Eurographics Conference on Visualization (EuroVis), Posters Track (2017)
A. Puig Puig and T. Isenberg (Editors)

“Them again?” Dynamic communities in the media

Haolin Ren[1,2], Marie-Luce Viaud[2] and Guy Melançon[2]

[1]National Audiovisual Institute, Paris, France
[2]Université de Bordeaux, CNRS UMR 5800 LaBRI, Talence, France

1. Introduction

Television, radio and the press invite people to publicly debate or discuss. Commonly enough, the audience experiences the “Them again!” impression, that is the impression to either ear the same stories being discussed or see the same actors and commentators on the medias. Conversely, politicians – among other public figures – may complain of not getting proper, or worse, biased coverage or presence in the media. Is this just an impression or is it a reality, at least partly?

In order to develop an objective vision on how news is discussed in the media, we investigate data specifying when often actors and commentators are co-invited in TV or radio shows. Participants – actors and commentators – being co-invited form an affiliation network, with a group of participants being often co-invited over time forming dynamic “communities”. Indeed, co-invitation links between commentators last over short periods of time and get renewed each time a TV or radio show takes place.

Typical questions raised in this context aim at observing whether the exact same group of commentators are invited altogether at the same TV or radio shows over a longer period of time? If so, how is this part of a editorial policy of a news channel, for instance? Communities studied in the literature usually emerge from more persistent “social” communities (personal social ties or family ties, for instance). As a consequence, traditional community identification approaches are not enough to capture the dynamic character of the media communities we study here. Supplementary analytical tools must be designed and assembled to properly investigate these dynamic communities.

2. Looking at Media Bias from a Different Perspective

Media bias has mostly been addressed in terms of topic detection and information retrieval in order to identify the use of biased vocabulary [KWN12], often using textual sources (reports, social media content) compared with known opinions [STCL13]. Natural language processing together with mining or learning approaches are indeed quite relevant when trying to identify and even measure bias emerging from topics addressed or terms used [FCG09]. Topic models have proven useful when trying to measure bias through the comparison of the mainstream news content with similar content in the social medias [YQK*12]. CompareClouds is an example visual analytics tools building on past work in text mining, based on a mapping of word prevalence and context information between two corpora supporting their comparison by content experts [DZS13] [DES*15].

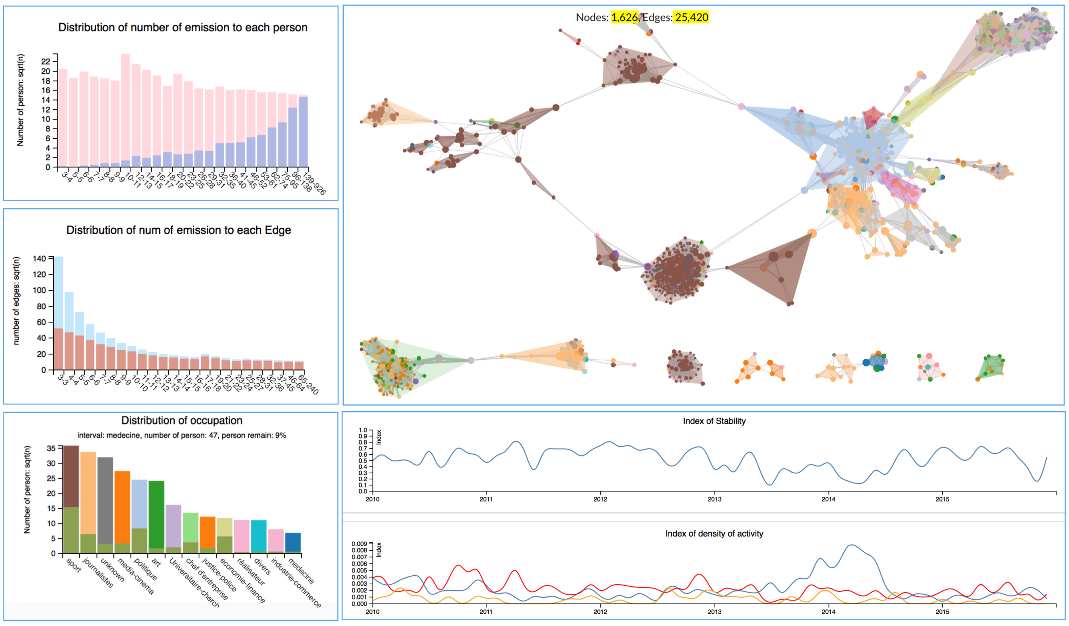
Together with domain experts, we adopt a different perspective and rather look at the participants that take part in TV or radio shows. Participants in the media build the public discourse, at least in part, on numerous societal issues such as politics. A reasonable hypothesis we make is that opinions expressed by a commentator, on any given topic, remain stable over time. Indeed, a political or military advisor, for instance, is expected to express the same view whatever the show or host he is discussing with. Similarly, a same set of participants being invited on different shows will collectively discuss a topic in a similar way and develop a similar perspective on the issue. As a consequence, we may see media bias as being induced from the repeated choice of the same participants on panels.

In this context, a main task we need to support is to identify groups of participants that more often get co-invited, as an indicator of a potential bias. Our work thus relies on the ability to detect communities (clusters) in dynamic graphs – seen as a sequence of graphs G_0, G_1, ... associated with timestamps $t = 0, 1,$ ... Approaches that have been proposed to tackle this problem fall in different categories [AFGW13]: (a) communities C^i_k in G_i are computed for each timestamp and matched (between consecutive timestamps); (b) communities $C^{i+1}_{k'}$ are somehow “deduced” from communities C^i_k; (c) communities are computed on the whole dataset $G_0 \cup G_1 \cup \cdots$ and projected back on each timestamp; (d) communities $C^{i+1}_{k'}$ are inferred from transformations occurring in the transition $G_i \to G_{i+1}$.

3. Media Presence and Dynamic Communities

French Radio and television programs are archived, referenced and stored by the French National Audiovisual Institute (INA: `www.ina.fr`). A *program* usually denotes a segment of content intended for broadcast (other than a commercial, trailer, etc.). It may be a single production, or more commonly, a series of related productions (also called a television *show*). For each TV or radio program, structured information about the program is stored – channel, show (if applicable), broadcast hours, theme (politics, arts, sports, etc.), type (late show, magazine, etc.) and most im-

DOI: 10.2312/eurp.20171173



portantly the list of guests that participated to the show. *Participants* to a show can be hosts (running the show and conducting the discussion), journalists, advisors, public figures (politicians, movie actors, etc.), etc. We also include as part of the participants the persons that are the focus of the discussion – this information is also part of the material archived by INA. For instance, if François Hollande or Donald Trump are being discussed (them personally or their policy), we then consider them part of the list of participants. Participants are also sometimes called *commentators* (more usually in the context of a discussion on public, often political, issues).

The data we use in this study gathers information about (daily, weekly, monthly or special events) programs broad casted on French radio and TV between 2011 and 2016: 491000 programs on 129 channels with 216647 guests and 14000 hosts.

4. A Visual Analytics Approach – data operations, tasks and visual encodings

The media communities we consider differ from those usually studied in the literature as they lack persistence in time. Indeed, the probability of the exact same guests being invited on a show the next day is close to null; the chance they get together again the next week or month is higher. We aim at assessing whether, over the 5 year period we consider, they get together often enough that they can indeed be considered a community (and thus induce media bias).

Our major visual analytics challenge is thus to allow expert users inspect the behavior of communities through time, question their stability and homogeneity.

The non-persistent nature of our data led us to use a community detection approach of type (c) on the overall dataset using [BGLL08]. However, the size and complexity of the whole dataset makes it necessary to filter out edges and work on the Simmelian backbone structure of the graph [NLCB13]. The backbone then allows to identify well defined communities. Edges are mapped to their timestamp to allow time-based query and navigation; discarded edges are brought back on demand (according to specific user interaction).

Synchronized views and interaction are key elements in the dashboard design (see teaser image *and* companion video). Timelines are used to graph the evolution of two statistics on communities: *density of activity* (*how much* do participants co-participate) and *dispersion* (*how* they link outside their community). Tighter communities are found by brushing the timeline, spotting time interval $[t,t']$ with high activity density and low dispersion. Sankey diagrams provide information on transformations communities undergo through time. Node-link animations allow to visually evaluate the overall stability (see companion video).

A yearly, small multiples view of communities is helpful to visualize periodicity patterns (often linked to repeated yearly events such as sport competitions). Drag and dropping a community from year y to year y' directly show whether most members of a community keep together or on conversely splatters around and dissolves.

Conclusion With expert users at INA, we are in the process of collecting user feedback, validating the tasks–visual encodings/interaction coupling and running story telling sessions.

References

[AFGW13] AYNAUD T., FLEURY E., GUILLAUME J.-L., WANG Q.: *Communities in Evolving Networks: Definitions, Detection, and Analysis Techniques*. Springer, 2013, pp. 159–200. 1

[BGLL08] BLONDEL V. D., GUILLAUME J.-L., LAMBIOTTE R., LEFEBVRE E.: Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment* (2008), P10008. 2

[DES*15] DIAKOPOULOS N., ELGESEM D., SALWAY A., ZHANG A., HOFL K.: Compare clouds: Visualizing text corpora to compare media frames. In *Proceedings IUI Workshop on Visual Text Analytics* (2015). 1

[DZS13] DIAKOPOULOS N., ZHANG A., SALWAY A.: Visual analytics of media frames in online news and blogs. In *Proceedings IEEE InfoVis Workshop on Text Visualization* (2013). 1

[FCG09] FORTUNA B., CRISTIANINI N., GALLEGUILLOS C.: *Detection of Bias in Media Outlets with Statistical Learning Methods*. Chapman & Hall/CRC Data Mining and Knowledge Discovery Series. Chapman and Hall/CRC, 2009, pp. 27–50. 1

[KWN12] KRESTEL R., WALL A., NEJDL W.: Treehugger or petrolhead?: identifying bias by comparing online news articles with political speeches. In *Proceedings of the 21st International Conference on World Wide Web* (2012), ACM, pp. 547–548. 1

[NLCB13] NICK B., LEE C., CUNNINGHAM P., BRANDES U.: Simmelian backbones: Amplifying hidden homophily in facebook networks. In *Advances in Social Network Analysis and Mining (ASONAM)* (2013), pp. 525–532. 2

[STCL13] SAEZ-TRUMPER D., CASTILLO C., LALMAS M.: Social media news communities: gatekeeping, coverage, and statement bias. In *Proceedings of the 22nd ACM international conference on Conference on information & knowledge management* (2013), ACM, pp. 1679–1684. 1

[YQK*12] YOUNUS A., QURESHI M. A., KINGRANI S. K., SAEED M., TOUHEED N., O'RIORDAN C., GABRIELLA P.: Investigating bias in traditional media through social media. In *Proceedings 21st International Conference on World Wide Web* (2012), WWW '12 Companion, ACM, pp. 643–644. 1