PLMP: A Method to Map the Linguistic Markers of the Social Discourse onto its Semantic Network

Tomaso Erseghe, Leonardo Badia
Dept. of Information Engineering
University of Padova, Italy
tomaso.erseghe@unipd.it
leonardo.badia@unipd.it

Lejla Dzanko
Center for Research on Social Relations
University of Social Sciences and
Humanities (SWPS), Poland
lejladzanko94@gmail.com

Caterina Suitner
Dept. of Developmental
Psychology and Socialization
University of Padova, Italy
caterina.suitner@unipd.it

***Abstract*—A modern interdisciplinary analysis of social networks implies detecting and investigating relevant socio-psychological linguistic markers that carry insight on the nature and characteristics of the social discourse. Associating markers to specific words is a further important step, allowing for an even richer interpretation. By taking as a working example the social discourse in Twitter, we propose a scalable method called PageRank-like marker projection (PLMP) following a rationale inspired by PageRank to fully exploit the interdependencies in a semantic network, so as to meaningfully project markers from a social discourse level (tweets) to its semantic elements (words). The effectiveness of PLMP is shown with an application example on calls to online collective action.**

***Index Terms*—Data analysis; Computational linguistics; Projection algorithms; Social networking; Sociology-psychology; #FridaysForFuture; #MeToo.**

I. Introduction

Online social networks connect people and convey ideas faster than every meeting platform, real or virtual [1]. Their analysis can capture the zeitgeist and predict evolving trends, but necessitates a strong crossbreeding among disciplines, since it requires a balanced blend of quantitative analysis, technological know-how, social sciences, and mathematical formalization. The online corpus of micro-blogging platforms is a melting pot of content, often very noisy and bubbly, that challenges researchers to extrapolate underlying meaning with analytical methods [2]. The specific challenge tackled within the present paper is the projection of socio-psychological linguistic markers from the holistic perspective of the social discourse to its semantic elements, specifically words. Building on a seminal proposal available in [3], we propose PageRank-like marker projection (PLMP), a technique that cohesively accounts for these inter-dependencies in the semantic network. In PLMP, a bipartite network of Tweets/words is stimulated by a PageRank-inspired approach [4], [5] that lets the information freely flow through the network interdependencies. Unlike the standard PageRank, an update matrix that is row-normalized (as opposed to column-normalized) is used to preserve coherence with the final aim. Mathematically proving its exactness implies a number of complicacies, that can be managed along the lines of [6], with proper modifications.

IEEE/ACM ASONAM 2022, November 10-13, 2022
978-1-6654-5661-6/22/$31.00 © 2022 IEEE

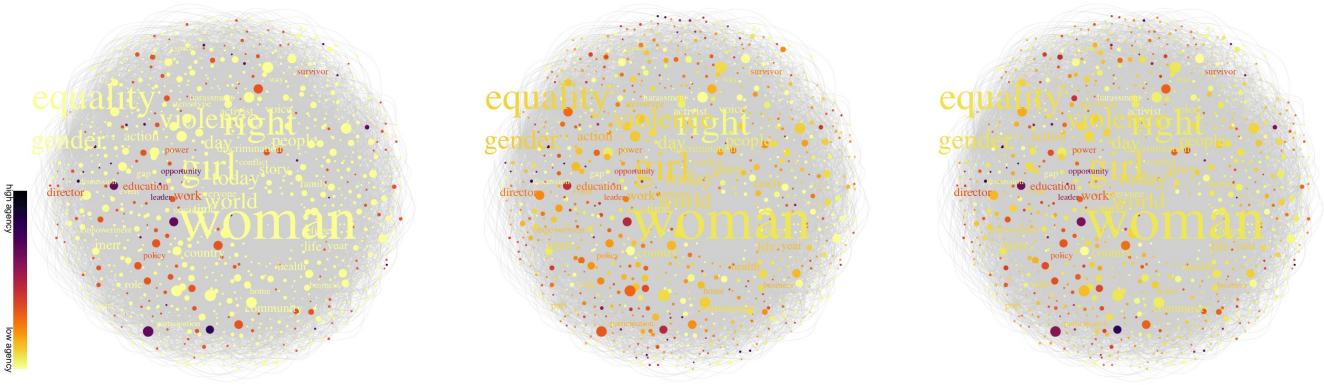
We tested PLMP on two online collective actions happening on Twitter: (1) #MeToo (October, 2017): a social movement encouraging sexual harassment victims (usually young women) to break the silence; and (2) #FridaysForFuture (August, 2018): school strikes demanding action from political leaders to prevent climate change. We investigate the rhetoric of online calls to action in the social discourse within these scenarios, capturing the structural changes of the semantic network revolving around these topics. We show through PLMP that the pertinent keywords have a translational change in relevance and meaning, which reflects in a better information flow enabling the collective action. Furthermore, we demonstrate how PLMP is able to highlight significant differences in the two calls to action.

The rest of the paper is organized as follows. In Section II we identify the semantic networks under study and detail the socio-psychological linguistic markers that can be automatically extracted from it. Section III introduces the PLMP approach, by further comparing it to a number of alternative solutions (based on the same idea) that however turned out to be less reliable. Its application to the analysis of a call-to-action is given in Section IV. Section V concludes the paper.

II. Preliminaries and Tools

A. Dataset

Among the social media, we chose Twitter as a well suited reality mirror for our analysis [1], because of its widespread usage and the option to easily access data through its APIs. The corresponding semantic network can be differently structured to coherently relate to the specific target of the study. Unlike the common approach of building a bipartite graph of tweets and hashtags [7], we rely on a representation where *tweets* are connected to all the *words* that appear inside them, so as to better capture the inter-dependencies between words in the social discourse. We limited our scope to tweets in the English language, and we sampled two groups of 5000 tweets before and after a main call to action, in two contexts:
(1) #MeToo – Tweets from the @UN_Women pages in the two periods April 1-June 30, 2017 and April 1-June 30, 2018;
(2) #FridaysForFuture – Tweets in the two periods March 1-April 19, 2018 and March 1-April 19, 2019 by using the neutral hashtag #climatechange in the search [8].



(a) agency in isolation (b) agency projection [3] (c) PLMP agency projection

Fig. 1. Agency methods (#MeToo, nouns only): node size is proportional to the PageRank value in the network, color corresponds to the level of agency.

After getting a significant corpus to read as a semantic network, identification and marking was achieved by means of Python's part of speech (POS) non deterministic tagger [2], [9]. We applied a post-processing where: all words not tagged as nouns, adjectives, adverbs, or verbs were discarded; contracted forms were expanded and non meaningful composed words were split; stopwords were removed. The remaining words were lemmatized preserving the associated POS tag.

B. Linguistic markers

The resulting semantic networks were analyzed under a socio-psychological lens that defines a collective action as any effort addressing a common goal that surpasses the individual interest, to improve group's conditions [10]. The most relevant socio-psycological drives for engaging in collective action, recognized in [11], correspond to the ideas of *social identity* and *collective efficacy*, i.e., associating with the topic or consider it important, and believing that the actions can contribute to a broader change. They match with the socio-psychological concepts of *affiliation* and *agency* [8], which can be inferred from tweets by means of Linguistic Inquire and Word Count (LIWC) 2015 [12], a well-established tool for detecting linguistic proxies of psychological processes in text samples [13]. LIWC performs a dictionary-based quantitative content analysis where every message receives a score on several categories derived from the number of words belonging to the specific category adjusted for the total number of words within the message. The LIWC "affiliation" entry was used, while a proxy for agency was identified [8], [14] as the average of "power," "achieve," "reward," "insight," and "cause."

III. Projection Method

A. The PLMP approach

Projections of social interconnections and relationships require carefully crafted solutions depending on the context [15], [16]. We review the problem at hand and illustrate the PLMP solution with the help of Fig. 1 by taking as a reference the projection of agency on nouns in the #MeToo call to action (2018 data). Fig. 1.(a) depicts in different colors the level of agency in the absence of a social discourse, i.e., it measures the *in isolation* agentic meaning of nouns as inferred from applying LIWC to single words, without resorting to the social discourse expressed by tweets. We let vector $\tilde{\boldsymbol{m}}_w$ (words markers, in our case agency) carry these *in isolation* information values. The limit of such an approach is evident in that, in a specific context (e.g., the @UN_Women feminist discussion on Twitter) some words that are neutral in agency might carry a high level of agency that is driven by the social discourse (e.g., the word *woman*). That is, the agency assigned to words in isolation does not capture the agency conveyed in the complexity of a social discourse. As a consequence, we argue that the level of agency is better captured through its context-based extraction of meaning within the semantic interaction of tweets. We let vector $\tilde{\boldsymbol{m}}_t$ (tweets markers) carry this context-specific information. Our aim is to identify an algorithm that reliably assigns agency to words by accounting for the information available from both $\tilde{\boldsymbol{m}}_w$ and $\tilde{\boldsymbol{m}}_t$.

To this end, we aim to exploit the indirect effect of socio-psychological markers, in the form suggested by the PageRank algorithm [4], [5], i.e., by letting information (iteratively) flow through the semantic network. This corresponds to declaring that words are affected by the average values contained in the tweets they belong to, and, conversely, that tweets are affected by the average values of the words they carry. If we denote the (nonnormalized) adjacency matrix linking words to tweets as $\boldsymbol{B}$, then averages can be inferred from matrices

$$\begin{aligned} \boldsymbol{B}_1 &= \operatorname{diag}((\boldsymbol{B}\mathbf{1})^{-1}) \cdot \boldsymbol{B} \\ \boldsymbol{B}_2 &= \operatorname{diag}((\boldsymbol{B}^T\mathbf{1})^{-1}) \cdot \boldsymbol{B}^T \,, \end{aligned} \tag{1}$$

which are the row-normalized and the column-normalized-and-transposed counterparts to $\boldsymbol{B}$, respectively. The above rationale can be formalized through the steady-state equation

$$\underbrace{\begin{bmatrix} \boldsymbol{m}_w \\ \boldsymbol{m}_t \end{bmatrix}}_{\boldsymbol{m}} = \alpha \underbrace{\begin{bmatrix} \mathbf{0} & \boldsymbol{B}_1 \\ \boldsymbol{B}_2 & \mathbf{0} \end{bmatrix}}_{\boldsymbol{M}} \begin{bmatrix} \boldsymbol{m}_w \\ \boldsymbol{m}_t \end{bmatrix} + (1-\alpha) \underbrace{\begin{bmatrix} \tilde{\boldsymbol{m}}_w \\ \tilde{\boldsymbol{m}}_t \end{bmatrix}}_{\boldsymbol{q}} , \tag{2}$$

where $\alpha \in (0, 1)$ is a mixing parameter that controls the spreading of information. Incidentally, (2) directly works on

the bipartite network, this being a particularly welcome approach that avoids a projection onto a network of words, which would conversely discard essential information [17].

A solution to (2) can be obtained by a standard power iteration method, by successively applying

$$\boldsymbol{m}_k = \alpha \boldsymbol{M} \boldsymbol{m}_{k-1} + (1-\alpha)\boldsymbol{q} \ , \tag{3}$$

starting from the initial state $\boldsymbol{m}_0 = \boldsymbol{q}$, a necessary initialization to convey the correct solution. We note that, although (2) appears to be a standard PageRank equation, here matrix $\boldsymbol{M}$ is row-normalized and not column-normalized, which implies a number of technical complicacies for proving the exactness of (3) as well as for assessing the validity of (2). We discuss these technical aspects in more detail in Section III-B.

The result of PLMP is graphically shown in Fig. 1.(c), while Fig. 1.(b) displays the outcome of the state-of-the-art solution [3], which corresponds to $\boldsymbol{m}_w = \boldsymbol{B}_1 \tilde{\boldsymbol{m}}_t$. Note how PLMP is able to combine the agency *in isolation* of (a) and the *one-hop* projection [3], in such a way that words that are agentic in themselves keep their agency level, see the darker colours in Fig. 1.(a) and (c), and so do words that acquire agency from the social discourse, as can be inferred from the comparison of Fig. 1.(b) with (c). Interestingly, thanks to the flow through the semantic network, some words acquire more agency from the social discourse under PLMP than under [3]; e.g., compare the word *right* in Fig. 1.(b) and (c).

B. Sketch of the proof

We investigate the solutions to the steady-state equation (2) for a row-normalized square matrix $\boldsymbol{M}$, i.e., $\boldsymbol{M}\boldsymbol{1} = \boldsymbol{1}$, and for $\boldsymbol{q} \geq \boldsymbol{0}$, $\boldsymbol{q} \neq \boldsymbol{0}$. Assume that $\boldsymbol{M}$ is *irreducible*, which is the case of interest in our context, and the mixing parameter α lies in the open range $(0, 1)$, then mimic the PageRank analysis of [6], with proper modifications due to the different normalization of $\boldsymbol{M}$. Let $\boldsymbol{v}$ be the left eigenvector corresponding to the right eigenvector $\boldsymbol{1}$, i.e., related to eigenvalue 1. From the Perron-Frobenius theorem [18], without loss of generality we assume that $\boldsymbol{v} > \boldsymbol{0}$; then, (2) provides

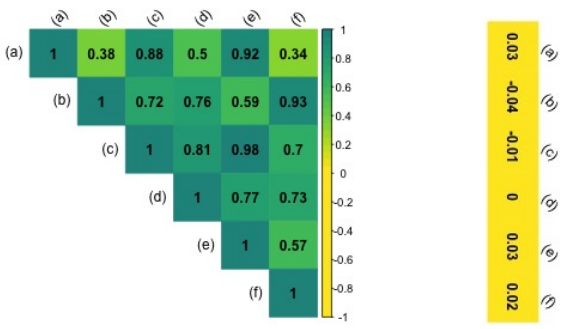
$$\boldsymbol{v}^T \boldsymbol{m} = \alpha \underbrace{\boldsymbol{v}^T \boldsymbol{M}}_{\boldsymbol{v}^T} \boldsymbol{m} + (1-\alpha)\boldsymbol{v}^T \boldsymbol{q} \ , \tag{4}$$

so that $\boldsymbol{v}^T \boldsymbol{m} = \boldsymbol{v}^T \boldsymbol{q} > 0$. This allows rewriting the steady-state equation (2) in the form

$$\boldsymbol{m} = \underbrace{\left[\alpha \boldsymbol{M} + (1-\alpha)\frac{\boldsymbol{q}\boldsymbol{v}^T}{\boldsymbol{v}^T\boldsymbol{q}}\right]}_{\boldsymbol{M}_1} \boldsymbol{m} \ , \tag{5}$$

where matrix $\boldsymbol{M}_1$ satisfies $\boldsymbol{v}^T \boldsymbol{M}_1 = \boldsymbol{v}^T$ by construction, i.e., $\boldsymbol{v}^T$ is a left eigenvector of $\boldsymbol{M}_1$ associated to eigenvalue 1. The corresponding right eigenvector, providing the solution $\boldsymbol{m}$, is instead different from $\boldsymbol{1}$. Since the irreducibility property on $\boldsymbol{M}$ implies that also $\boldsymbol{M}_1$ is irreducible, then the Perron-Frobenius theorem ensures that the eigenvalues of $\boldsymbol{M}_1$ satisfy $|\lambda| \leq 1$, and that eigenvalue 1 has multiplicity 1.

The *left* eigenvector $\boldsymbol{v}$ can be then used to characterize the *right* eigenvectors of $\boldsymbol{M}_1$. Let the Jordan form of $\boldsymbol{M}_1$ be $\boldsymbol{M}_1 = \boldsymbol{R}\boldsymbol{J}\boldsymbol{R}^{-1}$ where $\boldsymbol{R}$ collects the (generalized) right eigenvectors of $\boldsymbol{M}_1$, and, conversely, $\boldsymbol{R}^{-1}$ collects the right eigenvectors. We have



(a) correlation among methods (b) correlation with PageRank

Fig. 2. (a) Pearson's correlation matrix comparing projection methods and (b) their relation with PageRank centrality; all values are statistically significant.

$$\underbrace{\boldsymbol{v}^T \boldsymbol{M}_1}_{\boldsymbol{v}^T} \boldsymbol{R} = \boldsymbol{v}^T \boldsymbol{R}\boldsymbol{J} \ , \tag{6}$$

and therefore $\boldsymbol{v}^T \boldsymbol{R}(\boldsymbol{I} - \boldsymbol{J}) = \boldsymbol{0}$. Since in $\boldsymbol{M}_1$ there is only one eigenvalue equal to 1, this implies that all right eigenvectors, but the one related to $\boldsymbol{v}$, satisfy $\boldsymbol{v}^T \boldsymbol{r}_i = 0$. Hence, we have

$$\boldsymbol{M}_1 \boldsymbol{r}_i = \alpha \boldsymbol{M} \boldsymbol{r}_i + (1-\alpha)\frac{\boldsymbol{q}\boldsymbol{v}^T}{\boldsymbol{v}^T\boldsymbol{q}} \boldsymbol{r}_i = \alpha \boldsymbol{M} \boldsymbol{r}_i \tag{7}$$

so that the right (generalized) eigenvector $\boldsymbol{r}_i$ of $\boldsymbol{M}_1$ is also a right (generalized) eigenvector of $\alpha \boldsymbol{M}$, and as such it is related to an eigenvalue $|\lambda_i| \leq \alpha$. Thus, the eigenvalues of $\boldsymbol{M}_1$ include 1 as well as other eigenvalues with absolute value smaller than or equal to α. This ensures the convergence of (3) to the desired solution; observe the importance of setting $\boldsymbol{m}_0 = \boldsymbol{q}$ to guarantee $\boldsymbol{v}^T \boldsymbol{m}_k = \boldsymbol{v}^T \boldsymbol{q}$ throughout the iterations.

C. Algorithm comparison

To fully assess the validity of the PLMP method, we compare it with a few meaningful alternatives based on either [3] or the PageRank rationale. These further approaches added to those displayed in Fig. 1(a)–(c) are as follows.

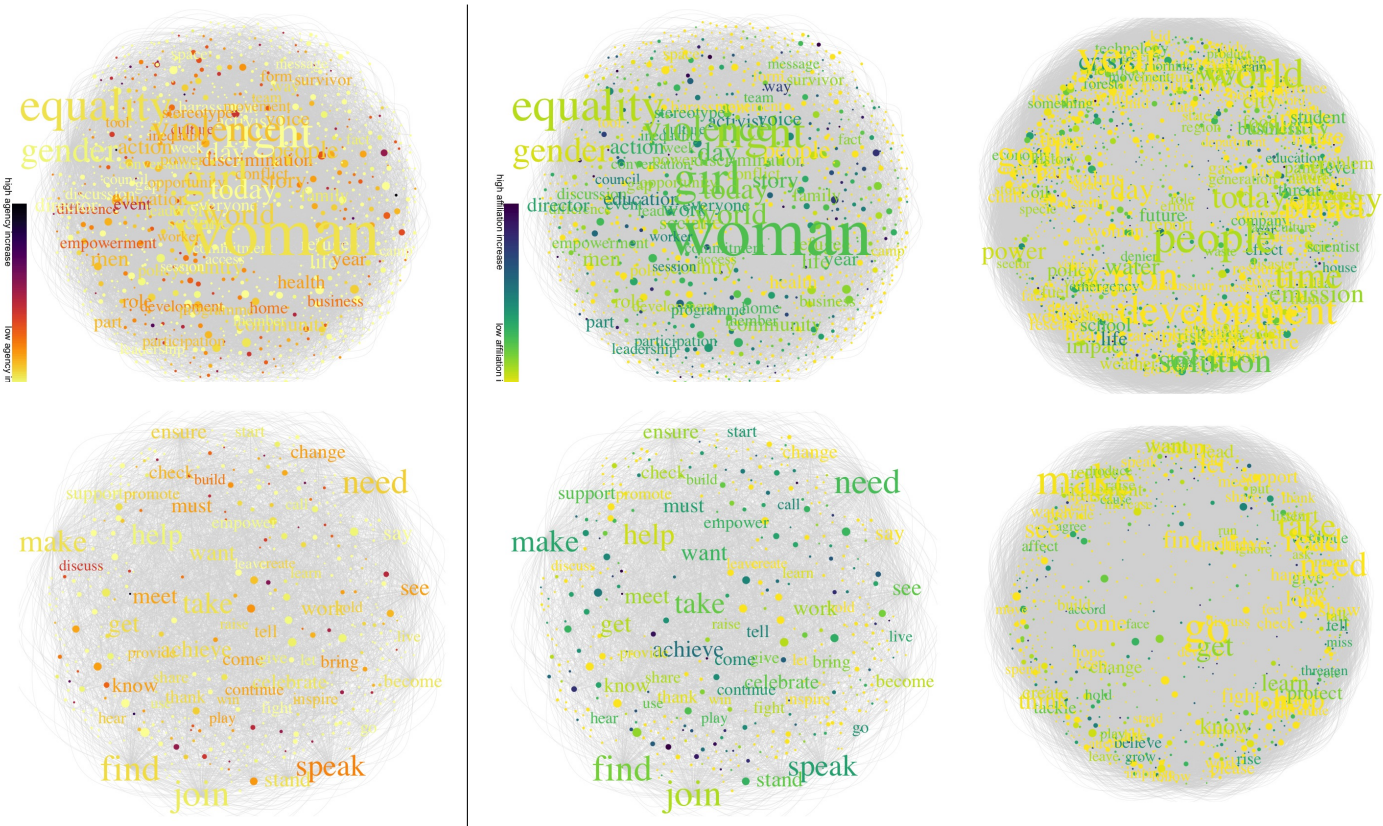
(d) We apply [3] to the sub-network of words whose (projected) adjacency matrix takes the form $\boldsymbol{M}_w = \boldsymbol{B}_1 \boldsymbol{B}_2$ [19], so that agency projection is inferred from $\boldsymbol{m}_w = \boldsymbol{M}_w \tilde{\boldsymbol{m}}_w$.

(e) If we generalize the inspiration from PageRank of (d), we get a steady-state equation $\boldsymbol{m}_w = \alpha \boldsymbol{M}_w \boldsymbol{m}_w + (1-\alpha)\tilde{\boldsymbol{m}}_w$, which gives rise to another different approach.

(f) Conversely, if we exploit $\tilde{\boldsymbol{m}}_t$ in a PageRank-like context, we obtain a projection $\boldsymbol{M}_t = \boldsymbol{B}_2 \boldsymbol{B}_1$ on tweets, and have

$$\begin{aligned} \boldsymbol{m}_t &= \alpha \boldsymbol{M}_t \boldsymbol{m}_t + (1-\alpha)\tilde{\boldsymbol{m}}_t \\ \boldsymbol{m}_w &= \boldsymbol{B}_1 \boldsymbol{m}_t \ . \end{aligned} \tag{8}$$

The inter-dependencies between the approaches (a)-(c) of Fig. 1 and the alternatives (d)–(f) are available in Fig. 2(a), showing Pearson's correlations between different options. The best balance between the *in isolation* information (a) and the *one hop* spreading [3] (b) is provided by the PLMP method.



(a) agency increase, #MeToo (b) affiliation increase, #MeToo (c) affiliation increase, #FridaysForFuture

Fig. 3. Graphical representation of the increase of agency (left) and affiliation (right) in both #MeToo and #FridaysForFuture, using the prestige measure (9) and the PLMP method: node size is proportional to the PageRank value in the considered network, while the color corresponds to the prestige level.

Fig. 2(b) further investigates the relation with the network's PageRank centrality, showing an absence of correlation that, especially for approaches based on a PageRank-like approach, is a guarantee of adequateness as it shows that the flow of agency through the network is not redundant with the nodes centrality. The consistency of Fig. 2 over different semantic networks recommends PLMP as the most reliable solution.

IV. Example: Calls to Action

An application of the PLMP method is illustrated in Fig. 3 with respect to the two fundamental markers of agency (left) and affiliation (right). Both outcomes for names and verbs are separately shown in Fig. 3. The semantic networks displayed refer to tweets published after the main event, and colors highlight the increase of agency with respect to the semantic network built on tweets published before the main event. Specifically, the color is set according to the *prestige* measure

$$p = \frac{m_{\text{post}} - m_{\text{pre}}}{m_{\text{post}} + m_{\text{pre}}}, \tag{9}$$

that is positive (dark color) if we observe an increase in the PLMP value (i.e., if the marker is bigger after the event); the darker the color, the more relevant the increase.

As shown in Table I, the average per-tweet levels of both agency and affiliation increase for both calls-to-action, with a slightly stronger trend for #MeToo, and all the increases being statistically relevant. Fig. 3 shows a difference in the way these increases relate to words. In #MeToo, the collective action markers increase is strongly associated with central words in the social discourse (e.g., *woman*, *speak*, *equality*, *change*, *achieve*, *violence*). Instead, most of the central words in #FridaysForFuture are not strongly associated to an increase in agency/affiliation (see their brighter colors). This effect is investigated in Fig. 4 showing a dependency between the PageRank centrality of words and their agency/affiliation increase, which is especially evident for highly ranked words. The environmentalist call to action #FridaysForFuture is characterized by an enhanced agency/affiliation in negative relation with PageRank ($B = -.34$), whereas the feminist call #MeToo features an increase in agency/affiliation as the words acquire importance in the discourse.

TABLE I
Impact of the Event on Agency and Affiliation

#MeToo	pre	post	variation
agency	1.67	1.83	+9.7%
affiliation	3.33	3.70	+10.9%
#FridaysForFuture	pre	post	variation
agency	1.56	1.65	+6.1%
affiliation	2.09	2.29	+9.5%

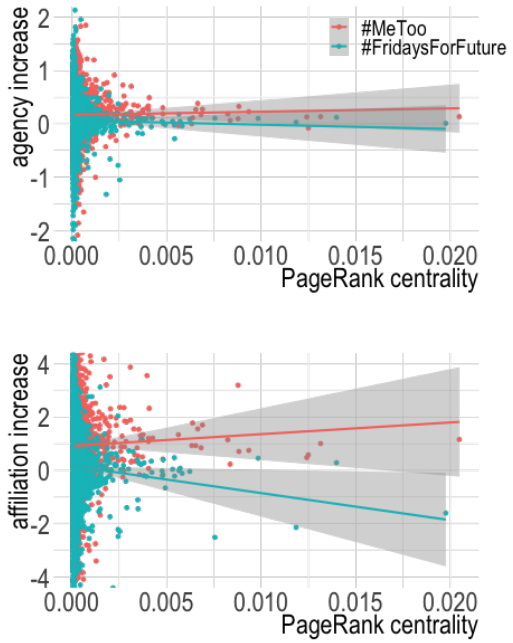


Fig. 4. PageRank centrality of words (after the main event) vs. agency (above) or affiliation (below) increases in #MeToo and #FridaysForFuture.

The statistical reliability of this result was verified by running a mixed full-factorial linear model [20] associating the *increase* in agency/affiliation (nested within words, which was included as a random factor) with the collective action (#MeToo or #FridaysForFuture) and the PageRank centrality of words post the event. This shows that, compared to #FridaysForFuture, #MeToo changed more over time, $F(1, 5391, 9) = 156.66$, $p < .001$. We also observed more change in affiliation than in agency, $F(1, 2455) = 131.15$, $p < .001$. The interaction between the two effects, $F(1, 5335.7) = 92.19$, $p < .001$, showed that the change over time is striking for the increase of affiliation in the #MeToo movement. Furthermore, the increase in both projected agency and affiliation is linked to PageRank in a different way according to the call to action, $F(1, 5464.5) = 5.01$, $p < .001$.

Figs. 3 and 4 capture a structural difference between the two calls. #FridaysForFuture appears as a sparser discourse that concentrates the increase in agency and affiliation on the periphery of the speech (lower PageRank centrality). #MeToo is instead characterized by very prominent hub words (e.g., *women*, associated to a PageRank value twice that of the hubs in #FridaysForFuture) embedded in the socio-psychological features of agency and affiliation. This result allows a finely tuned characterization of the rhetoric of movements, which may offer novel insights into social influence phenomena.

V. Conclusions

We proposed PLMP, a method to extrapolate holistic information from a semantic network by mapping linguistic markers onto the network elements, i.e., the words. PLMP exploits a PageRank-like rationale to capture the network interdependencies, hence better tracking the subtleties of the social discourse. Its application to the study of calls to collective action revealed that PLMP is able to characterize many call-specific aspects. Future studies may investigate if these properties are replicated in other calls focusing on human rights (like #MeToo) or scientific matters (like #FridaysForFuture). It is nevertheless relevant that, comparing the alternative methods of Section III-C, only the methods built upon a PageRank-like flow, namely PLMP and method (e), are able to capture this fine characterization in a statistically reliable way (i.e., $p < .05$ in the tests), confirming the validity of our approach.

References

[1] D. Boyd and K. Crawford, "Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon," *Inf. Commun. Soc.*, vol. 15, no. 5, pp. 662–679, 2012.

[2] K. Gimpel, N. Schneider, B. O'Connor, D. Das, D. Mills, J. Eisenstein, M. Heilman, D. Yogatama, J. Flanigan, and N. Smith, "Part-of-speech tagging for Twitter: Annotation, features, and experiments," in *Ann. Meet. Ass. Computat. Ling.*, 2011, vol. 2, pp. 42–47.

[3] L. Badia, D Clementel, T. Erseghe, L. Iacovissi, M. Migliorini, B. G. Salvador Casara, and C. Suitner, "Structural and semantic impact of online collective action," in *NetSci Conf.*, 2020.

[4] L. Page, S. Brin, R. Motwani, and T. Winograd, "The pagerank citation ranking: Bringing order to the web," Tech. Rep., Stanford, 1999.

[5] D. F. Gleich, "Pagerank beyond the web," *SIAM Review*, vol. 57, no. 3, pp. 321–363, 2015.

[6] T. Haveliwala and S. Kamvar, "The second eigenvalue of the google matrix," Tech. Rep., Stanford, 2003, http://ilpubs.stanford.edu:8090/582/.

[7] I. Hellsten and L. Leydesdorff, "Automated analysis of actor–topic networks on twitter: New approaches to the analysis of socio-semantic networks," *J. Assoc. Inf. Sc. Tech.*, vol. 71, no. 1, pp. 3–15, 2020.

[8] C. Suitner, L. Badia, D. Clementel, L. Iacovissi, M. Migliorini, B. G. Salvador Casara, D. Solimini, M. Formanowicz, and T. Erseghe, "The rise of #climateaction in the time of the fridaysforfuture movement: a semantic network analysis," *Soc. Netw.*, 2022, to be published.

[9] O. Owoputi, B. O'Connor, C. Dyer, K. Gimpel, N. Schneider, and N. Smith, "Improved part-of-speech tagging for online conversational text with word clusters," in *Proc. NAACL-HLT*, 2013, pp. 380–390.

[10] M. van Zomeren and A. Iyer, "Introduction to the social and psychological dynamics of collective action," *J. Soc. Issues*, vol. 65, no. 4, pp. 645–660, 2009.

[11] M. van Zomeren, M. Kutlaca, and F. Turner-Zwinkels, "Integrating who "we" are with what "we"(will not) stand for: A further extension of the social identity model of collective action," *Eur. Rev. Soc. Psych.*, vol. 29, no. 1, pp. 122–160, 2018.

[12] J. W. Pennebaker, R. L. Boyd, K. Jordan, and K. Blackburn, "The development and psychometric properties of LIWC2015," Tech. Rep., 2015.

[13] R. C. Hawkins II and R. L. Boyd, "Such stuff as dreams are made on: Dream language, LIWC norms, and personality correlates," *Dreaming*, vol. 27, no. 2, pp. 102–121, 2017.

[14] A. Pietraszkiewicz, M. Formanowicz, M. Gustafsson Sendén, R. L. Boyd, S. Sikström, and S. Sczesny, "The big two dictionaries: Capturing agency and communion in natural language," *Eur. J. Soc. Psych.*, vol. 49, no. 5, pp. 871–887, 2019.

[15] A. V. Guglielmi and L. Badia, "Social communication to improve group recognition in mobile networks," in *Proc. IEEE Globecom Wkshps*, 2015.

[16] T. Erseghe, "A distributed and scalable processing method based upon ADMM," *IEEE Sig. Proc. Lett.*, vol. 19, no. 9, pp. 563–566, 2012.

[17] Y. Fan, M. Li, P. Zhang, J. Wu, and Z. Di, "The effect of weight on community structure of networks," *Physica A*, vol. 378, no. 2, pp. 583–590, 2007.

[18] F. R. Gantmacher, *The Theory of Matrices, vol. 2*, AMS Chelsea Publishing Company, New York, 2000.

[19] T. Zhou, J. Ren, M. Medo, and Y. C. Zhang, "Bipartite network projection and personal recommendation," *Phys. Rev. E*, vol. 76, no. 4, pp. 046115, 2007.

[20] E. P. George, J. S. Hunter, W. G. Hunter, R. Bins, K. Kirlin IV, and D. Carroll, *Statistics for experimenters: Design, innovation, and discovery*, vol. 2, Wiley New York, NY, USA:, 2005.