# Tie Strength Persistence and Transformation

Michele A. Brandão, Pedro O. S. Vaz de Melo & **Mirella M. Moro**
Universidade Federal de Minas Gerais - UFMG, Brazil

WIC@2017

# Introduction

## Context

- Social networks $\longrightarrow$ interactions among individuals
- Their models and patterns allow to solve different problems
- Time: fundamental aspect


- Social ties: the strength of ties
- Tie strength persistence and transformation → *time*

# Relevance

- Relevant knowledge
- Different applications
- Academic context
  - Ranking
  - Productivity analysis

# Example
# Dengue @ Brazil

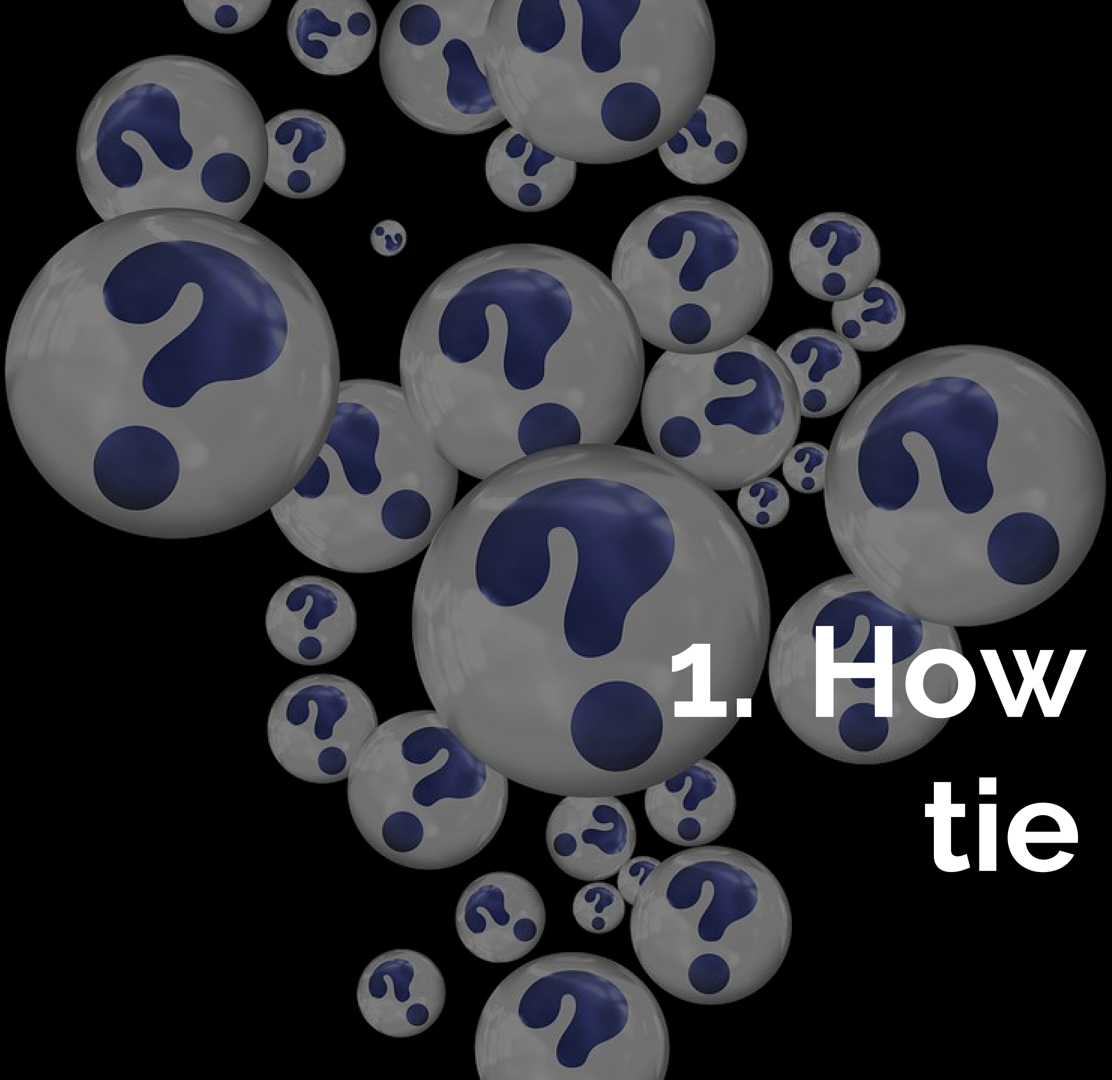
Mining over Twitter data

Outbreaks detection

Predict future outbreaks

Plan properly

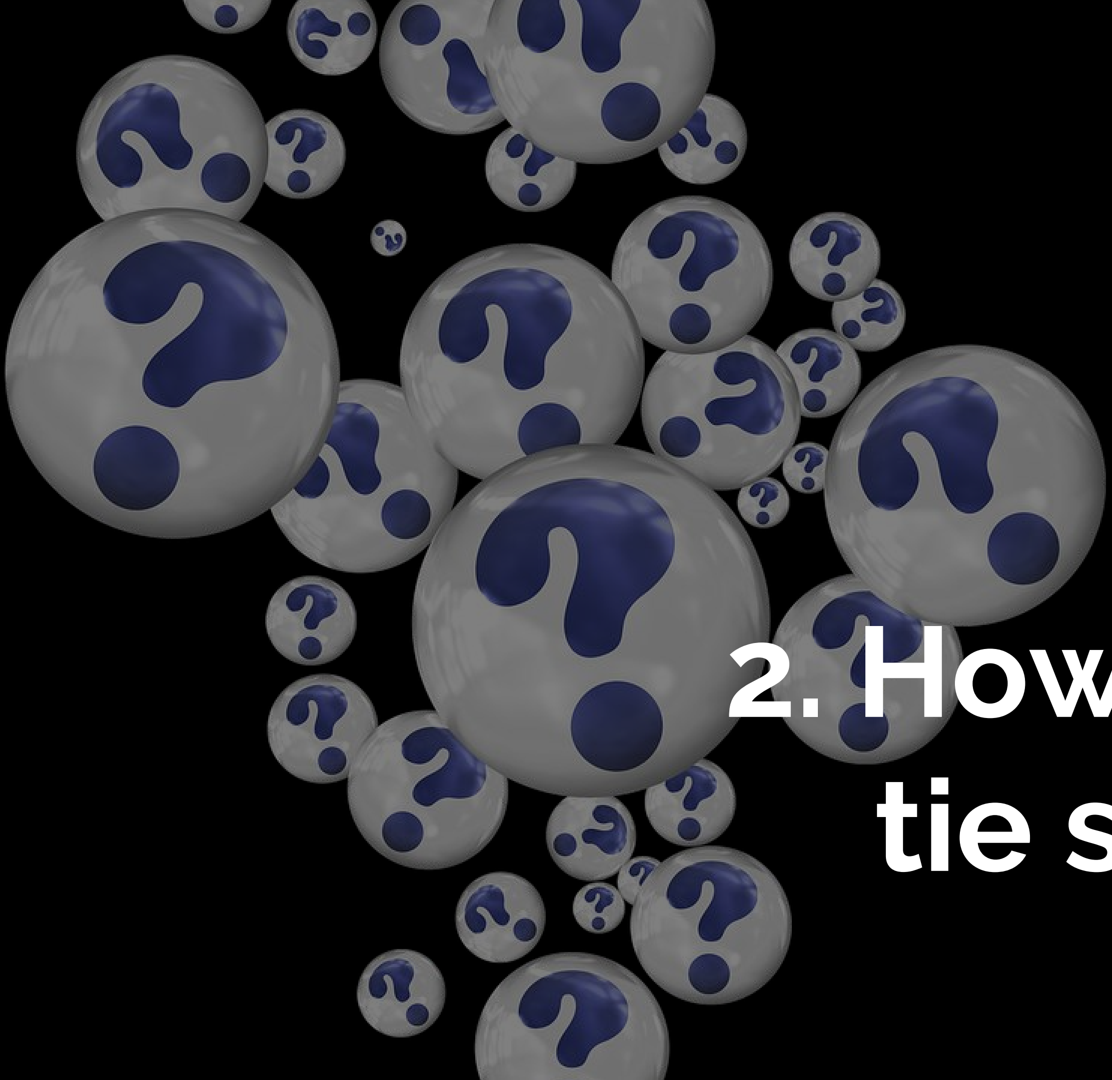

pixabay.com



# 1. How to measure tie strength?

"Tie strength may be measured by a combination of the amount of time, the cooperation intensity and the reciprocal services that characterize the tie "

[Granovetter, 1973]



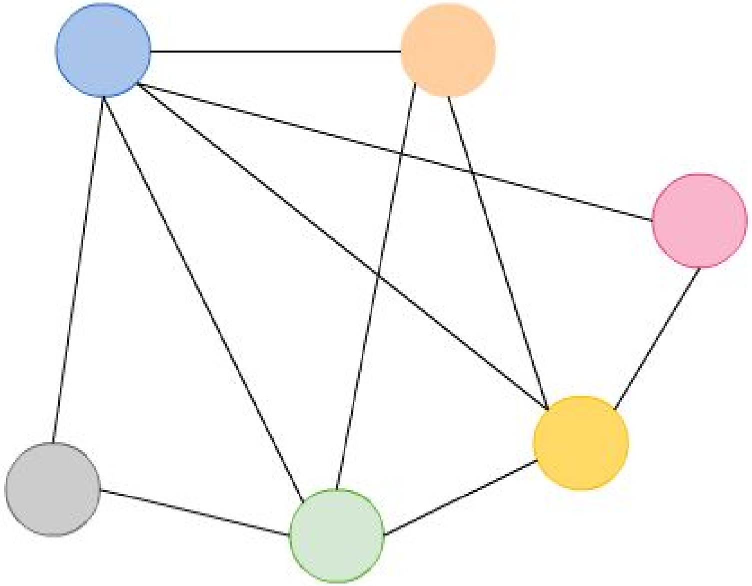
# 2. How to analyze tie strength?

# Goals

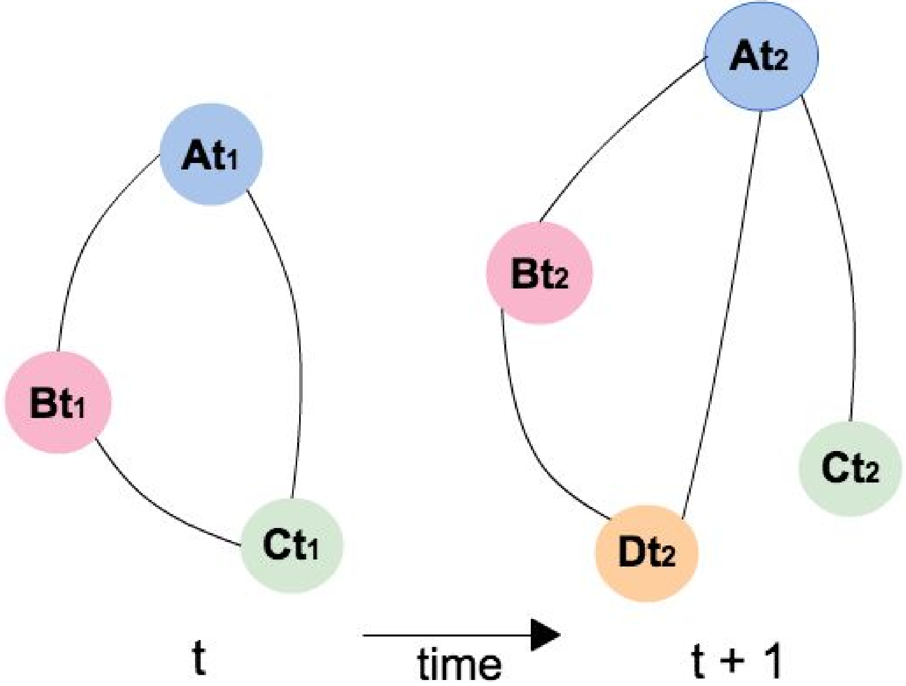
1. Formalize tie strength + new algorithm (called fast-RECAST) to measure it in large SN in acceptable run-time
2. Tie strength dynamism analyses over co-authorship networks of different areas
   a. **Link persistence** analyses: most co-authorships are symbiotic
   b. **Link transformation** analyses: in a long period of time, most ties tend to disappear

# Background

## Static aggregated graph



## Dynamic graph



# Background

## Tie Strength

- What is a **weak** tie?
- A **strong** tie?

# Measuring Tie Strength

## RECAST

Random rElationship ClASsifer sTrategy

- Random graphs
- Two features
  - Neighborhood overlap
  - Edge persistence
- Small networks

# RECAST

Random rElationship ClASsifer sTrategy

Four classes

- Friends (strong)
- Acquaintances (weak)
- Bridges
- Random

# fast-RECAST

- Build more than one random graph at a time
- Compute edge persistence and topological overlap in parallel
- Optimize memory use

**Algorithm 1** Multiprocessing RECAST (fast-RECAST): a parallelized code to classify edges of $G_t$ as random or social – strong, weak or bridge.

**Require:** $p_{rnd} \geq 0$

**return** $class(i,j)\ \forall (i,j) \in U_t E_t$

Construct $G_t^R$ and set $\mathbf{RND}(G_1), \ldots, \mathbf{RND}(G_t)$ using **T−RND** with different processes for each $\mathrm{RND}(G_k)$ (**pool.map_async**)

Get $\overline{F}_{to}(x)$ and $\overline{F}_{per}(x)$ from $G_t^R$ using a two-dimensional size mutable data structure (**pandas dataframe**)

Get $\overline{x}_{to} | \overline{F}_{to}(\overline{x}_{to})$ and $\overline{x}_{per} | \overline{F}_{per}(\overline{x}_{per}) = p_{rnd}$ with a process for each network feature (**pool.apply_async**)

**for all** edges$(i,j) \in E_t$ **do**

    **if** $per(i,j) > \overline{x}_{per}$ and $to(i,j) > \overline{x}_{to}$ **then**

        $class(i,j) \longleftarrow Strong$

    **else if** $per(i,j) > \overline{x}_{per}$ and $to(i,j) \leq \overline{x}_{to}$ **then**

        $class(i,j) \longleftarrow Bridges$

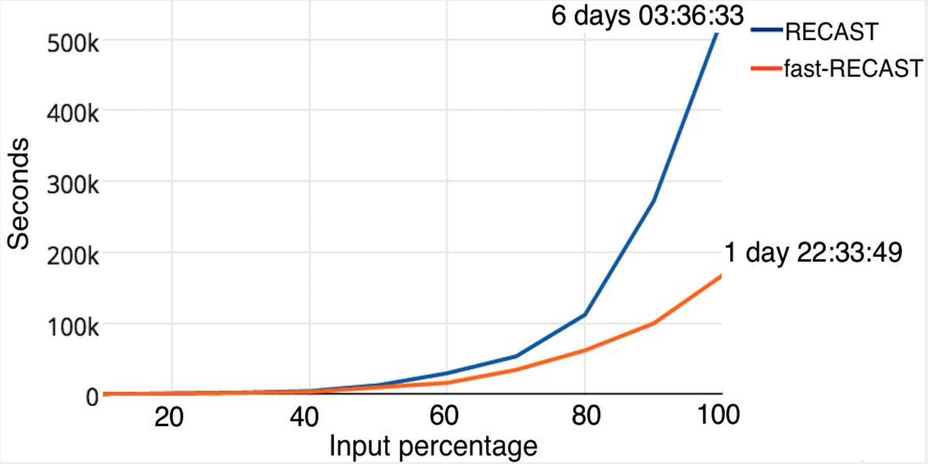
    **else if** $per(i,j) \leq \overline{x}_{per}$ and $to(i,j) > \overline{x}_{to}$ **then**

        $class(i,j) \longleftarrow Weak$

    **else**

        $class(i,j) \longleftarrow Random$

# fast-RECAST




For PubMed dataset

# Experiments and Results

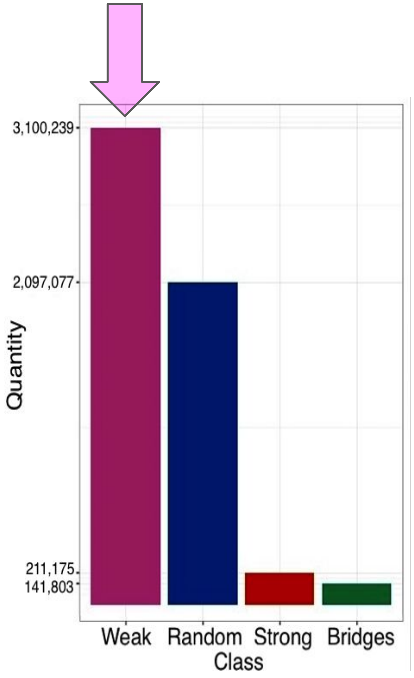
## Datasets

## Co-authorship Networks

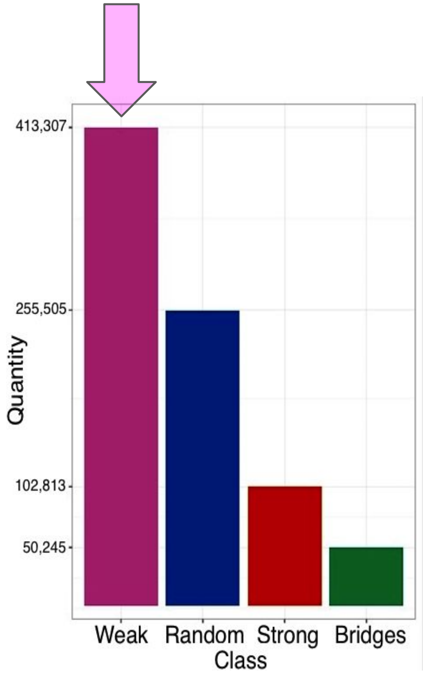
| Dataset | #nodes | #edges |
|---|---|---|
| DBLP Articles | 837,583 | 2,935,590 |
| DBLP Inproceedings | 945,297 | 3,760,247 |
| PubMed | 443,784 | 5,550,294 |
| APS | 180,718 | 821,870 |

# fast-RECAST

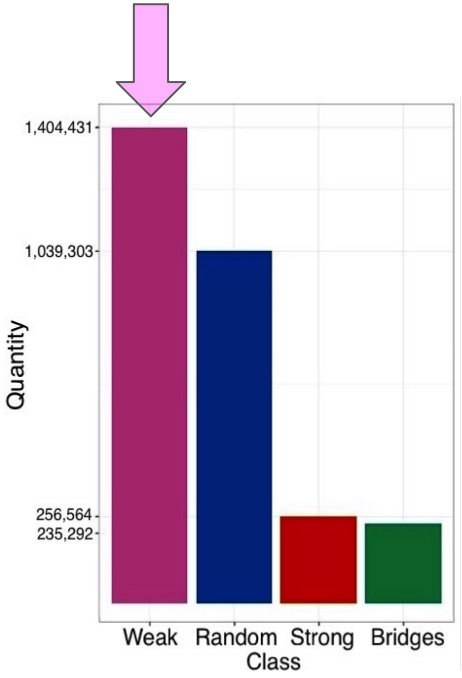
## Relationship classes: quantity per class



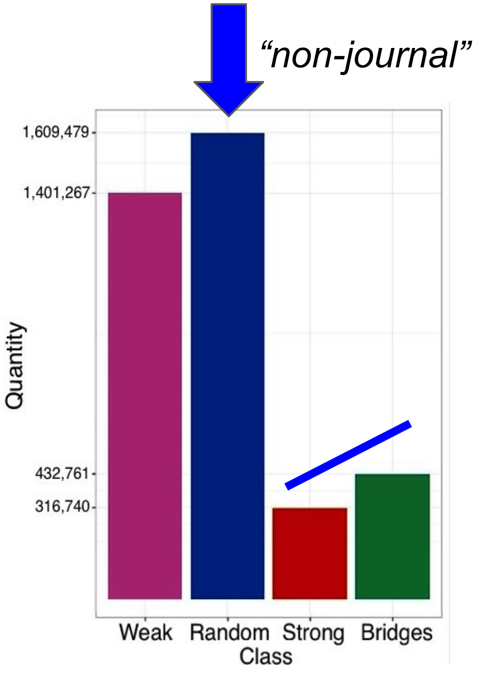

PubMed




APS




DBLP Articles




DBLP Inproceedings

# Tie Strength

Varying over Time

## Tie persistence

- Two time windows: 80%-20% & 70%-30%
- Strong ties and bridges persist more than weak and random ones

# Tie Strength

## Varying over Time

Tie transformation

- Two time windows of 50% of the timestamp
- Most ties tend to disappear

# Main Conclusion

Fast-RECAST can be applied to

- Automatically detect relationship classes
- Differentiate random from social relationships

# Future Work

- Investigate the discovered patterns
- Add other social network features to fast-RECAST
- Expand the study to other collaboration social networks

# Tie Strength Persistence and Transformation

Michele A. Brandão, Pedro O. S. Vaz de Melo & **Mirella M. Moro**
{micheleabrandao, olmo, **mirella**}@dcc.ufmg.br

Research funded by FAPEMIG CAPES CNPq Conselho Nacional de Desenvolvimento Científico e Tecnológico