A (dis)similarity index for comparing two character networks

François Bavaud
SLI / UNIL
francois.bavaud@unil.ch

Coline Métrailler
SLI / UNIL
coline.metrailler@unil.ch

1 Introduction

Very often, networks of fictional characters exist in two or more versions. For instance (section 4), network A is built form a novel, and network B form a movie adaptation. Besides the main characters common to both versions, there are characters proper to a single version only. Also, the importance of common characters (as e.g. measured by their relative occurence), is bound to vary between the two versions, as is the strength of their mutual relations (as e.g. measured by their relative co-occurence). Hence, the networks A and B differ both along character weights (nodes) and interaction weights (edges).

Our contribution proposes the definition of a single index measuring the overall similarity between A and B. In particular, $\texttt{RV} \in [0,1]$, with $\texttt{RV} = 1$ iff A and B are identical, and $\texttt{RV} = 0$ iff A and B have no character in common. This similarity coefficient can be transformed into a dissimilarity coefficient, which can be additively decomposed into five interpretable components.

2 Formalism (RV index and MDS)

The proposed similarity index is a weighted version of the so-called RV-coefficient (Robert and Escoufier, 1976), initially defined as the *cosine similarity* between the covariances of two sets of multivariate features $\boldsymbol{X}_A$ and $\boldsymbol{X}_B$.

In the dual, individual-centered perspective relevant for tackling the comparaison of character networks, the proposed similarity index reads

$$\texttt{RV} = \frac{c_{AB}}{\sqrt{c_{AA}\, c_{BB}}} \qquad c_{AB} = \text{trace}(\boldsymbol{K}_A \boldsymbol{K}_B) \quad (1)$$

where $\boldsymbol{K}_A$ is the matrix of weighted scalar products between individuals or *kernel*, computed as

$$\boldsymbol{K}_A = \text{diag}(\sqrt{\boldsymbol{f}_A})\, \boldsymbol{B}_A\, \text{diag}(\sqrt{\boldsymbol{f}_A}) \quad (2)$$

$$\boldsymbol{B}_A = -\frac{1}{2}\boldsymbol{H}_A \boldsymbol{D}_A \boldsymbol{H}_A^\top \quad (3)$$

where $\boldsymbol{f}_A$ is the vector of individual relative weights, normalized to one, $\boldsymbol{D}_A$ is the matrix of squared Euclidean dissimilarities between individuals (as computed from their features $\boldsymbol{X}_A$), $\boldsymbol{H}_A = \boldsymbol{I} - \boldsymbol{1}\boldsymbol{f}_A^\top$ is the centering matrix, and $\boldsymbol{B}_A$ is the usual matrix of scalar products. Spectral decomposition of kernels $\boldsymbol{K}_A$ (and $\boldsymbol{K}_B$) enables the *weighted classical multidimensional scaling* (MDS), permitting to extract individual coordinates reproducing the dissimilarities $\boldsymbol{D}_A$, and whose low-rank restriction expresses a maximum amount of the configuration dispersion or *inertia* $\Delta_A = \frac{1}{2}\boldsymbol{f}_A^\top \boldsymbol{D}_A \boldsymbol{f}_A$.

3 The case of character networks

Adapting the above formalism to the comparison of character networks raises interesting formal challenges:

(1) character networks have to be converted into weighted configurations $(\boldsymbol{f}, \boldsymbol{D})$, where $\boldsymbol{D}$ is a squared Euclidean dissimilarity between characters; many available competing definitions and corresponding algorithms (commuting times, heat kernels, Sinkhorn, Metropolis) permit to compute the latter.

(2) crucially, relative weights $\boldsymbol{f}_A$ and $\boldsymbol{f}_B$ (set to the uniform weights $1/n$ in most multivariate proposals) may *differ* to a spectacular extent: their supports $\text{supp}(\boldsymbol{f}_A)$ and $\text{supp}(\boldsymbol{f}_B)$ do not even coincide in general, since version A may contain characters absent in version B, and vice-versa.

Define the *compromise weight* $\mathbf{h}$ as

$$h_i = \frac{\sqrt{f_i^A f_i^B}}{Z} \quad \text{where } Z = \textstyle\sum_j \sqrt{f_j^A f_j^B} \in [0,1] \quad . \quad (4)$$

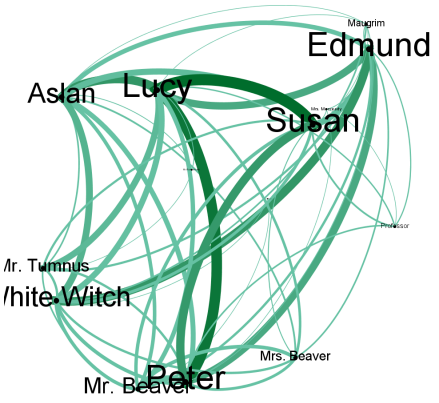

Figure 1: character network of the book *The Lion, the Witch and the Wardrobe*.

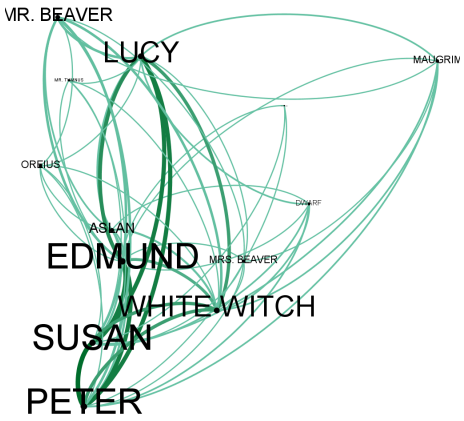

Figure 2: character network of the movie *The Chronicles of Narnia*.

By construction, $h_i = 0$ unless character i appears in both versions. A little algebra demonstrates the numerator of the similarity index (1) to express as

$$\frac{c_{AB}}{Z^2} = \text{trace}(\boldsymbol{K}_{hA}\boldsymbol{K}_{hB}) + \kappa_{AB} \qquad (5)$$

$$\text{where} \quad \kappa_{AB} = D^A_{hf_A} D^B_{hf_B} + 2\sum_i h_i \ell^A_i \ell^B_i \ .$$

Here $\boldsymbol{K}_{hA}$ is the kernel associated to configuration $(\boldsymbol{h}, \boldsymbol{D}_A)$ and $D^A_{hf_A}$ is the squared Euclidean distance between the gravity centers of $(\boldsymbol{h}, \boldsymbol{D}_A)$ and $(\boldsymbol{f}_A, \boldsymbol{D}_A)$ (terms involving B are defined analogously). Term κ_{AB}, which can negative, represents a correction due the non-coincidence of the $\boldsymbol{f}_A$- and $\boldsymbol{h}$-centroids in configuration A (respectively the $\boldsymbol{f}_B$- and $\boldsymbol{h}$-centroids in configuration B); its second component involves a weighted covariance between the $\boldsymbol{h}$-centered vectors

$$\boldsymbol{\ell}_A = \boldsymbol{B}_{hA}\boldsymbol{f}_A \qquad \text{and} \qquad \boldsymbol{\ell}_B = \boldsymbol{B}_{hB}\boldsymbol{f}_B$$

4 Illustration: The Chronicles of Narnia

The Lion, the Witch and the Wardrobe was the second of the seven novels of the *The Chronicles of Narnia*, written by C. S. Lewis in 1950, and adapted in a film directed by A. Adamson released in 2005. Among a total of 37 characters identified in the pre-treatments (manual annotation of all named entities throughout the book and the movie script, then gathered into groups of aliases to create a list of distinct characters), 16 are common to both the book and the movie (i.e. they constitute the support of $\boldsymbol{h}$ in (4), with $Z = 0.890$), 8 occur in the book only, and 13 in the movie only.

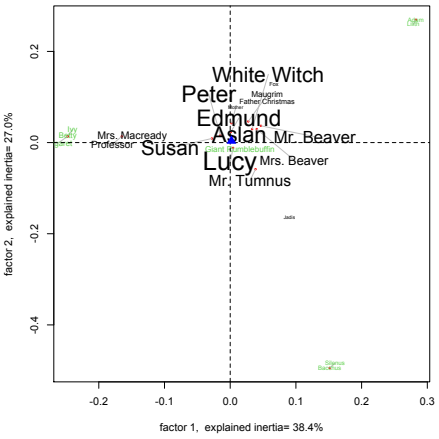

Figure 3: first MDS coordinates (x^A_i, y^A_i) of the 24 book characters (the 8 green ones appear in the book only), centered on $(\bar{x}^A_{f_A}, \bar{y}^A_{f_A}) = (0, 0)$. The blue dot depicts the centroid $(\bar{x}^A_h, \bar{y}^A_h) = (0.002, 0.004)$, fairly close to the origin: $D^A_{hf_A} = 5.6 \cdot 10^{-5}$.

The symmetric *book cross-count matrix* $\boldsymbol{N}^A = (n^A_{ij})$ counts the number of co-occurrences of characters i and j within a window of 5 paragraphs (each paragraph being delimited by a line break), with $n^A_{ii} = 0$. The associated character networks obtain by defining n^A_{ij} as the absolute weight of the unoriented edge ij, and $n^A_{i\bullet} = \sum_j n^A_{ij}$ as the absolute weight of node i. The relative book weights $\boldsymbol{f}_A$ of the previous section are defined as $f^A_i = n^A_{i\bullet}/n^A_{\bullet\bullet}$. The same treatment yields the movie graph form the movie cross-count matrix $\boldsymbol{N}^B$.

The cross-counts matrices permit to define a so-called *instantaneous jump process*, permitting to navigate from node i to node j with a rate given by (minus) the components of the *normalized Lapla-*

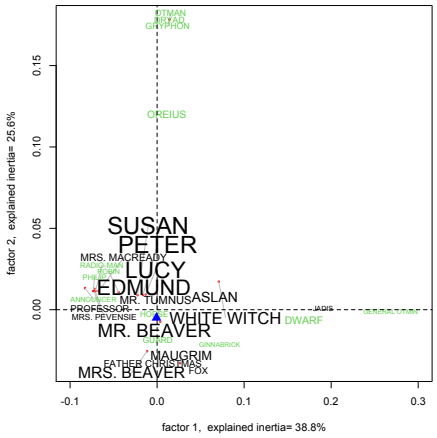

Figure 4: first MDS coordinates (x_i^B, y_i^B) of the 29 movie characters (the 13 green ones appear in the movie only), centered on $(\bar{x}^B_{f_B}, \bar{y}^B_{f_B}) = (0, 0)$. The blue dot depicts the centroid $(\bar{x}^B_h, \bar{y}^B_h) = (-0.0007, -0.005)$, fairly close to the origin again: $D^B_{hf_B} = 4.0 \cdot 10^{-5}$.

cian $L^A_{ij} = \delta_{ij} - n^A_{ij}/\sqrt{n^A_{i\bullet} n^A_{\bullet j}}$. They define in turn a positive semi-definite *weighted diffusion* or *heat kernel* by $\boldsymbol{K}_A(t) = \exp(-t\boldsymbol{L}_A) - \sqrt{\boldsymbol{f}_A}\sqrt{\boldsymbol{f}_A}^\top$, obeying $\boldsymbol{K}_A(t)\sqrt{\boldsymbol{f}_A} = 0$ as it must (see e.g. Bavaud, 2013). The free parameter $t > 0$ is the diffusion time, and spectral decomposition of $\boldsymbol{K}_A(t)$ (for $t = 10$) yields the MDS representation of the book characters in figure 3. Figure 4 depicts the analogous MDS representation of the movie characters constructed from $\boldsymbol{N}^B$.

At this stage, the book and the movie network of characters have been expressed into kernels, namely $\boldsymbol{K}_A$ and $\boldsymbol{K}_B$; alternative kernel constructions are possible, their discussion being beyond the scope of this paper. Applying the above formalism yields a coefficient of similarity (1) of $\mathtt{RV} = 0.113$. Since $\boldsymbol{h}$ centroids are close to the $\boldsymbol{f}_A$-, respectively $\boldsymbol{f}_B$-centroids, the relative importance of both terms in the r.h.s. of (5) is very contrasted, namely 100.11% and -0.11% (κ_{AB} being negative).

5 An exact decomposition formula

A similarity coefficient such as $\mathtt{RV} \in [0, 1]$ can be converted into a *dissimilarity* coefficient $d \in [0, \infty)$ by $d = -\ln \mathtt{RV}$. Applying the transformation to (1) yields the following exact decomposition for the dissimilarity between character networks A and B:

$$d_{AB} = -\ln \mathtt{RV} = -\ln \mathtt{RV}_h - 2\ln Z - \\ -\frac{1}{2}\ln \Gamma_A - \frac{1}{2}\ln \Gamma_B - \ln(1+\epsilon) \quad (6)$$

where

- $\mathtt{RV}_h = \dfrac{\mathrm{trace}(\boldsymbol{K}_{hA}\boldsymbol{K}_{hB})}{\sqrt{\mathrm{trace}(\boldsymbol{K}^2_{hA})\mathrm{trace}(\boldsymbol{K}^2_{hB})}} \in [0, 1]$ measures the similarity between dissimilarities $\boldsymbol{D}_A$ and $\boldsymbol{D}_B$ in the common compromise weighting $\boldsymbol{h}$.
- $Z \in [0, 1]$ in (4) is a measure of similarity between weights $\boldsymbol{f}_A$ and $\boldsymbol{f}_B$.
- $\Gamma_A = \dfrac{\mathrm{trace}(\boldsymbol{K}^2_{hA})}{\mathrm{trace}(\boldsymbol{K}^2_A)}$ is a measure of the ratio of the (quartic) dispersion of configuration $\boldsymbol{D}_A$ in the compromise weighting $\boldsymbol{h}$ to the dispersion of $\boldsymbol{D}_A$ in the original weighting $\boldsymbol{f}_A$ (Γ_B is defined analogously). One expects $\Gamma_A \in [0, 1]$ due to the loss of diversity entailed by the disappearance of A-specific characters in the compromise weighting.
- $\epsilon = \dfrac{\kappa_{AB}}{\mathtt{RV}_h\sqrt{\Gamma_A\Gamma_B\mathrm{trace}(\boldsymbol{K}^2_A)\mathrm{trace}(\boldsymbol{K}^2_B)}}$ is a normalized measure of the centroid correction occurring in (5), whose magnitude is expected to be small since main characters are likely to occur in both versions A and B.

In the present study, (6) yields (in order)

$$2.1829 = 0.9385 + 0.2323 + 0.4349 + 0.5762 + 0.0011 \ .$$

References

François Bavaud. 2013. Testing spatial autocorrelation in weighted networks: the modes permutation test. *Journal of Geographical Systems*, 15(3):233–248.

Paul Robert and Yves Escoufier. 1976. A unifying tool for linear multivariate statistical methods: the RV-coefficient. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 25(3):257–265.