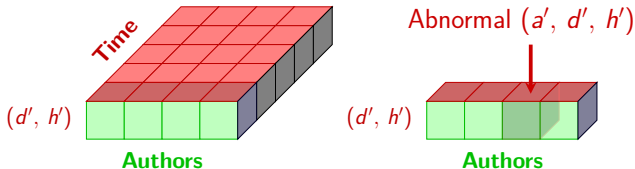**Authors**

Local Comparison

- Principle: fixing a dimension in an entity and comparing its value to all others obtained by varying the remaining variable(s).

Example:
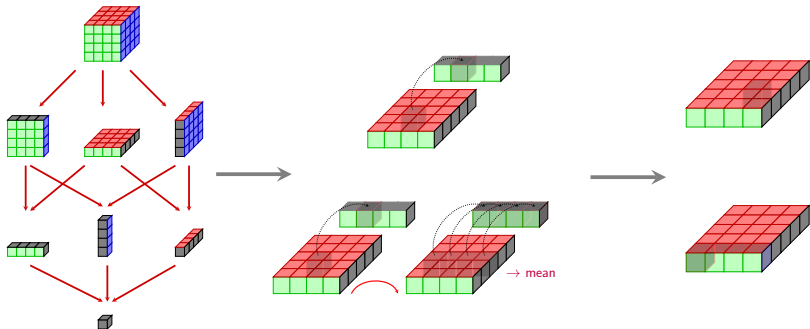Activity function: $v(., a, d, h)$
Entity: (a, d, h)
Compared Entities: $(a, d', h'), \forall a \in V$

Abnormal (a', d', h')

(d', h') **Time** **Authors**

(d', h') **Authors**

Multilevel & Multidimensional Method

To find irregular patterns:

– 1 st step: choice of an entity and measure of the quantity of interaction (information),

– 2 nd step (optional): normalisation of the quantity of interaction,

– 3 rd step: choice of the comparison.

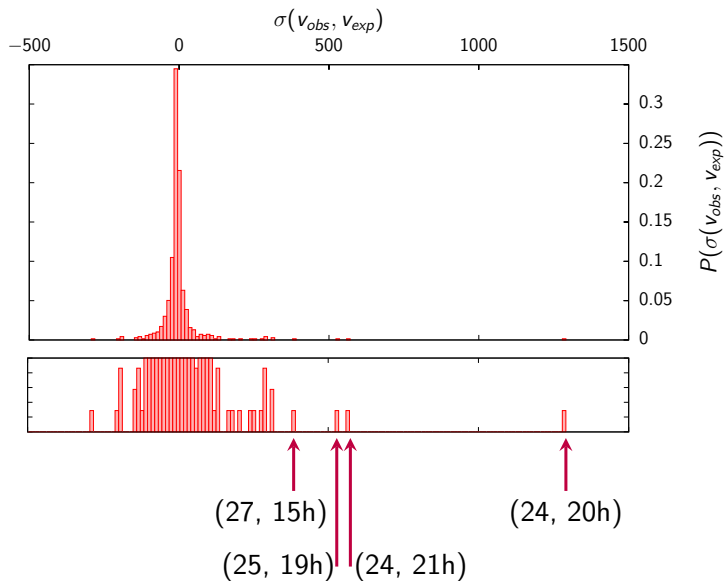# Irregular Patterns

Abnormal (d, h) - **Global Comparison**
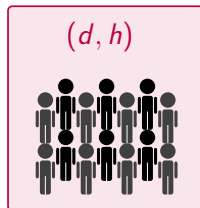
Entity: (d, h),
Information: $v(., ., d, h)$,
Normalisation: $v(., ., d, h)/v(., ., d, .)$ vs v_{exp},
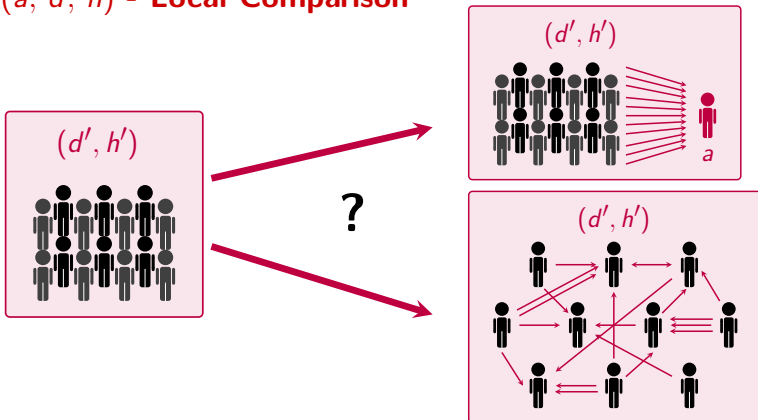Comparison: Global.

Audrey Wilmet, Robin Lamarche-Perrin

Abnormal (d, h) - Results

$(27, 15h)$

$(25, 19h)$ $(24, 21h)$

$(24, 20h)$

(a, d, h) - **Local Comparison**

(d', h')

?

(d', h')

a

(d', h')

Entity: (a, d, h),
Normalisation: $v(., a, d, h)/v(., ., d, h)$ vs v_{exp},
Comparison: (d', h') fixed.

(d', h')

Audrey Wilmet, Robin Lamarche-Perrin

Abnormal (a, d, h) - Results
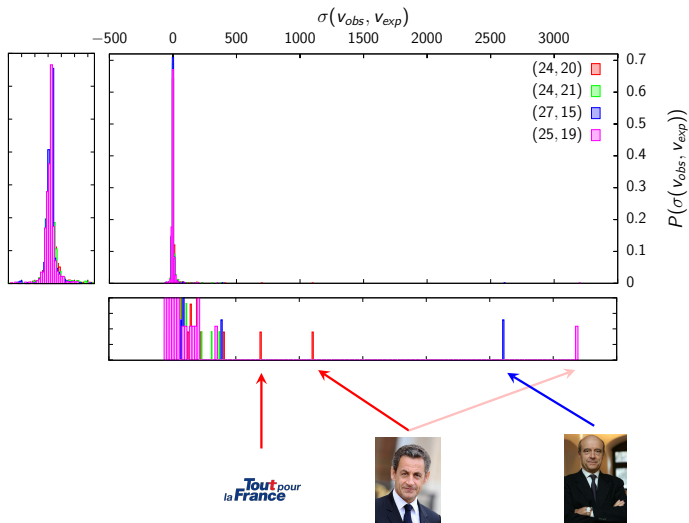
Abnormal (a, d, h) - **Analysis**

Other approach

(d', h')

a' s
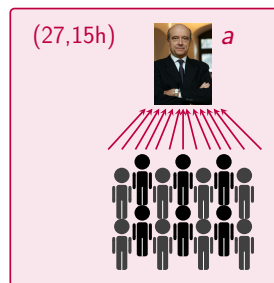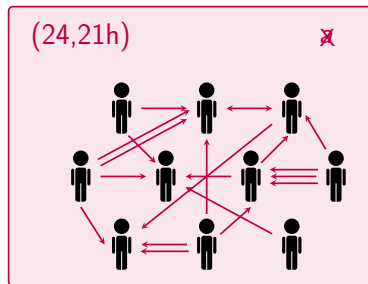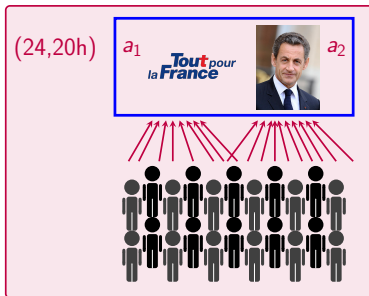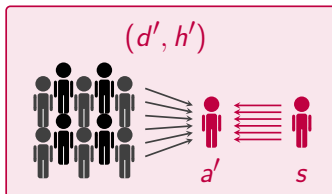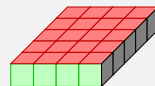
Entity: (a, d, h),
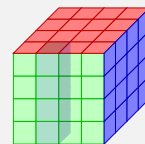Normalisation: $v(., a, d, h)/v(., ., d, h)$ vs v_{exp},
Comparison: Global.

Entity: (s, a, d, h),
Normalisation: $v(s, a, d, h)/v(., a, d, h)$ vs v_{exp},
Comparison: Local, (a', d', h') fixed.

(a', d', h')

Conclusion

• Multidimensional and multilevel method:

– numerous ways to explore data,

– $\neq$ types of irregular patterns: dimensions, normalities,

– Aggregate accounts belonging to the same │ user category,
                                           │ political groups.

$\implies$ difficulties in finding a relevant study approach.

Thanks for your attention !
Questions

Acknowledgement