

Is there a correlation between importance and centrality?

Evaluation of characters with social networks in plays of the Spanish Silver Age

Teresa Santa María, Elena Martínez Carro,
Concepción Jiménez y José Calvo Tello

Index

1. Introduction
2. Initial questions
3. Data
4. Results
5. Conclusions

Introduction

Biblioteca Electrónica Textual del Teatro en Español de 1868-1936 (BETTE)

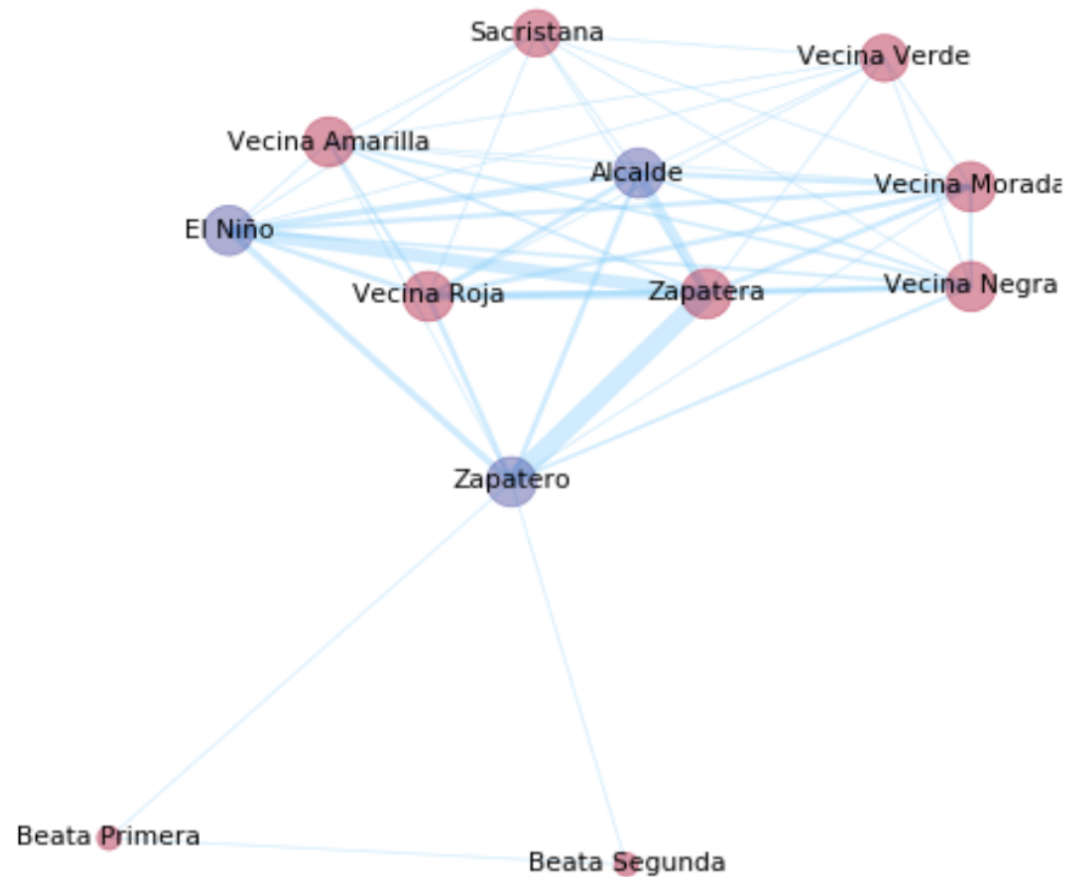
DOI 10.5281/zenodo.995969

Introducción

La *Biblioteca Electrónica Textual del Teatro en Español* (BETTE) es una colección de textos teatrales de la Edad de Plata española, codificada en XML-TEI por el grupo GHEDI de la Universidad Internacional de La Rioja (UNIR). En esta primera fase (BETTE 1.0) hemos publicado 25 textos de 8 autores (Lorca, Valle, Munoz, Echegaray, Unamuno, Valera, Clarin y Dicenta).

<https://github.com/GHEDI/BETTE>

Introduction



García Lorca *Zapatero prodigioso*’s. Ghaph by Python

Initial questions

1. How strong is the correlation between the **centrality measures** and the **importance of the character**?
2. Do the most important characters appear at the beginning of the ***dramatis personae***?
3. Is there a correlation between **importance** and **textual values** (quantities of textual units of the character)?
4. What **values** could we use to **distinguish the main characters** from the rest?

Quantitative textual values

- **Position** in the *dramatis personae* (castList)
- Amount of **text that he pronounces**
- Number of **interventions**
- Number of **references** to his name
- Number of **scenes** in which **he appears**

The importance value's criteria:

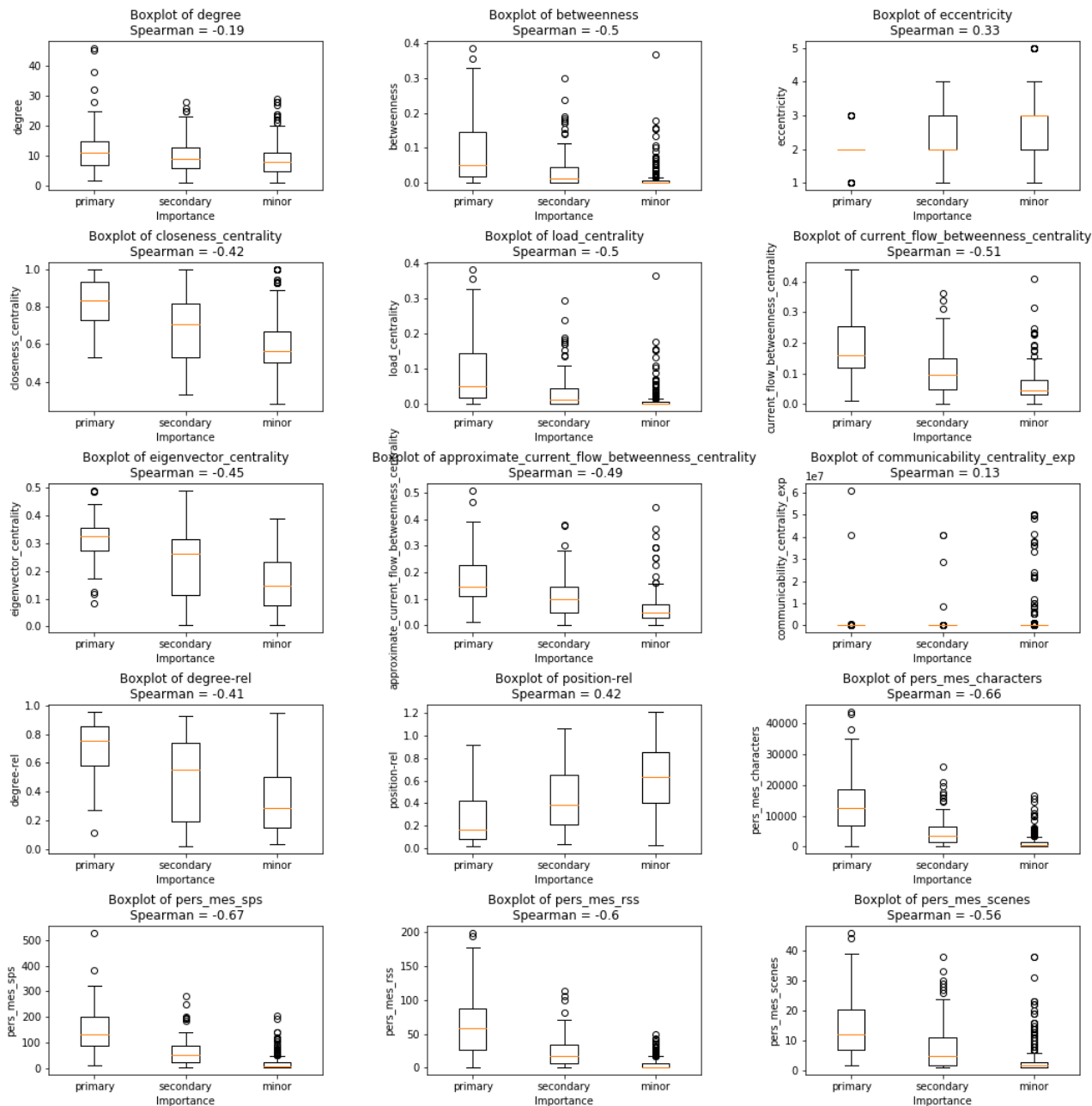
- **Minor:** if the character does not appear on the summary (also contained in the TEI file)
- **Secondary:** if he appears on the summary
- **Primary:** if he belongs to the group of between two and four essential characters

Each of
the 516
characters
has:

1. A value of his importance on the play (which can be used as *ground truth*)
2. Different textual quantitative values
3. Position in the *dramatis personae*
4. Different values according to measures of centrality

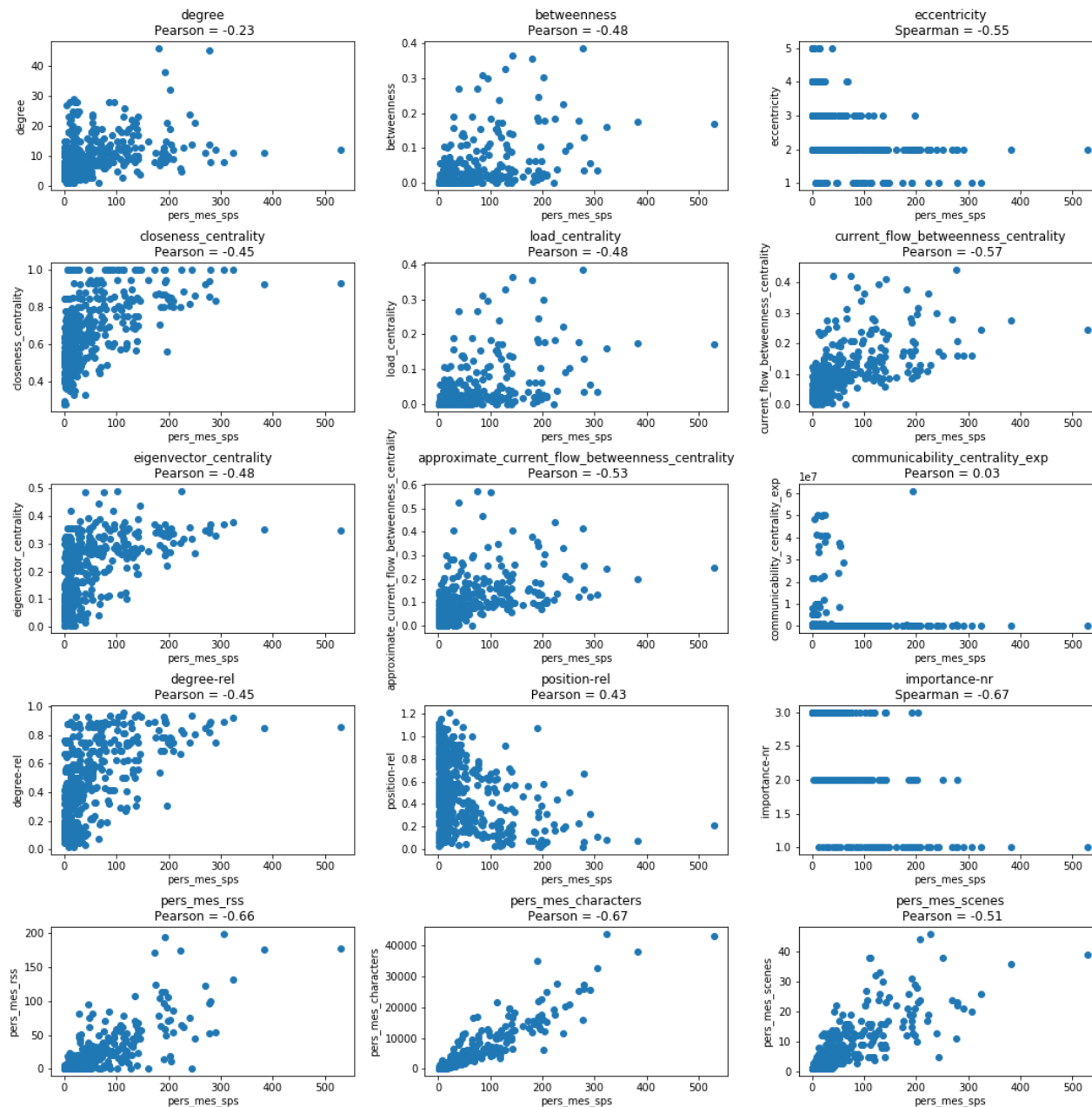
- Degree
- Betweenness centrality
- Eccentricity
- Closeness centrality
- Load centrality
- Current flow betweenness centrality (information centrality)
- Eigenvector centrality
- Approximate current flow betweenness centrality
- Communicability centrality exp.

Results



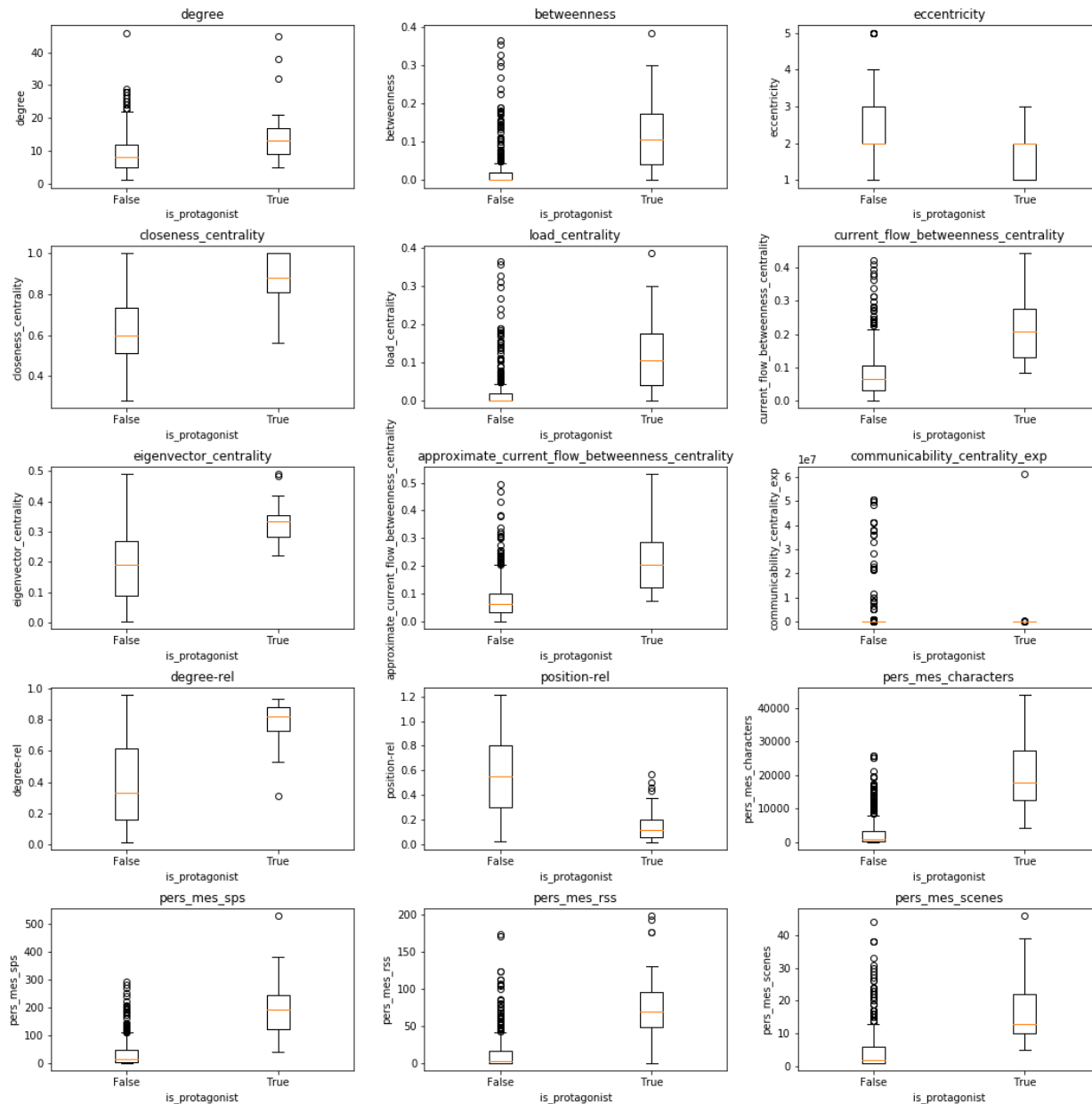
Boxplots
and
correlations
with
importance
of all
BETTE's
plays

Results



Scatterplots
with correlation
between the
times a
character
speaks (<sp>)
and other
values

Results



Boxplots of
protagonists
in front of the
other
characters

Conclusions

1. The **importance** has a correlation only between weak or moderate with any formalization of centrality, being the strongest correlation with the information centrality
2. The position in the *dramatis personae* can be used as a mean of **differentiation between primary and minor characters**, but not to differentiate these from secondary ones
3. **Textual quantitative values** have the strongest correlations with the importance and are also those that best **distinguish characters** between protagonists and not protagonists.
4. **Textual units**, simpler than networks, provide more information on both the importance of the characters and their role as protagonists.

Thank you very much!!!



DH2018

Mexico City | 26-29 JUNE