



HAL
open science

Exploring network distances for movies comparison

Majda Lafhel, Hocine Cherifi, Mohammed El Hassouni

► **To cite this version:**

Majda Lafhel, Hocine Cherifi, Mohammed El Hassouni. Exploring network distances for movies comparison. International Conference on Complex Networks and Their Applications, Nov 2021, Madrid, Spain. hal-03503915

HAL Id: hal-03503915

<https://hal.science/hal-03503915>

Submitted on 28 Dec 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Exploring network distances for movies comparison

Majda Lafhel¹, Hocine Cherifi², and Mohammed El Hassouni¹

¹ LRIT URAC 29, Mohammed V University
majdalafhel1@gmail.com, mohamed.elhassouni@gmail.com

² LIB, University of Burgundy
hocine.cherifi@u-bourgogne.fr

1 Introduction

In this paper, we investigate the similarity between movies. Rather than looking at the visual content [7][4] [9] [10], we exploit the movie script to build a multilayer network representation. Once this movie summary is built, the problem reduces to comparing networks. This work departs from previous studies concentrating on network movie representations. Indeed, work reported in the literature usually focuses on interactions between the characters [2]. To enrich the representation, we have previously proposed a multilayer network model capturing the key elements of the movie story [5][6]. It considers three types of entities, *characters* C , *keywords* K , and *locations* L , together with their interactions. The multilayer network is automatically extracted from the scripts of the movies to compare. At this point, the similarity issue reduces to compute the similarities between the layers. In previous work [3], we used the Multilayer Network Model to extract Star Wars movie networks. Then, we investigated Portrait Divergence [1] in comparing the six episodes. This paper extends our work to research appropriate measures for comparing movie networks. We investigate three classes of methods from the three main categories (statistical, embedding, and spectral approaches) to compare four movie categories (horror, adventure, romance, and comedy). The statistical class includes the Network Portrait Divergence [1] and D-distance [11]. In the embedding category, we have the Network Matrix Factorization (NetMF) [8]. The spectral section contains the Laplacian Spectra [13] and the Network Laplacian Spectral Descriptor (NetLSD) [12].

T

2 Experimental evaluation

For each movie category, we have at least 3 movies. Due to lack of space, we present only horror and romance movies. The horror category includes the movie scripts of three episodes of the Scream Saga³. In the Romance movies, we have Titanic⁴ and Episodes I and II of Twilight⁵.

³ Scream movies: Episode I (SC1) in 1995, Episode II (SC2) in 1997, and Episode III (SC3) in 1999.

⁴ Titanic movie in 1997

⁵ Twilight movies: Fascination (TW1) in 2008, and New Moon (TW2) in 2009

We show the outputs related to character layers of horror and romance movies. A quick look at character layers of horror movies in fig. 1 and character layers of romance movies in fig. 2. Each figure is composed of five heat maps, where each heat map represents the divergence of character layers using a distance measure. The experiments show the high performance of the Portrait Divergence in comparing the distance between character layers in each of horror, romance, adventure, and comedy movies. D-measure shows good results regarding horror, comedy, and romance movies. However, it fails in comparing characters in romance movies. Also, the measure NetLSD succeeds in measuring the distance between characters of horror and comedy categories, but it fails in comparing romance and adventure movies. The embedding method NetMF can only compare characters of horror movies. In brief, the Portrait Divergence outperforms other measures in comparing character layers in each of horror, romance, adventure, and comedy movies. All the approaches present good results in comparing keyword layers in Horror movies. D-measure performs great in Adventure and Comedy keyword layers. In contrast, all the measures fail in analyzing keyword and location layers in Romance movies. The portrait divergence outperforms the other approaches in comparing the similarity between location layers in Comedy movies. The Network Portrait Divergence and D-measure are accurate for Adventure movies. D-measure and NetLSD are quite effective in comparing location layers for Horror movies. NetMF and Laplacian Spectra show poor performance in investigating location layers.

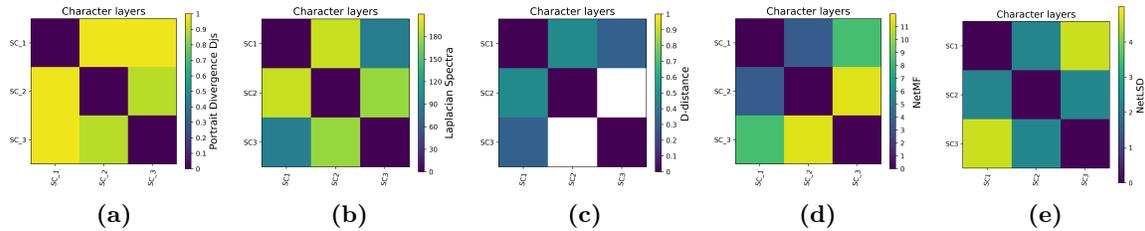


Fig. 1: Divergence of the *characters* in the 3 episodes of the Scream saga. (a) *Portrait Divergence*. (b) *Laplacian Spectra*. (c) *D-distance*. (d) *NetMF*. (e) *NetLSD*. Each cell of the heat-map represents the divergence between a couple of episodes.

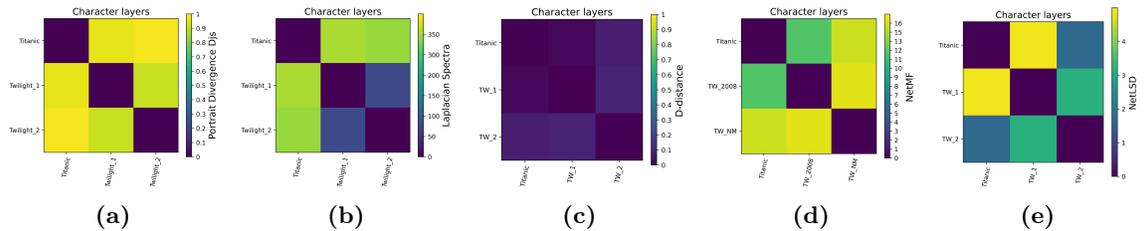


Fig. 2: Divergence of the *characters* in the 2 episodes of the Twilight and Titanic saga. (a) *Portrait Divergence*. (b) *Laplacian Spectra*. (c) *D-distance*. (d) *NetMF*. (e) *NetLSD*. Each cell of the heat-map represents the divergence between a couple of episodes.

References

1. Bagrow, J.P., Bollt, E.M.: An information-theoretic, all-scales approach to comparing networks. *Applied Network Science* 4(1), 45 (2019)

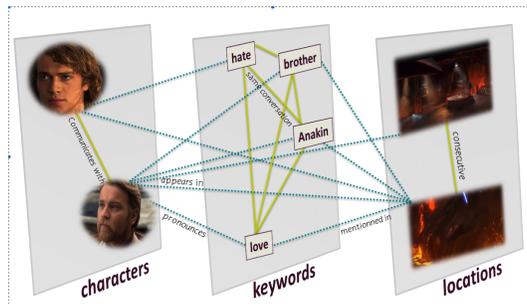


Fig. 3: Multilayer network model.

2. Labatut, V., Bost, X.: Extraction and analysis of fictional character networks: A survey. *ACM Computing Surveys (CSUR)* 52(5), 1–40 (2019)
3. Lafhel, M., Cherifi, H., Renoust, B., El Hassouni, M., Mourchid, Y.: Movie script similarity using multilayer network portrait divergence. In: *International Conference on Complex Networks and Their Applications*. pp. 284–295. Springer (2020)
4. Lasfar, A., Mouline, S., Aboutajdine, D., Cherifi, H.: Content-based retrieval in fractal coded image databases. In: *Proceedings 15th International Conference on Pattern Recognition. ICPR-2000*. vol. 1, pp. 1031–1034. IEEE (2000)
5. Mourchid, Y., Renoust, B., Cherifi, H., El Hassouni, M.: Multilayer network model of movie script. In: *International Conference on Complex Networks and their Applications*. pp. 782–796. Springer (2018)
6. Mourchid, Y., Renoust, B., Roupin, O., Văn, L., Cherifi, H., El Hassouni, M.: Movienet: a movie multilayer network model using visual and textual semantic cues. *Applied Network Science* 4(1), 1–37 (2019)
7. Pastrana-Vidal, R.R., Gicquel, J.C., Colomes, C., Cherifi, H.: Frame dropping effects on user quality perception. *Proc. 5th Int. WIAMIS* (2004)
8. Qiu, J., Dong, Y., Ma, H., Li, J., Wang, K., Tang, J.: Network embedding as matrix factorization: Unifying deepwalk, line, pte, and node2vec. In: *Proceedings of the eleventh ACM international conference on web search and data mining*. pp. 459–467 (2018)
9. Rital, S., Bretto, A., Cherifi, H., Aboutajdine, D.: A combinatorial edge detection algorithm on noisy images. In: *International Symposium on VIPromCom Video/Image Processing and Multimedia Communications*. pp. 351–355. IEEE (2002)
10. Rital, S., Cherifi, H., Miguet, S.: Weighted adaptive neighborhood hypergraph partitioning for image segmentation. In: *International Conference on Pattern Recognition and Image Analysis*. pp. 522–531. Springer, Berlin, Heidelberg (2005)
11. Schieber, T.A., Carpi, L., Díaz-Guilera, A., Pardalos, P.M., Masoller, C., Ravetti, M.G.: Quantification of network structural dissimilarities. *Nature communications* 8(1), 1–10 (2017)
12. Tsitsulin, A., Mottin, D., Karras, P., Bronstein, A., Müller, E.: Netlsd: hearing the shape of a graph. In: *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. pp. 2347–2356 (2018)
13. Wilson, R.C., Zhu, P.: A study of graph spectra for comparing graphs and trees. *Pattern Recognition* 41(9), 2833–2841 (2008)