On Designing a Time Sensitive Interaction Graph to Identify Twitter Opinion Leaders

Marco Furini
Università degli Studi di Modena e Reggio Emilia
Modena, Emilia Romagna, Italy
marco.furini@unimore.it

Luca Mariotti
Università degli Studi di Modena e Reggio Emilia
Modena, Emilia Romagna, Italy
mariottiluca1@hotmail.it

Riccardo Martoglia
Università degli Studi di Modena e Reggio Emilia
Modena, Emilia Romagna, Italy
riccardo.martoglia@unimore.it

Manuela Montangero
Università degli Studi di Modena e Reggio Emilia
Modena, Emilia Romagna, Italy
manuela.montangero@unimore.it

ABSTRACT

What happened on social media during the recent pandemic? Who was the opinion leader of the conversations? Who influenced whom? Were they medical doctors, ordinary people, scientific experts? Did health institutions play an important role in informing and updating citizens? Identifying opinion leaders within social platforms is of particular importance and, in this paper, we introduce the idea of a *time sensitive interaction graph* to identify opinion leaders within Twitter conversations. To evaluate our proposal, we focused on all the tweets posted on Twitter in the period 2020-21 and we considered just the ones that were Italian-written and were related to COVID-19. After mapping these tweets into the graph, we applied the PageRank algorithm to extract the opinion leaders of these conversations. Results show that our approach is effective in identifying opinion leaders and therefore it might be used to monitor the role that specific accounts (i.e., health authorities, politicians, city administrators) have within specific conversations.

KEYWORDS

Opinion leader, Twitter conversations, COVID-19

ACM Reference Format:

Marco Furini, Luca Mariotti, Riccardo Martoglia, and Manuela Montangero. 2022. On Designing a Time Sensitive Interaction Graph to Identify Twitter Opinion Leaders. In *Conference on Information Technology for Social Good (GoodIT'22), September 7–9, 2022, Limassol, Cyprus.* ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3524458.3547268

1 INTRODUCTION

People use social media for many different reasons (i.e., to get informed, to share opinions, to influence other people, to look for products) and for many different topics (i.e., TV-shows, stock-market, sports, politics, leisure, health) [2, 10, 11]. In particular, a recently observed phenomenon is the use of social media to talk about health-based topics [7, 25]: people make personal decisions based on what they read on social platforms [19], even though health professionals have warned that social media conversations might contain misleading medical information [16, 23]. This paper analyzes what people talked about during the COVID-19 period.

In December 2019 in Wuhan (China), a never-before-seen virus (later, named COVID-19) began to spread among the population. The health authorities were groping in the dark. Nobody knew how the virus was born, how it spreads and, above all, how to cure it. Within days, the virus caused thousands of deaths. Within a few weeks, the virus reached Italy, the first European country to be affected by the virus. The Italian population turned out to be defenseless. The first COVID-19 case was officially discovered on February 20, 2020. The infections rose exponentially, hospitals became saturated, and the number of dead grew. Italy entered lockdown (the first Western country to take this measure) on March 8, 2020. Grounded at their homes, citizens looked for information and used social platforms to try to understand more. What was written in posted conversations? Who did post tweets? Ordinary people? Politicians? Medical doctors? Health authorities? Being a respiratory virus, it is expected that people refer to medical authorities, but is it really so?

To understand all this, we analyze Italian-written conversations posted on Twitter from March 2020 to Dec 2021 with respect to the COVID-19 topic. Tweets have been the basis of many different analyses (from stock market prediction [4] to TV promotions [5], from geographical discovery [6] to cultural heritage promotion [15]). In this paper, building on past expertise in graph and social analytics in different contexts (e.g., cultural heritage [13–15, 21] and HRM [3, 9]) we use Twitter conversations to address the following research questions:

- RQ Who are the opinion leaders within these conversations? Are they medical doctors, ordinary people, scientific experts? Do health institutions play an important role in informing and updating citizens?

It is worth mentioning that opinion leaders are those who influence conversations with their posts. It is also important to emphasize that an opinion leader is not necessarily an influencer. Indeed, while the latter have a high number of followers and post generic or advertising tweets, the figure of the opinion leader is not linked to the number of followers, but to the number of interactions (e.g., likes, replies, retweets) that his posts got [12].

GoodIT'22, September 7–9, 2022, Limassol, Cyprus

ACM ISBN 978-1-4503-9284-6/22/09…$15.00
https://doi.org/10.1145/3524458.3547268

In this paper, we propose a method to map Twitter conversations into a graph with the final aim to analyze this graph to get the opinion leaders. We evaluate our proposal using a dataset of 13 million Italian-written tweets focused on COVID-19 and posted in from March 2020 to December 2021, and using the well-know Page-Rank algorithm to extract the opinion leaders from the graph. Results show that the proposed approach might be helpful to identify opinion leaders within specific time-windows and thus our approach might be used to monitor the role that specific accounts (i.e., health authorities, politicians, city administrators) have within specific conversations.

The reminder of the paper is organized as follows: in Section 2 we briefly review studies that focused on social media and health; in Section 3 we provide details of our proposal; Sections 4 and 5 present the experimental setting and the results obtained while evaluating our proposal over a Twitter dataset; Section 6 discusses the strengths and limits of the approach, finally conclusions are drawn in Section 7.

2 RELATED WORK

Ever since social networks became popular, many different studies proposed methods to identify the most influential people on Twitter. Initially, it was always thought that the most influential people were those with the largest number of followers, but it was soon realized that this is not true [22]. Fake accounts and/or farm-bot accounts alter the reality of numbers. For this reason, recent studies have focused on designing more effective methodologies in identifying the so-called opinion leaders [12].

Before describing some recent studies, it should be noted that in the literature there is no precise definition of opinion leaders. While many agree that an opinion leader is a person capable of influencing public opinion and the habits of others, some argue that the opinion leader is an expert on the topic they are dealing with [20], while others argue that an opinion leader might not have a specific knowledge of the subject [17]. Far from settling the debate, we simply observe that an opinion leader writes tweets that other users appreciate through likes, retweets, replies and mentions.

In literature, some proposals focused on measuring user influence through graph data structures and through specific algorithms. [26] proposed PersonalizedPageRank, a ranking algorithm to search for influencer of specific topics. [8] showed the effectiveness of the PageRank algorithm when used to search for the most popular users. With respect to the COVID-19 pandemic, many studies were conducted on the Twitter Social Network: [1] focused on Twitter Arabic and designed a graph structure where nodes are the users and arches with predefined weights represent the relationships between users; [18] focused on the #CoronavirusPandemic trend and measured the influence of Twitter users by combining the attributes of the users' profile with the underlying network type formed among users. [24] focused on a possible correlation between the number of Twitter mentions and the number of new COVID-19 cases. [12] proposed a novel metric to measure the influence a user might have on specific conversations and, therefore, to identify opinion leaders within Twitter conversations.

3 OUR PROPOSAL: THE TIME SENSITIVE INTERACTION GRAPH

Identifying opinion leaders within social platforms is of particular importance because it allows understanding who can influence the users' opinions. The idea that an opinion leader is a user with many followers is deeply wrong for several reasons: i) there are farm bots that sell packages of followers; ii) followers have no temporal connection with the account followed (several followers may no longer be active); iii) followers are related to the user not to the content posted (the so-called influencers post very different contents and what interests some users may not interest others).

To overcome these difficulties, in this paper we introduce the idea of *time sensitive interaction graph*, a graph appropriately designed to represent conversations within the Twitter platform. Thanks to this representation, it is possible to apply algorithms (such as PageRank) to identify users who engage (like, retweet, replies, mentions) more people. These users are the opinion leaders of that graph. Note that, since the graph is time and content sensitive, our approach allows identifying opinion leaders of specific topics within specific time windows (weeks, months, years).

The following properties are considered:

- *Authors and mentions:* we extract the Twitter account names of the authors of the tweets and of the accounts mentioned in the tweets;
- *Like count:* for each time window, we associate to each Twitter account the number of likes its tweets got;
- *Tweet classification:* we classify each tweet into one of the following non overlapping categories: reply, retweet, mention (i.e., containing a mention to another account), quote (i.e., containing a quote to a tweet of another account), tweet (i.e., a tweet not falling into any of the previous categories);
- *Relations:* according to tweet classification, we build directed binary relations between pairs of accounts to highlight interactions between them: account Ta_1 is associated to Ta_2 if Ta_1 replied (resp. retweeted, quoted) to a (resp. a, a) tweet authored by Ta_2. Analogously, Ta_1 is associated to Ta_2 if Ta_1 mentioned Ta_2 in at least one of its tweets. Also relations are defined for each time window.

The *time sensitive interaction graph* is then built in the following way:

- There is one node for each Twitter account being author of a tweet or being mentioned in a tweet. Nodes label is *User*;
- Fig 1 shows the properties associated to nodes. There are some self-explaining properties, and then we have one property for each time window under consideration. The value of these properties count the number of likes received by the tweets posted by the account in that specific time window;
- There is a directed edge between user U_1 and user U_2 if their corresponding accounts are in a relation given by reply, retweet, mention, or quote, as explained before. Edges label is *Interacted*;
- Fig 1 shows the properties associated to edges. Again, total counts of interactions (by type) are also split by time window counts.

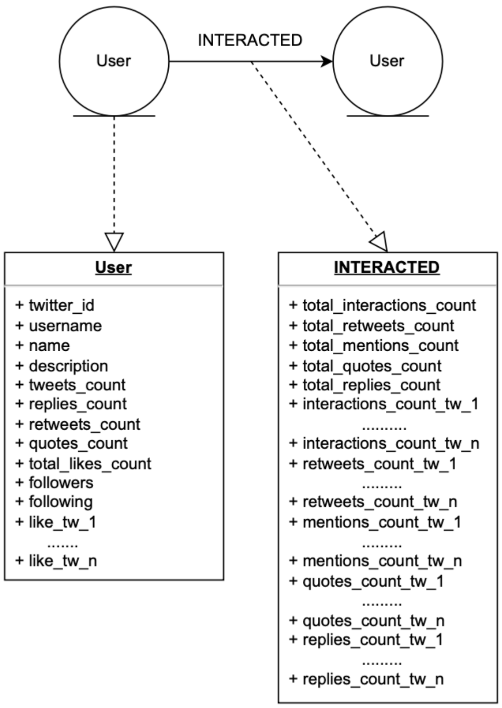

Figure 1: Data Base graph scheme (tw stands for time-window).

Once the graph has been built, it can be analyzed to determine how users' opinions influence others as time goes by, where time units are given by time windows.

4 EXPERIMENTS

To evaluate our proposal, we consider a dataset of 13M ca. tweets (Italian-written and Covid-19 related) posted on Twitter from March 2020 to December 2021. We analyzed this dataset month by month (i.e., we built a time sensitive interaction graph for each month). Then, we applied to these graphs the PageRank algorithm to identify the opinion leaders.

4.1 Data set and interaction graph construction

The data set is composed of almost 13 millions tweets in Italian that are related to the Covid-19 pandemic and have been posted between March 2020 and December 2021 (5.5M tweets ca. in 2020, and 7.5M ca. in 2021). These tweets are linked to the Covid-19 topic as they contain one or more then one of the following hashtags: #coronavirus, #covid19, #covid_19, #sarscov2, #astrazeneca, #pfizer, #moderna, #vaccino (#vaccine), #vaccini (#vaccines), #vaccinazione (#vaccination), #vaccinazioni (#vaccinations), #vax, #novax, #greenpass, #nogreenpass, #terzadose (#thirddose), #mrna.

Figure 2 shows the monthly tweet distribution. We observe peaks in the tweets counts in correspondence with particular pandemic events in Italy. On March 4th, the Italian government forced all the Italian schools to close; on March 9 the whole Italy went to **lockdown**. People found themselves in a situation never seen before

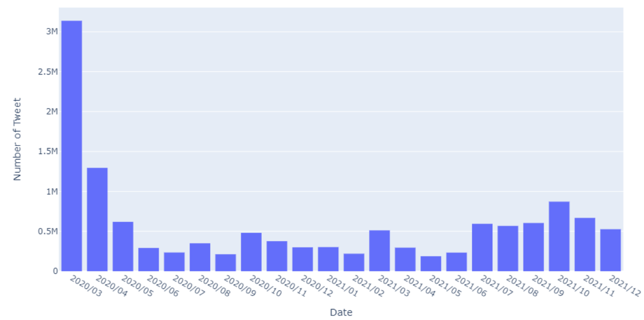

Figure 2: Tweets distribution over time, by monthly counts.

and everybody resorted to social platforms to get news about the new virus (ca 1,137K tweets). Lockdown lasted until the 18th of May 2020 and then the pandemic situation seemed to get better. **August** is the traditional month for **holidays** in Italy, and people had to adjust to new **restrictions** ruling holidays activities. Schools reopened in September, but in **October** 2020 the **second wave** started and Covid-19 related numbers started to grow again. With the beginning of **2021** the **vaccine campaign** began, organized in four phases (according to the categories of people that had right to get vaccinate): from January for healthcare personnel and elderly people), from March for educational personnel, fragile people, over 60s), form April for over 18s, from July for teenagers over 13. **March** also saw the beginning of the **third wave** and lockdowns in many northern regions. In **July** the Italian government announced the introduction of the **green pass** starting **August** 6th to access public places, and in September all working (public and private) places starting **October** 15th. Finally, **November** sees the announcement of the **third dose** for over 40s starting December the 1st and the beginning of the **fourth wave**.

Figure 3 shows the tweet classification counts for 2020 and 2021 separately. We observe that, in both cases, the majority of tweets are retweets of other's account tweets. On average, each original tweet was retweeted about three times (resp. almost five times) in 2020 (resp. in 2021). The other categories are definitively smaller, accounting for about 11% of the total in both years, mentions and quotes being the most populated classes.

The distribution of tweets in categories highlights that a large number of users was very active in retweeting others' contents, but only a small number of them wrote original contents. Other types of interactions were limited.

To have more detailed insights, we generated the interaction graph for each year (2020 and 2021) separately, using Twitter APIs and Neo4j graph database management system. In 2020 (resp. 2021), we have 450,703 (resp. 265,752) nodes and 2,992,056 (resp. 2,340,368) edges. Time windows were set to calendar months, therefore nodes have 23 properties and relations 65.

4.2 Tests

Here, we present our experiments to determine the Italian Twitter opinion leaders during the COVID-19 pandemic in the years 2020 and 2021. After building the time sensitive interaction graphs, we

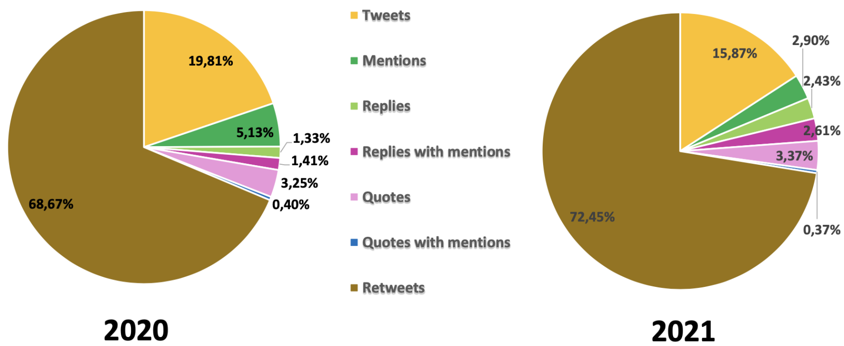

Figure 3: Tweets categories in 2020 (left) and 2021 (right).

applied the PageRank algorithm to it. PageRank assigns a numerical weight to the nodes of the graph to measure the relative importance of a node within the set of nodes, by exploiting relations between nodes as given by the edges of the graph. In our case, importance is *transferred* from one node to another with interactions directed from the former to the latter. Therefore, the most important nodes are those that made others interact with them, i.e., opinion leaders. The Neo4j PageRank implementation assigns a weight to edges (i.e., relations): we used the total number of interactions in the given time window.

5 RESULTS

The results represent the degree of influence of user accounts on a month-by-month basis. Therefore, they might be used to highlight the trend of Italian opinion leaders during the pandemic period, giving an answer to our research question RQ. The analysis of the results is divided into two sections, in Section 5.1 the most important results for the year 2020 are analyzed, while Section 5.2 focuses on year 2021.

5.1 Year 2020

Looking at the situation as a whole (Figure 4 left side, showing the PageRank value for all users in 2020), it is possible to observe that, in the month of March, the scores are on average much higher than in the other months of the year. On 11 March the status of a Pandemic for the new Coronavirus SARS-Cov-2 was declared by the **World Health Organization**, therefore the month sees the highest number of tweets posted by users (more than four gigabytes). The politics ranking of the month sees the Italian **Prime Minister** at that time at first place (Figure 5 left side), followed by **Right-Wing Politicians**, **Minister of Health** and **Left-Wing Politicians**. As regards health institutions and other institutes whose task is to safeguard the health of citizens, **Ministry of Health**, **Civil Protection Department**, **World Health Organization**, **Italian Red Cross** and **Higher Institute of Health** were fundamental in March and throughout the rest of the year, even if with a lower score (Figure 6 left side). Being the month of March the first official month of Pandemic, this explains the importance obtained by the above mentioned institutions. Other institutions, including **Palazzo Chigi** (Italian Government), **Quirinale** (residence of the President of the Italian Republic), **Viminale** (Headquarters of the Ministry of the Interior) and **Farnesina** (Ministry of Foreign Affairs and International Cooperation) get high marks, together with the **Embassy of the People's Republic of China in Italy**: As the Chinese government sent medical equipment and a team of medical experts to Italy on March 12 to help manage the pandemic, there was a high level of interaction and appreciation towards the embassy (Figure 7 left side). In April, the second month of Pandemic, the average score was halved and the ranking still sees the **Prime Minister** among the first accounts. During the month of April, the Prime Minister held a multitude of press conferences aimed at keeping citizens updated on the evolution of regulations and the pandemic situation. The importance of the **Prime Minister** gradually decreases until the month of September, as shown in Figure 5 left side, and then increases again (among first three accounts) in the months of October and November, months in which the Prime Minister signed two new Ministerial Decrees. As to influential non-institutional accounts, the account that appears to dominate the ranking (even w.r.t. other kinds of accounts) from July to December is the one of **N. C.**, surgeon and president of a well-known Italian scientific foundation (Figure 8 left side). This means that this account played a central role in the Italian Twitter scene during the pandemic in 2020 and, thanks to his qualifications and skills, it was the point of reference for the multitude of people who found themselves facing a pandemic for the first time in their life. An important role, especially in the initial months of the pandemic, was also played by an Italian virologist, immunologist and academic **R. B.**. (see Figure 8 left side, comparing this and No-Vax user account scores). Figure 9 left side, focuses on the Italian News Agencies and newspapers: among newspapers, **Repubblica** newspaper was the one with which users interacted the most, making it the favorite for disseminating news about the Pandemic. With less importance, but still relevant, follow **Sky Tg24**, **Il Sole 24 Ore**, **RaiNews** and **Corriere della Sera**, the latter especially significant in April.

5.2 Year 2021

The graph on the right side of Figure 4 has a very different shape compared to the one of the previous year (on the left side), highlighting the fact that there is no particular month in which the popularity of users is concentrated: as we can see, there are high

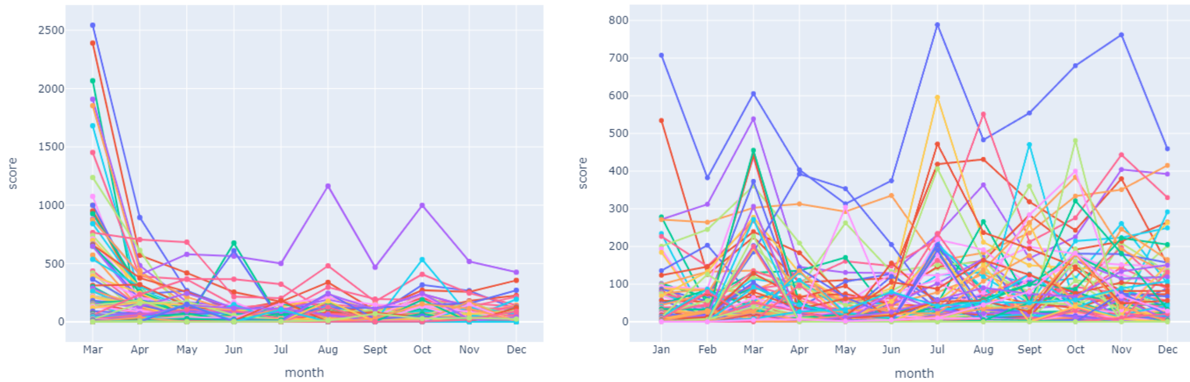

Figure 4: Overall PageRank results for years: 2020 on left, 2021 on right (colored lines represent different user accounts).

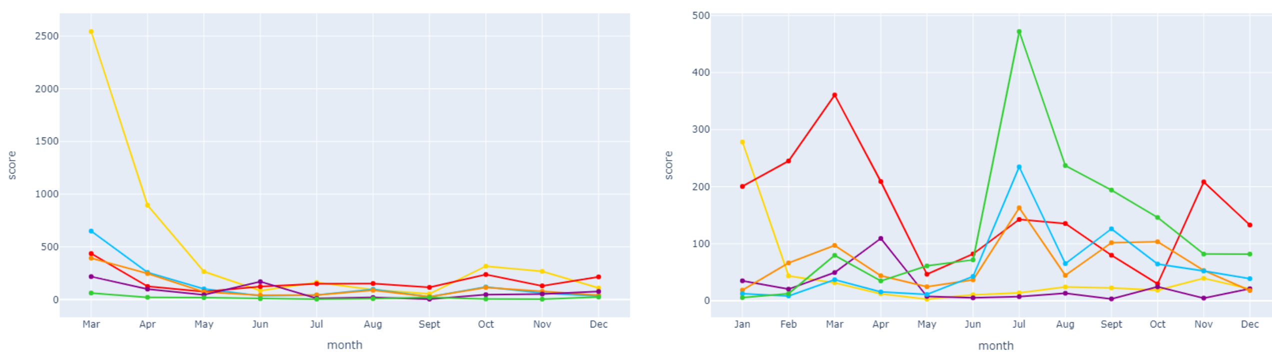

Figure 5: PageRank results for Italian politicians accounts in years 2020 (left) and 2021 (right): Prime Minister (yellow), Italian Minister for Health (red), Right-Wing Politicians (orange - blue - green) and Left-Wing Politicians (purple).

popularity values for all months, except for April, May and June. Talking about politicians, the Italian **Prime Minister** makes his last appearance in January, the month in which the Senate confirmed its trust in him, to be outclassed in the following months by right-wing politicians, opposed to the obligation of masks and green pass (Figure 5 right side). The Italian **Minister of Health** account obtained a good level of interaction in the months of February, March and April, therefore in full second wave of Covid-19. The **Ministry of Health** and **Higher Institute of Health** obtained a good response in terms of interactions, almost always placing themselves in the upper part of the ranking (Figure 6 right side); this because the vaccination plan for Italy was also announced in March. Figure 7 shows that in March the **Viminale** and **Quirinale** had an important role, because of announcing the rules about red and orange zones (rules dedicated to the containment of the coronavirus) and the imminent start of the third wave. As for 2020, 2021 sees the **N. C.** account at the top of the ranking, confirming its authority and role of reference for Twitter users in the pandemic period (Figure 8 right side). The figure highlights, unlike 2020, a strong importance given to No-Vax and No-GreenPass users **V. S.** and **F. M.**, almost reaching the popularity of medical doctors such as **N. C.**. As to Italian news agencies and newspapers, **Repubblica** confirmed its central role, followed by **Corriere della Sera**, **SkyTg24** and **Il Sole 24 Ore** (Figure 9 right side). Generally speaking, on average in 2021 almost all accounts obtained a lower value of importance than in 2020 (this can be seen, for instance, by comparing in Figure 4 the months of March, where the difference is particularly evident). The only users who have achieved an increase in their importance compared to the previous year are the No-Vax and No-GreenPass accounts, reaching a high level of audience (Figure 8).

5.3 Overall considerations

Summing up, during 2020, the central role that professors, medical doctors, institutions and newspapers have had in the dissemination of information and procedures regarding the Coronavirus is evident. During the Lockdown of March-May 2020, the medical doctors **N. C.** and **R. B.**, the **Prime Minister** and **Repubblica** were the most

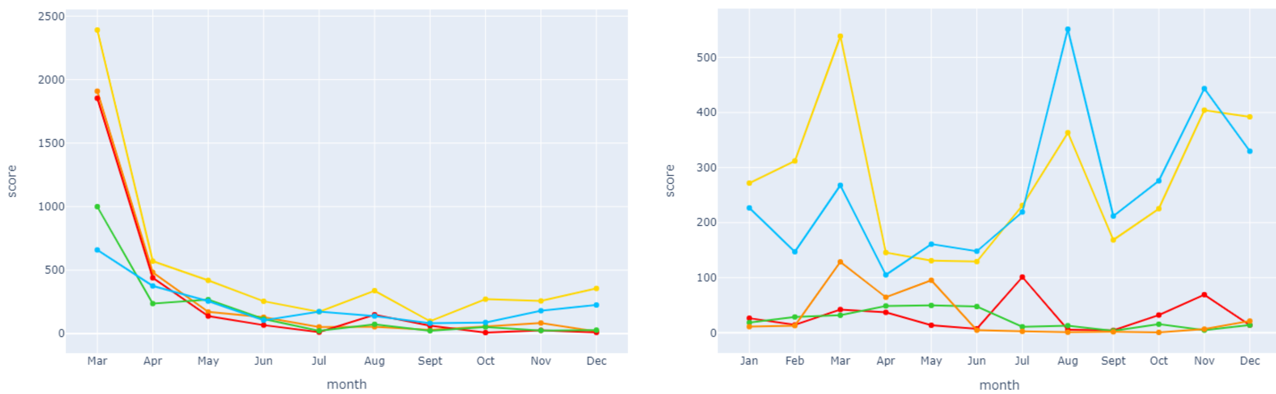

Figure 6: PageRank results for Health-related institutions accounts in years 2020 (left) and 2021 (right): World Health Organization (red), Civil Protection Department (orange), Ministry of Health (yellow), Higher Institute of Health (light blue) and Italian Red Cross (green).

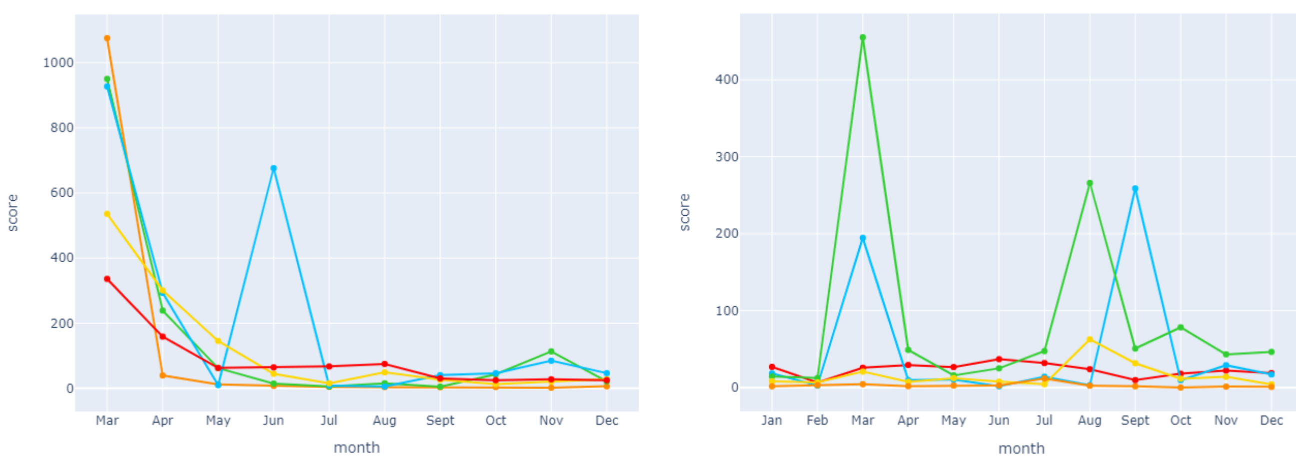

Figure 7: PageRank results for Government Institutions accounts in years 2020 (left) and 2021 (right): Embassy people's Republic of China in Italy (orange), Palazzo Chigi (green), Il Viminale (yellow), Farnesina (red) and Quirinale (light blue).

appreciated accounts by Twitter users, leading people in the battle against the common enemy, the virus. The relevance of these users has continued until the end of the first phase of the second wave of Covid-19 (October 2020 - February 2021). Since December 2020, exponents of philosophies of thought contrary to medical doctors, virologists and institutions, have gradually gained popularity on the social network, gaining a following of people whose trust in medical doctors or institutions is wavering. The most influential account among them was **V. S.**, who from December 2020, the month in which we started talking about vaccines, until the end of the year 2021, has always been one of the first three most influential users month by month, gaining an ever-growing following. Indeed, since the introduction of the green pass, a multitude of users opposed to government policies and vaccinations gained popularity, finding enormous success and outclassing medical doctors and institutions in the dissemination of news and information. The surgeon **N. C.** despite the growth in importance of the counterpart, has nevertheless maintained a high level of influence, highlighting in fact that a good part of users still have qualified and competent people as medical doctors as a point of reference.

6 DISCUSSION AND LIMITS

Results can certainly be interesting *per se*, as a novel data-driven interpretation of opinion-leader trends in the pandemic era, but

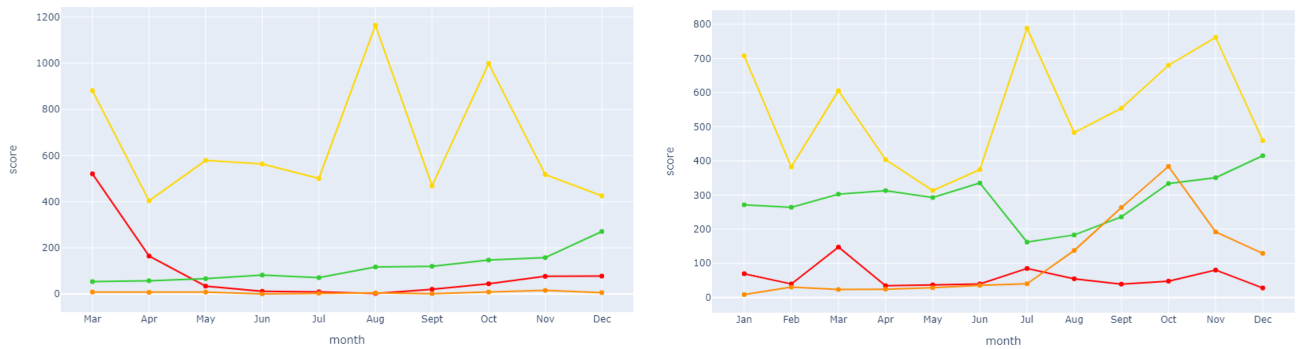

Figure 8: PageRank results in years 2020 (left) and 2021 (right) for influential non-institutional accounts: medics N. C. (yellow), R. B. (red) and no-vax users V. S. (green), F. M. (orange).

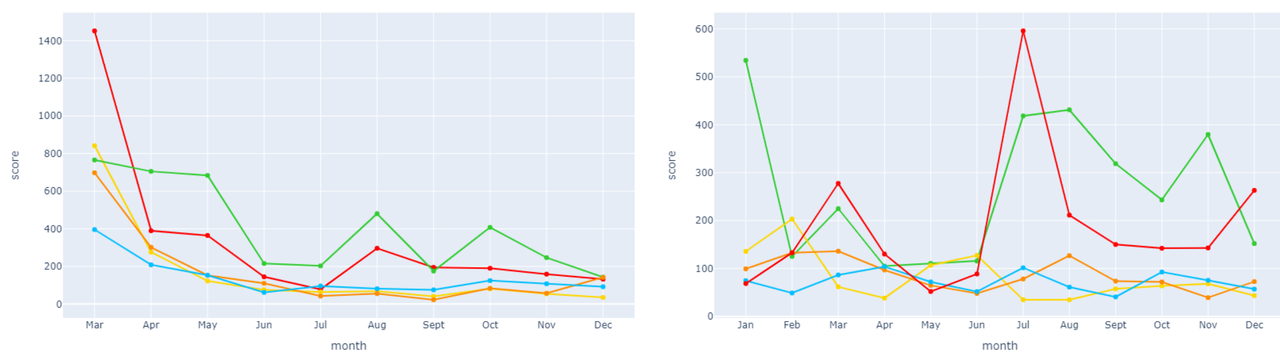

Figure 9: PageRank results for Italian news agencies accounts in years 2020 (left) and 2021 (right): Corriere della Sera (red), Il Sole 24 Ore (yellow), Sky Tg24 (orange), RaiNews (light blue) and Repubblica (green)

they are also key in testifying the importance of the employed methodology from different points of view:

- the methodology is primarily focused on discovering opinion leaders: this is the reason why the graph is designed in such a way to capture the appreciation other users give to messages (e.g., through retweets, reply, mentions);
- results are not directly influenced by the number of messages user accounts write (the messages contribute only if they generate some sort of reaction), allowing the analysis to go beyond a limited and often misleading "tweet count". The same applies to the number of followers of the involved accounts, which are not even considered in our method (differently from the many proposals simply focusing on discovering social network influencers);
- by using a graph and a PageRank algorithm, we are able to take reactions of users into account in a transitive way, going beyond "one-step" interactions between accounts;
- thanks to the temporal nature of the graph, the temporal dimension becomes crucial to the analysis, enabling the identification of several opinion-leader trends related to the different temporal periods.

Note that this is an initial proposal and we are also aware of some limits in the presented approach that we plan to address in future work: the employed PageRank algorithm does not allow full exploitation of the data embedded in the graph. Thus, we will consider additional algorithms that will enable for instance to exploit the number of likes (together with the already considered number of interactions) in nodes / edges weights. Moreover, we are aware of the importance of bringing into the discussion the textual content of the tweets, in order not only to discover opinion leaders but also the reasons why they became so.

7 CONCLUSIONS

In this paper, we proposed a novel approach to identify opinion leaders within the Twitter platforms. Our method is based on the design of a time sensitive interaction graph and on algorithms able to analyze such graph to extract users who were able to engage a large number of users. The obtained results showed the effectiveness of our proposal and therefore our approach might be considered as a first step towards a system capable of monitoring and identifying opinion leaders in real time.

REFERENCES

[1] Sarah Alqurashi, Abdulaziz Alashaikh, Mahmoud Nabil, Eisa Alanazi, Ahmed Sherif, and Waleed Alsmary. 2021. Identifying Information Superspreaders of COVID-19 from Arabic Tweets. In *2021 3rd IEEE Middle East and North Africa COMMunications Conference (MENACOMM)*. 109–114. https://doi.org/10.1109/MENACOMM50742.2021.9678304

[2] Onur Asan and Pascale Carayon. 2017. Human Factors of Health Information Technology—Challenges and Opportunities. *International Journal of Human–Computer Interaction* 33, 4 (2017), 255–257. https://doi.org/10.1080/10447318.2017.1282755

[3] Fabiola Bertolotti, Tommaso Fabbri, Federica Mandreoli, Riccardo Martoglia, and Annachiara Scapolan. 2020. Work datafication and digital work behavior analysis as a source of social good. In *IEEE Consumer Communications and Networking Conference (CCNC)*. IEEE.

[4] Armir Bujari, Marco Furini, and Nicolas Laina. 2017. On using cashtags to predict companies stock trends. In *2017 14th IEEE Annual Consumer Communications Networking Conference (CCNC)*. 25–28. https://doi.org/10.1109/CCNC.2017.7983075

[5] Roberta De Michele, Stefano Ferretti, and Marco Furini. 2019. On helping broadcasters to promote TV-shows through hashtags. *Multimedia Tools and Applications* 78, 3 (2019), 3279–3296. https://doi.org/10.1007/s11042-018-6510-7

[6] Roberta De Michele and Marco Furini. 2018. Smart City and Images: The Use of Image Hashtags to Get Insights on Citizens. In *Smart Objects and Technologies for Social Good*. 311–320.

[7] Mark Dredze. 2012. How Social Media Will Change Public Health. *IEEE Intelligent Systems* 27, 4 (July 2012), 81–84. https://doi.org/10.1109/MIS.2012.76

[8] Marcel Dunaiski, Jaco Geldenhuys, and Willem Visser. 2018. Author ranking evaluation at scale. *Journal of Informetrics* 12, 3 (2018), 679–702. https://doi.org/10.1016/j.joi.2018.06.004

[9] Tommaso Fabbri, Federica Mandreoli, Riccardo Martoglia, and Annachiara Scapolan. 2019. Employee attitudes and (digital) collaboration data: A preliminary analysis in the HRM field. In *International Workshop on Social Media Sensing (SMS'19 @ IEEE ICCCN)* (2019-08-05). IEEE.

[10] Stefano Ferretti, Marco Furini, Claudio E. Palazzi, Marco Roccetti, and Paola Salomoni. 2010. WWW recycling for a better world. *Commun. ACM* 53, 4 (2010), 139–143. https://doi.org/10.1145/1721654.1721692

[11] Marco Furini. 2021. Identifying the features of ProVax and NoVax groups from social media conversations. *Computers in Human Behavior* 120 (2021). https://doi.org/10.1016/j.chb.2021.106751

[12] Marco Furini and Emilia Flisi. 2022. Opinion Leaders and Twitter: Metric Proposal and Psycholinguistic Analysis. In *IEEE Symposium on Computers and Communications (ISCC)*.

[13] Marco Furini, Federica Mandreoli, Riccardo Martoglia, and Manuela Montangero. 2017. The Use of Hashtags in the Promotion of Art Exhibitions. In *Proc. 13th Italian Research Conference on Digital Libraries (IRCDL), Revised Selected Papers*. 187–198.

[14] Marco Furini, Federica Mandreoli, Riccardo Martoglia, and Manuela Montangero. 2018. Towards Tweet Content Suggestions for Museum Media Managers. In *Proc. of GOODTECHS18*. Association for Computing Machinery, 265–270.

[15] Marco Furini, Federica Mandreoli, Riccardo Martoglia, and Manuela Montangero. 2022. A Predictive Method to Improve the Effectiveness of Twitter Communication in a Cultural Heritage Scenario. *J. Comput. Cult. Herit.* 15, 2, Article 21 (apr 2022), 18 pages. https://doi.org/10.1145/3470786

[16] Marco Furini and Gabriele Menegoni. 2018. Public Health and Social Media: Language Analysis of Vaccine Conversations. In *2018 International Workshop on Social Sensing (SocialSens)*. 50–55. https://doi.org/10.1109/SocialSens.2018.00022

[17] Andrew J. Grimes and Phillip K. Berger. 1970. Cosmopolitan-Local: Evaluation of the Construct. *Administrative Science Quarterly* (1970).

[18] Somya Jain and Adwitiya Sinha. 2020. Identification of influential users on Twitter: A novel weighted correlated influence measure for Covid-19. *Chaos, Solitons and Fractals* 139 (2020), 110037. https://doi.org/10.1016/j.chaos.2020.110037

[19] S. Mo Jang, Brooke W. Mckeever, Robert Mckeever, and Joon Kyoung Kim. 2019. From Social Media to Mainstream News: The Information Flow of the Vaccine-Autism Controversy in the US, Canada, and the UK. *Health Communication* 34, 1 (2019), 110–117. https://doi.org/10.1080/10410236.2017.1384433

[20] Yanyan Li, Shaoqian Ma, Yonghe Zhang, Ronghuai Huang, and Kinshuk. 2013. An improved mix framework for opinion leader identification in online learning communities. *Knowledge-Based Systems* 43 (2013). https://doi.org/10.1016/j.knosys.2013.01.005

[21] Riccardo Martoglia and Manuela Montangero. 2020. An Intelligent Dashboard for Assisted Tweet Composition in the Cultural Heritage Area (Work-in-progress). In *Proc. of GOODTECHS20*. Association for Computing Machinery, 226–229.

[22] Manuela Montangero and Marco Furini. 2015. TRank: Ranking Twitter users according to specific topics. In *2015 12th Annual IEEE Consumer Communications and Networking Conference (CCNC)*. 767–772. https://doi.org/10.1109/CCNC.2015.7158074

[23] Marcella Tambuscio, Giancarlo Ruffo, Alessandro Flammini, and Filippo Menczer. 2015. Fact-checking Effect on Viral Hoaxes: A Model of Misinformation Spread in Social Networks. In *24th WWW Conference* (Florence, Italy). 6 pages. https://doi.org/10.1145/2740908.2742572

[24] Zainab Tariq Soomro, Sardar Haider Waseem Ilyas, and Ussama Yaqub. 2020. Sentiment, Count and Cases: Analysis of Twitter discussions during COVID-19 Pandemic. In *2020 7th International Conference on Behavioural and Social Computing (BESC)*. 1–4. https://doi.org/10.1109/BESC51023.2020.9348291

[25] Ryen W. White and Eric Horvitz. 2009. Cyberchondria: Studies of the Escalation of Medical Concerns in Web Search. *ACM Trans. Inf. Syst.* 27, 4 (Nov. 2009), 23:1–23:37. https://doi.org/10.1145/1629096.1629101

[26] Zeynep Zengin Alp and Şule Gündüz Öğüdücü. 2018. Identifying topical influencers on twitter based on user behavior and network topology. *Knowledge-Based Systems* 141 (2018), 211–221. https://doi.org/10.1016/j.knosys.2017.11.021