

Using Topological Properties to Measure the Strength of Co-authorship Ties

Michele A. Brandão, Matheus A. Diniz & Mirella M. Moro
Universidade Federal de Minas Gerais

BRASNAM@2016

Table of Contents

1. Introduction
 - 1.1. Concepts
 - 1.2. Research Questions
2. Related Work
3. Measuring the Strength of Ties: A New Metric
 - 3.1. Problems with NO and Weight
 - 3.2. Tieness: A New Metric
 - 3.3. Publications' Quality
4. Conclusions and Next Steps

Introduction

- Ties strength has been explored in different contexts
- Analyzing how strong a tie is → Investigating
 - different roles of relationships + impact at micro-macro level + influences in the patterns of communication
- Studying the strength of co-authorship ties

Introduction

- How may the tie strength be measured?
 - Amount of time, cooperation intensity, reciprocal services
 - **Neighborhood overlap** (Jaccard index)
 - **Weight** – the intensity of co-authorship between researchers

Introduction

Neighborhood Overlap

Weight

Advantages

- | | |
|--|---|
| <ul style="list-style-type: none">■ Its simplicity to compute■ The possibility to identify if ties are bridges or not■ The consideration of the neighbors to calculate the ties strength | <ul style="list-style-type: none">■ The simplicity to calculate such topological property■ The representation of the exact frequency of collaboration between ties |
|--|---|

Disadvantages

- We find a few problems in both metrics that complicate their sole use to measure the strength of co-authorship ties

Research Questions

- RQ1: How can we measure the strength of co-authorship ties?
 - Neighborhood overlap, weight and **tieness**
- RQ2: What is the relationship between the strength of ties and publications' quality?
 - Qualis rankings

Related Work

- Works that considers the intensity of interactions:

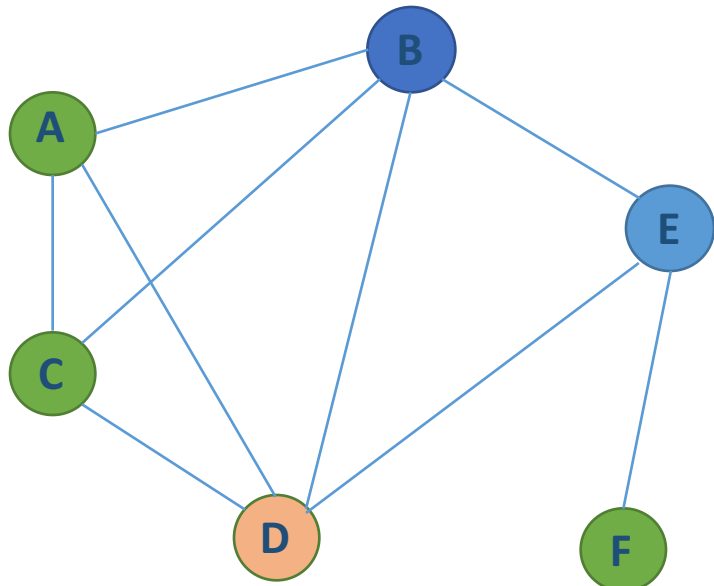
Gilbert, E. and Karahalios, K. (2009). Predicting tie strength with social media. In *Procs. of SIGCHI*, pages 211-220 (**Interactions history**)

Wiese et al., J. (2015). You never call, you never write: Call and sms logs do not always indicate tie strength. In *Procs. of CSCW*, pages 765-774 (**Problems with interactions history**)

Zignani et al., M. (2016). Predicting the link strength of "newborn" links. In *Procs. of WWW*, pages 147-148 (**Multiple features**)

Characterizing the Strength of Ties

- Neighborhood overlap (NO)

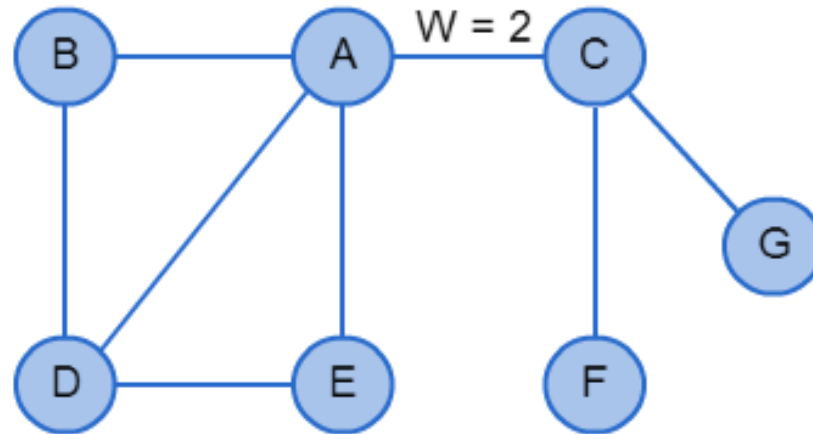


$$\frac{|A_{c_1} \cap A_{c_2}|}{(|A_{c_1} \cup A_{c_2}| - (|A_1, A_2, \textit{themselves}|))}$$

Example: B-E = 1/4

Problems with NO and Weight

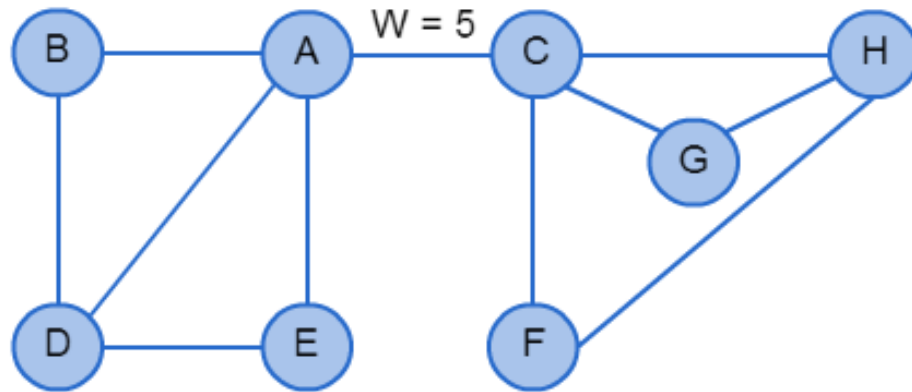
- Case 1: No common co-author



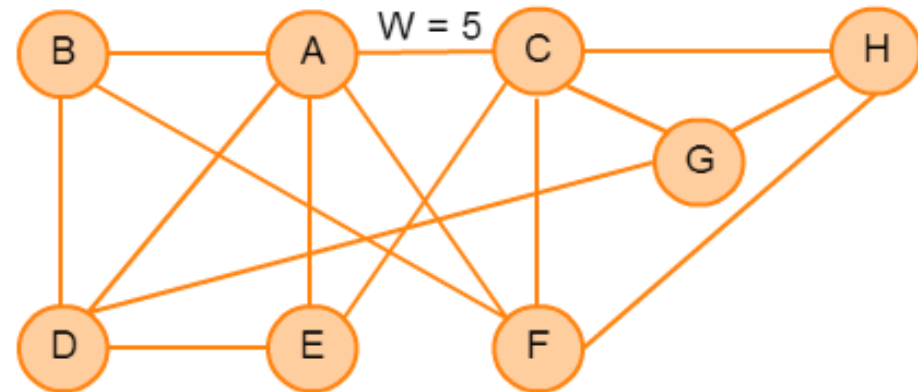
$$NO(A,C) = 0/5 = 0 \text{ and } W(A,C) = 2$$

Problems with NO and Weight

- Case 2: No representation if the tie is inside a community or not



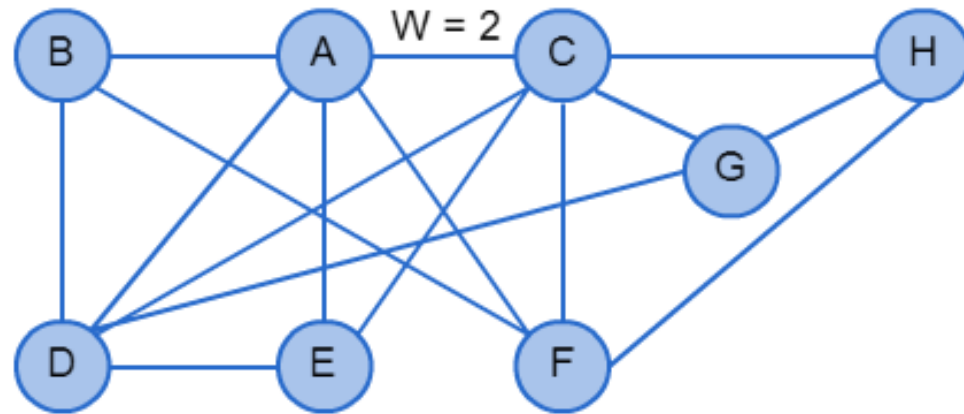
$$\text{NO}(A,C) = 0/6 = 0 \text{ and } W(A,C) = 5$$



$$\text{NO}(A,C) = 2/6 = 0.33 \text{ and } W(A,C) = 5$$

Problems with NO and Weight

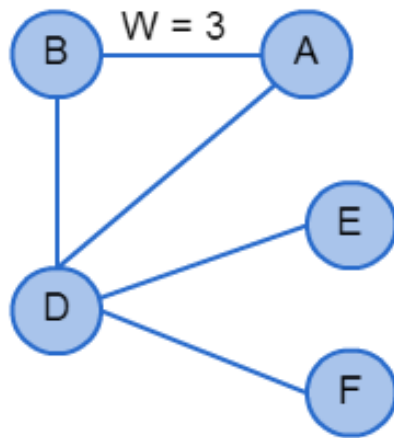
- Case 3: Many common co-authors



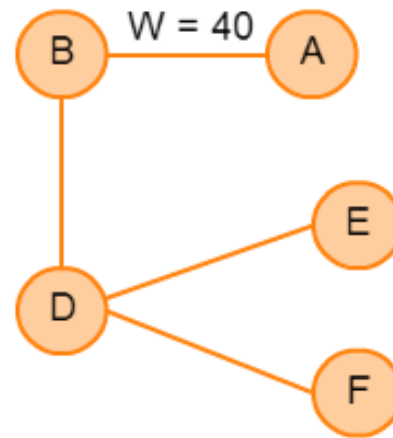
$$NO(A,C) = 3/6 = 0.5 \text{ and } W(A,C) = 2$$

Problems with NO and Weight

- Case 4: Results very small or too high



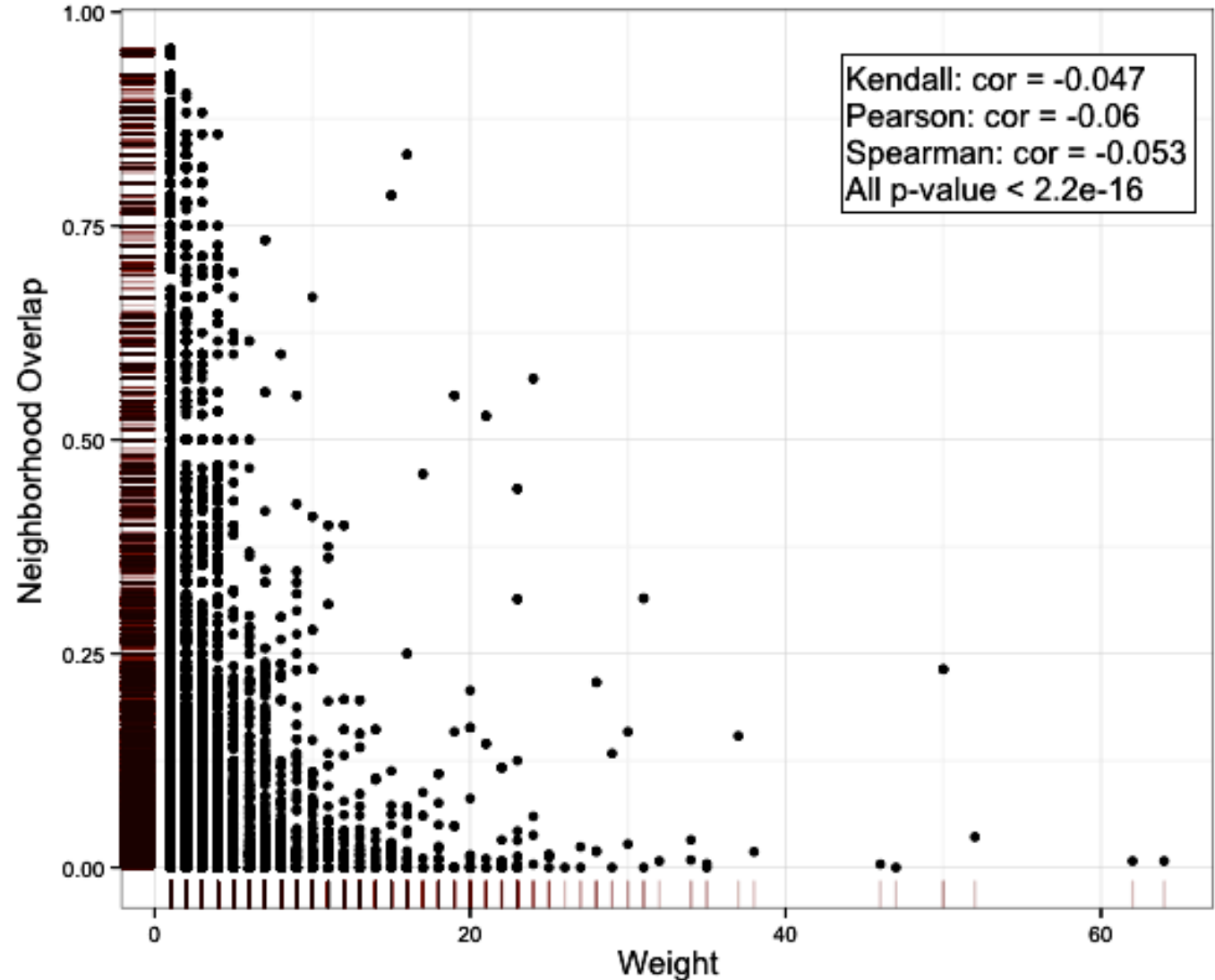
$$\text{NO}(A,B) = 1/1 = 1 \text{ and } W(A,B) = 3$$



$$\text{NO}(A,B) = 0/1 = 0 \text{ and } W(A,B) = 40$$

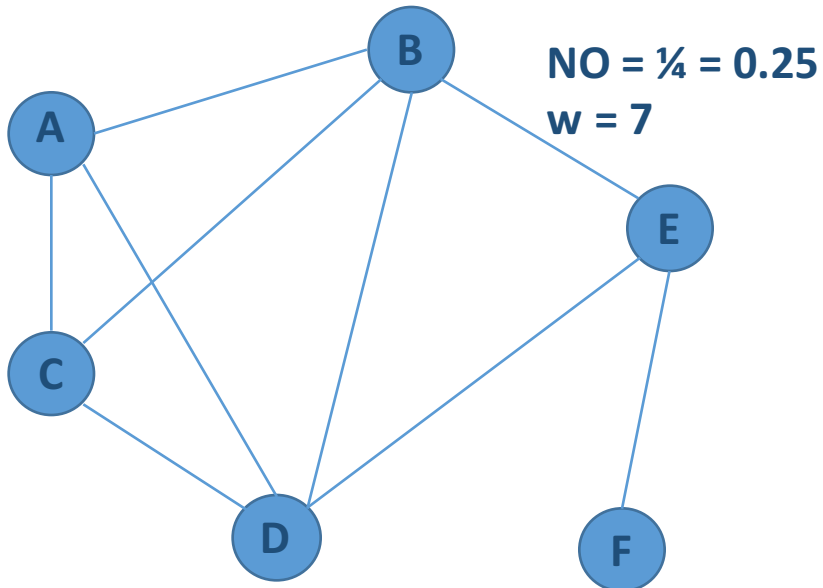
Dataset and SN Information

- DBLP
 - Publications from articles and inproceedings
 - From the year 1964 to 2015
 - 145,604 researchers
 - 208,808 links between them from 1,096,797 publications



Tieness: A New Metric

- A combination between a modification in NO and weight

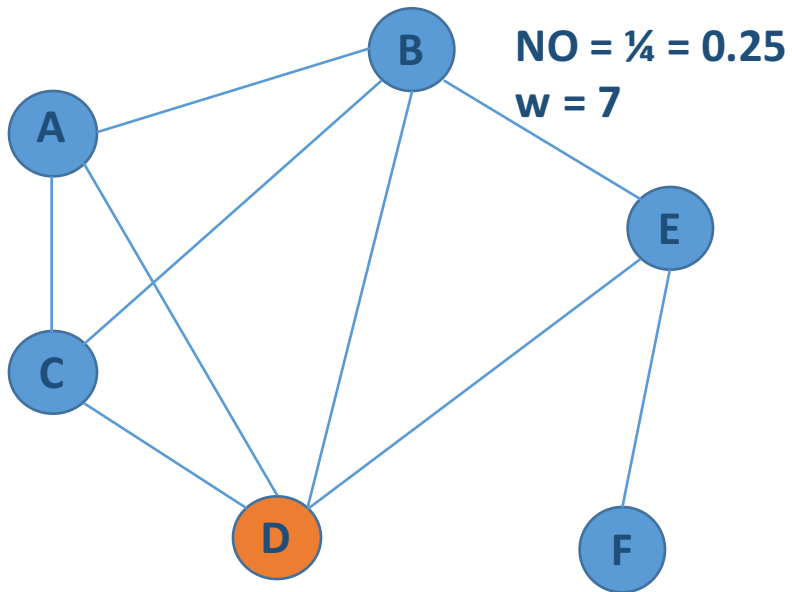


$$\frac{|A_{c_1} \cap A_{c_2} + 1|}{(|A_{c_1} \cup A_{c_2}| - (|A_1, A_2, \textit{themselves}|))} * \textit{weight}$$

Example: B-E = ?

Tieness: A New Metric

- A combination between a modification in NO and weight

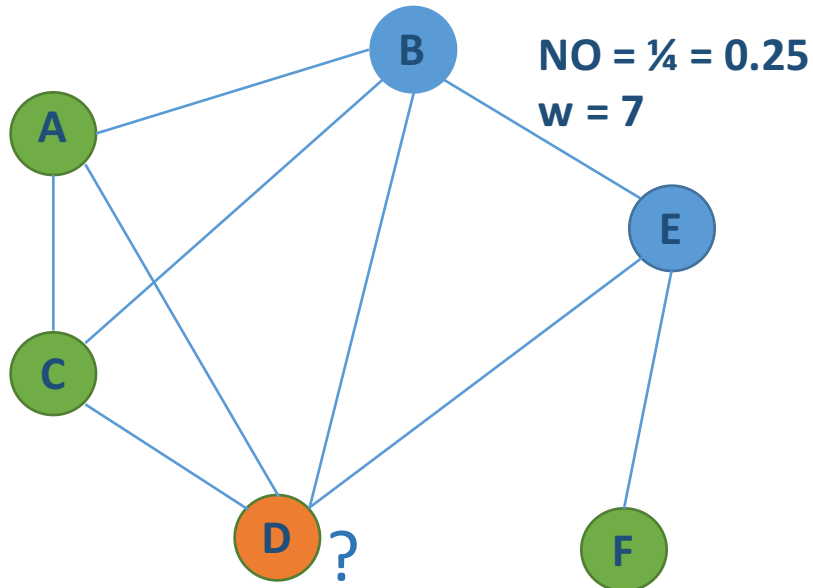


$$\frac{|A_{c_1} \cap A_{c_2} + 1|}{(|A_{c_1} \cup A_{c_2}| - (|A_1, A_2, \textit{themselves}|))} * \textit{weight}$$

Example: B-E = 2/? * 7

Tieness: A New Metric

- A combination between a modification in NO and weight

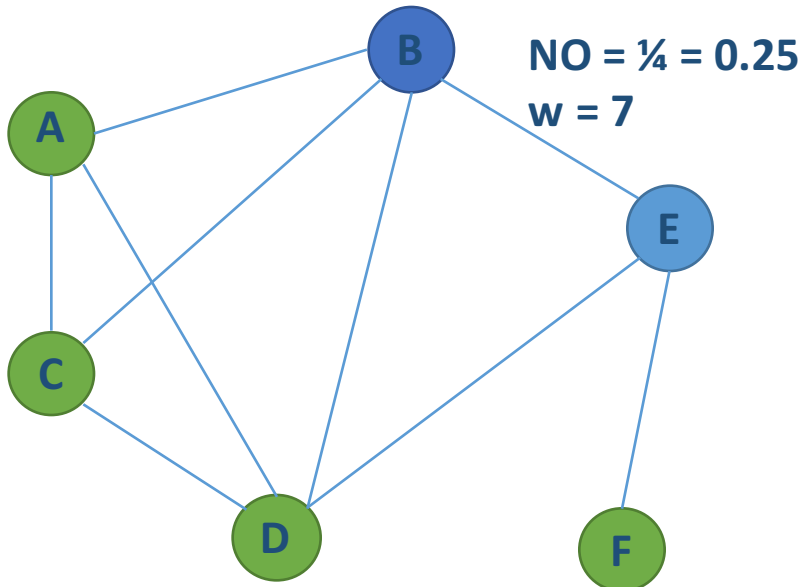


$$\frac{|A_{c_1} \cap A_{c_2} + 1|}{(|A_{c_1} \cup A_{c_2}| - (|A_1, A_2, \textit{themselves}|))} * \textit{weight}$$

Example: B-E = $2/? * 7$

Tieness: A New Metric

- A combination between a modification in NO and weight



$$\frac{|A_{c_1} \cap A_{c_2} + 1|}{(|A_{c_1} \cup A_{c_2}| - (|A_1, A_2, \textit{themselves}|))} * \textit{weight}$$

Example: B-E = $\frac{2}{4} * 7 = 3.5$

Tieness: A New Metric

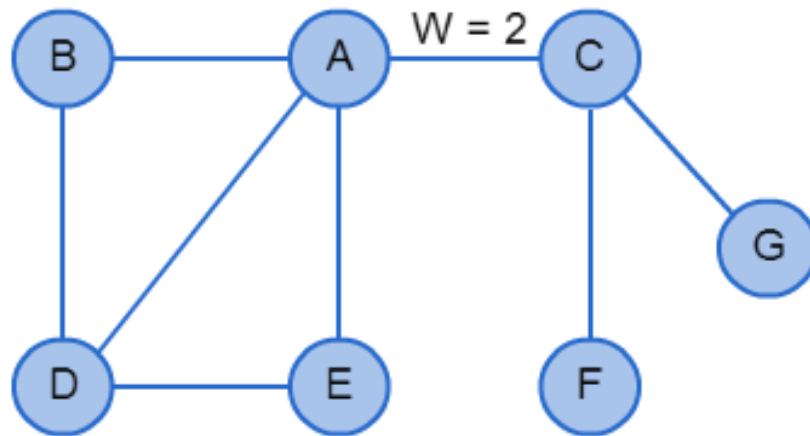
- Overall level of tieness in a network

$$\overline{tieness} = \frac{1}{|E|} \sum_{i=1}^{|E|} tieness_i$$

$|E|$ is the number of edges in the social network

Tieness: A New Metric

- Case 1: No common co-author

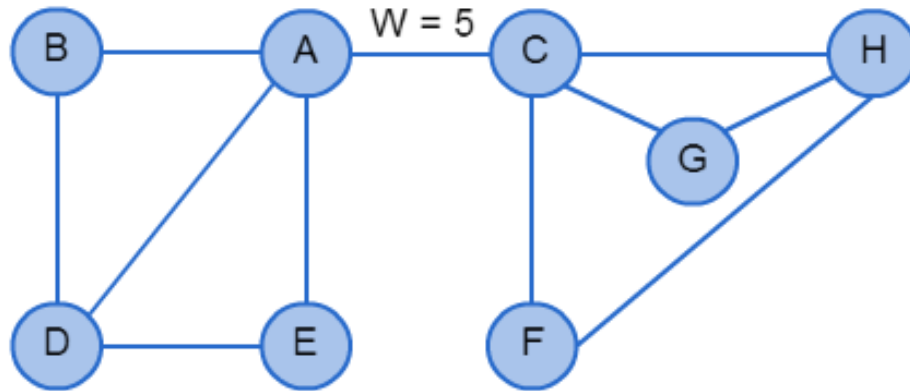


$$NO(A,C) = 0/5 = 0 \text{ and } W(A,C) = 2$$

Tieness = 0.4

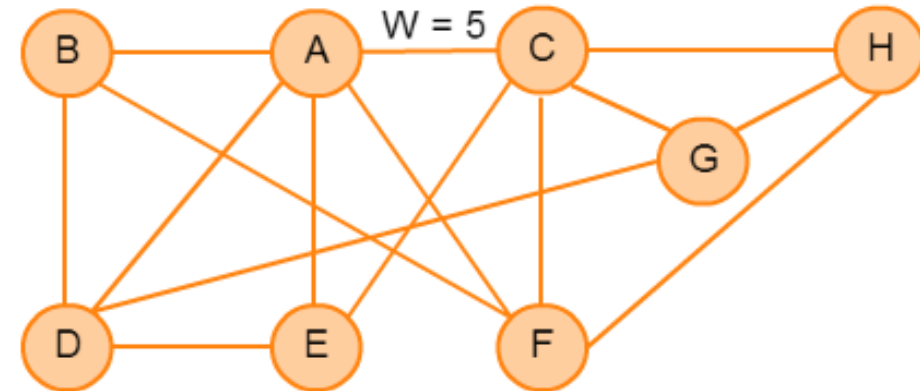
Tieness: A New Metric

- Case 2: No representation if the tie is inside a community or not



$$NO(A,C) = 0/6 = 0 \text{ and } W(A,C) = 5$$

Tieness = 0.83

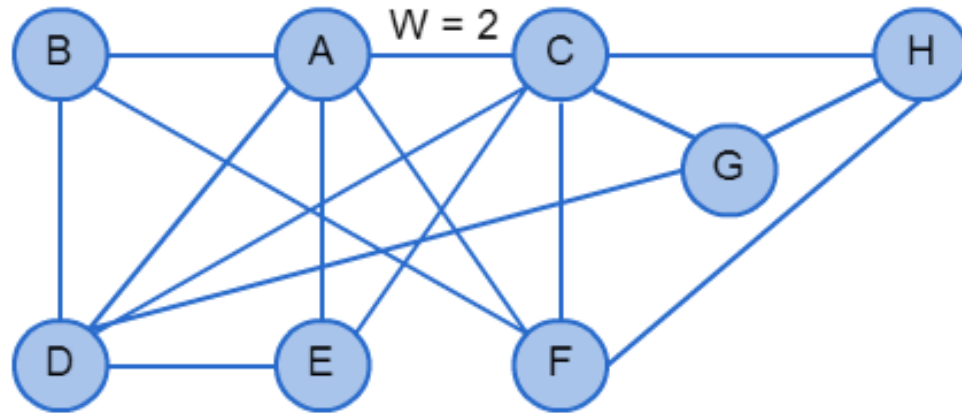


$$NO(A,C) = 2/6 = 0.33 \text{ and } W(A,C) = 5$$

Tieness = 2.5

Tieness: A New Metric

- Case 3: Many common co-authors



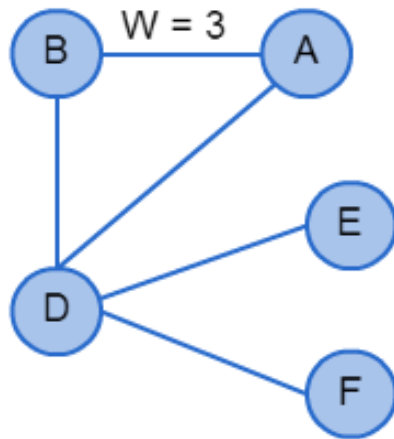
$$NO(A,C) = 3/6 = 0.5 \text{ and } W(A,C) = 2$$

Tieness = 1.33

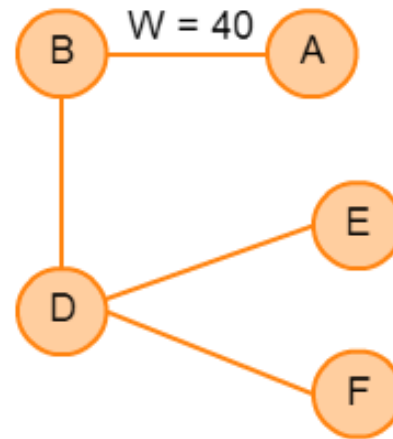
Tieness: A New Metric

- Case 4: Results very small or too high

Tieness = 6



$$NO(A,B) = 1/1 = 1 \text{ and } W(A,B) = 3$$



$$NO(A,B) = 0/1 = 0 \text{ and } W(A,B) = 40$$

Tieness = 40

Tieness: A New Metric

- Example: Real case

Top 10 Mirella M. Moro SN			
Co-Author	Weight	NO	Tieness
Alberto H. F. Laender	15	0.17	2.69
Zografoula Vagena	9	0.015	0.29
Vassilis J. Tsotras	9	0.031	0.43
Giseli Rabello Lopes	6	0.076	0.56
José Palazzo M. de Oliveira	6	0.055	0.37
Nina Edelweiss	6	0.097	0.69
Clesio Saraiva dos Santos	5	0.046	0.317
Michele A. Brandão	5	0.077	0.476
Thiago H. P. Silva	5	0.077	0.476
Altigran Soares da Silva	4	0.17	0.75

Tieness: A New Metric

- Example: Real case

Top 10 Mirella M. Moro SN			
Co-Author	Weight	NO	Tieness
Alberto H. F. Laender	15	0.17	2.69
Zografoula Vagena	9	0.015	0.29
Vassilis J. Tsotras	9	0.031	0.43
Giseli Rabello Lopes	6	0.076	0.56
José Palazzo M. de Oliveira	6	0.055	0.37
Nina Edelweiss	6	0.097	0.69
Clesio Saraiva dos Santos	5	0.046	0.317
Michele A. Brandão	5	0.077	0.476
Thiago H. P. Silva	5	0.077	0.476
Altigran Soares da Silva	4	0.17	0.75

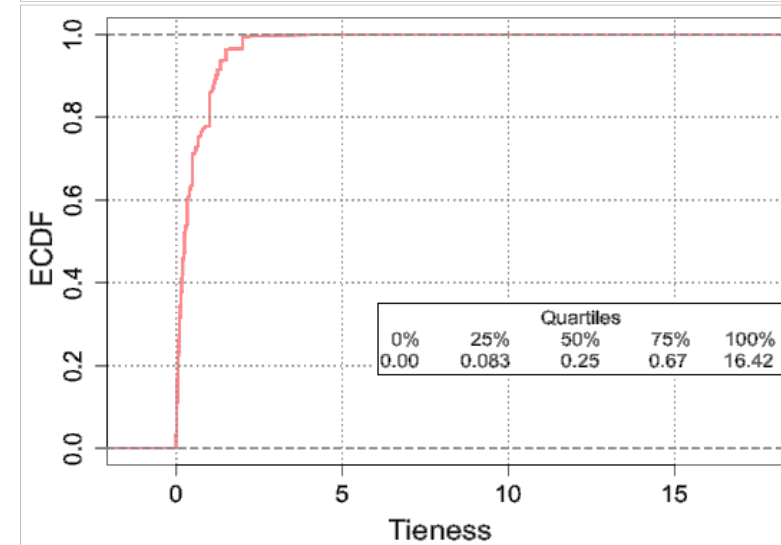
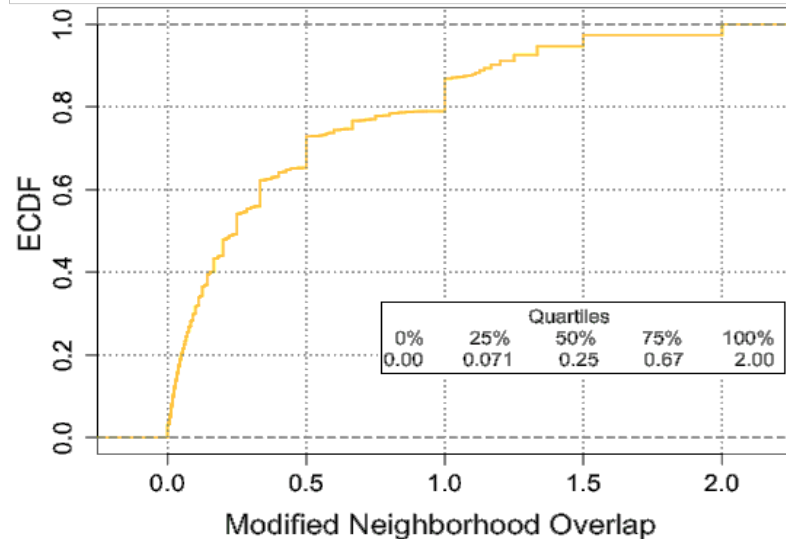
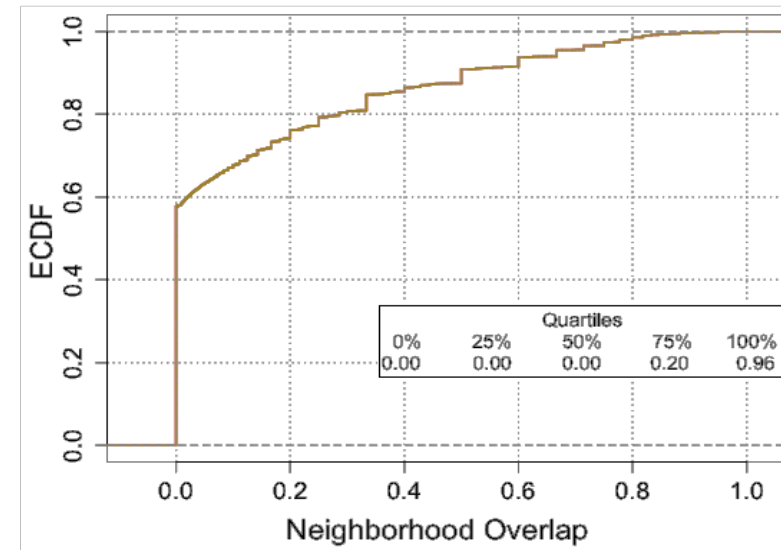
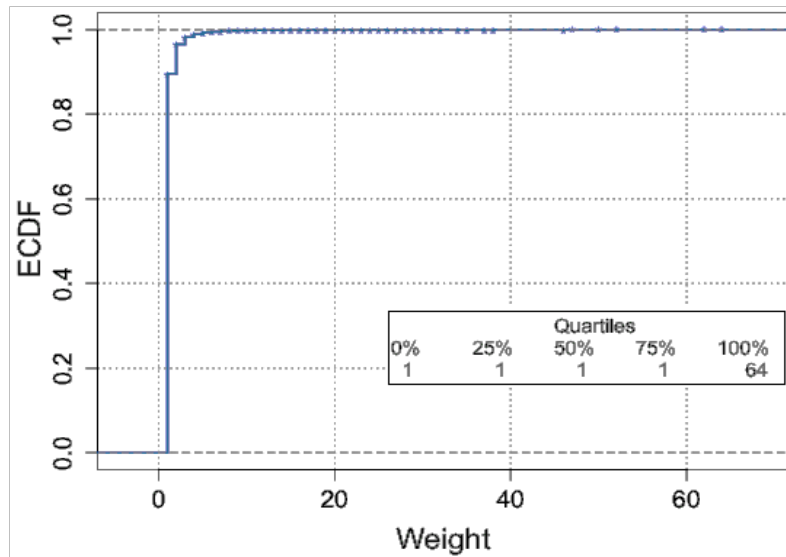
Tieness: A New Metric

- Example: Real case

Top 10 Mirella M. Moro SN			
Co-Author	Weight	NO	Tieness
Alberto H. F. Laender	15	0.17	2.69
Zografoula Vagena	9	0.015	0.29
Vassilis J. Tsotras	9	0.031	0.43
Giseli Rabello Lopes	6	0.076	0.56
José Palazzo M. de Oliveira	6	0.055	0.37
Nina Edelweiss	6	0.097	0.69
Clesio Saraiva dos Santos	5	0.046	0.317
Michele A. Brandão	5	0.077	0.476
Thiago H. P. Silva	5	0.077	0.476
Altigran Soares da Silva	4	0.17	0.75

Tieness: A New Metric

- Defining a nominal scale



Tieness: A New Metric

- Defining a nominal scale

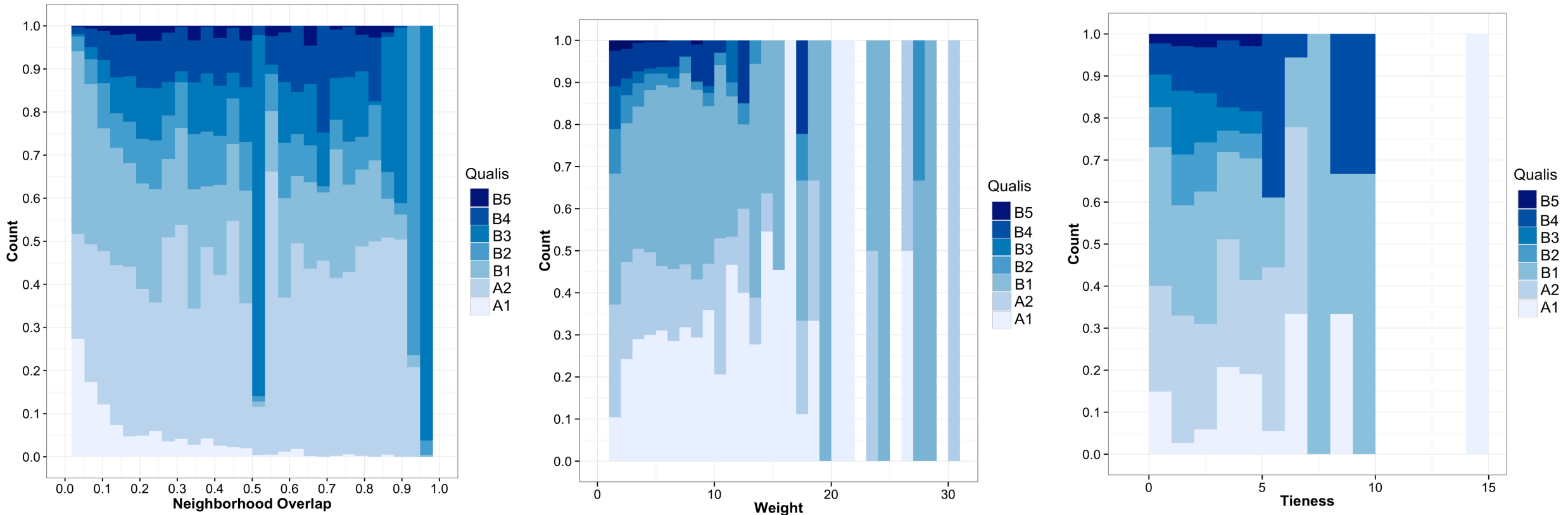
Quartiles				
0%	25%	50%	75%	100%
0.00	0.083	0.25	0.67	16.42

$$\left\{ \begin{array}{l} \textit{weak}, \textit{tieness} \leq 0.25 \\ \textit{moderate}, 0.25 < \textit{tieness} < 0.67 \\ \textit{strong}, \textit{tieness} \geq 0.67 \end{array} \right.$$

Granovetter's theory analysis → agreement

Publications' Quality

- Does the quality of venues related to the strength of ties?



Proportion of co-authorship in venues grouped by Qualis rankings

Summary of Results

RQ1: How can we measure the strength of collaboration ties?

RQ2: What is the relationship between the strength of ties and publications' quality?

- Some problems with weight and NO
- Weight + NO + tieness → first step toward classifying tie strength
- Nominal scale to tieness
- Analyses of the strength of ties in publications of different Qualis rankings
- Future work → new model + temporal aspect

Using Topological Properties to Measure the Strength of Co-authorship Ties

Michele A. Brandão, Matheus A. Diniz & Mirella M. Moro

Universidade Federal de Minas Gerais

{micheleabrandao, matheusad, mirella}@dcc.ufmg.br



BRASNAM@2016