

Received October 3, 2020, accepted October 5, 2020, date of publication October 14, 2020, date of current version October 28, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3031014

The Dynamical Interplay of Collective Attention, Awareness and Epidemics Spreading in the Multiplex Social Networks During COVID-19

MARIALISA SCATÁ^{ID}, (Member, IEEE), BARBARA ATTANASIO^{ID}, (Member, IEEE),
GRAZIA VERONICA AIOSA^{ID}, AND AURELIO LA CORTE^{ID}

Dipartimento di Ingegneria Elettrica, Elettronica ed Informatica (DIEEI), Università di Catania, 95125 Catania, Italy

Corresponding author: Marialisa Scatá (lisa.scata@dieei.unict.it)

This work was supported in part by the Italian Ministry of University and Research (MIUR)–PON REC 2014–2020 (TALISMAN) under Project ARS01_01116, and in part by the project MANGO–Piaceri 2020–2022–Linea 2, Università di Catania.

ABSTRACT Leveraging social and communication technologies, we can digitally observe that the collective attention typically exhibits a heterogeneous structure. It shows that people's interests are organized in clusters around different topics, but the rising of an extraordinary emergency event, as the coronavirus disease epidemics, channels the people's attention into a more homogenized structure, shifting it as triggered by a non-random collective process. The connectedness of networked individuals, on multiple social levels, impacts on the attention, representing a tuning element of different behavioural outcomes, changing the awareness diffusion enough to produce effects on epidemics spreading. We propose a mathematical framework to model the interplay between the collective attention and the co-evolving processes of awareness diffusion, modelled as a social contagion phenomenon, and epidemic spreading on weighted multiplex networks. Our proposed modeling approach structures a systematically understanding as a social network marker of interdependent collective dynamics through the introduction of the multiplex dimension of both networked individuals and topics, quantifying the role of human-related factors, as homophily, network properties, and heterogeneity. We introduce a data-driven approach by integrating different types of data, digitally traced as user-generated data from Twitter and Google Trends, in response to an extraordinary emergency event as coronavirus disease. Our findings demonstrate how the proposed model allows us to quantify the reaction of the collective attention, proving that it can represent a social predictive marker of the awareness dynamics, unveiling the impact on epidemic spreading, for a timely crisis response planning. Simulations results shed light on the coherence between the data-driven approach and the proposed analytical model.

INDEX TERMS Multiplex networks, social networks, epidemics spreading, social contagion, collective attention, collective awareness, COVID-19.

I. INTRODUCTION

The outbreak of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), a novel coronavirus, emerged in the city of Wuhan (Hubei, China) in early December of 2019, has posed a global public interest and it has raised concerns in most people worldwide on the future health and well-being [1]–[3]. This new betacoronavirus have shocked the world and has drawn global attention since it has caused a

The associate editor coordinating the review of this manuscript and approving it for publication was Derek Abbott^{ID}.

fast diffusion. The groups of cases of pneumonia of unknown cause detected in Wuhan was reported to the World Health Organization (WHO) on the 31st December 2019, and consequently this new virus was named 2019 novel coronavirus (2019-nCoV) [4] and SARS-CoV-2, by the International Committee on Taxonomy of Virus (ICTV) on the 11th February 2020, and the diseases caused by SARS-CoV-2 were named Corona Virus Diseases 2019 (COVID-19) on the same day [1], [4]. Since then, besides China, cases confirmed with COVID-19 had also been detected in many other countries and territories [4], [5], thus, from the city of Wuhan that

has taken unprecedented measures in response to the outbreak, including extended school and workplace closures, also the other countries follows these strategies through various national responses to the pandemic, including measures such as lockdowns, quarantines, curfews and other restrictions (stay-at-home orders, shelter-in-place orders, shutdowns/lockdowns) aimed at preventing further spread of COVID-19. On 30th January 2020 WHO officially proclaimed the outbreak of COVID-19 as a public health emergency of international concern, and establishes temporary recommendations, early indications, giving a boost to a different reaction, especially in terms of timing. The public health measures for managing the possible outbreaks count on the preparedness ability of each country, namely the capacity to prevent, detect, verify, assess, and react in accordance with the International Health Regulations for COVID-19 [6], [7]. The majority of methods, which take into considerations evidence-based public health prevention strategies weighs the people's interests and the ongoing development consciousness which makes it possible to shape the public awareness on behaviours [6]. The global public attention to this issue can reflect people's interests and their propensity to acquire knowledge on COVID-19 with the aim at taking precautionary actions. In fact, to understand an emergency event, it is insufficient to rely on numerical reports based only on confirmed cases and the spatial spread [8], but it is needed adding more information, integrating knowledge and data on people's behaviours to quantify other dominant variables that can influence and impact possible future outbreaks. Exploiting the social networks and communication technologies under this extraordinary emergency condition, a systematic understanding of social collective dynamics, based on the complex networks paradigms that integrate weighted multiplex networks mathematical representation, can be achieved. To do this, we need to detect how the public attention around COVID-19, could guide the situational awareness and the timing of intervention strategies, impacting on the epidemics spreading. A large corpus of scientific literature have studied the phenomenon of spreading deepen inside a better understanding on how it produces dynamics which depends on various structural and human-related factors [9], [10], suggesting that in social networks the dynamics depend on the nature of social ties [11]. Since behaviours [11], [12], misinformation [13], [14], infectious diseases [15], [16], distress [17], emotions [18] and competing processes [19], [20] spread through interactions of networked individuals, starting from classical epidemiological models [21]–[23], and based on an interdisciplinary approach which involves several research fields in complex network science [24]–[27], we are able to model the diffusion dynamics into the social networks. Taking into consideration the modeling approach of dynamics of social contagion and spreading processes in complex networks [16], [18], [28], we are interested to grasp the linkage among the shape of collective phenomena and their interpersonal spreading in networks structure, in response to a triggering force as an extraordinary event

able to shift the common outcomes to unexpected peaks and clusters. Many studies have investigated the spreading phenomena and its co-evolution, as the awareness, modelled as social contagion phenomena and the epidemics spreading unveiling the interplay between them, in a multiplex network but separating and constraining each of the processes to only one of the layers [9], [29], [30]. By contrast, in [16], [17], it has been explored and quantified the impact of the co-evolution of the two processes in all the layers considering the introduction of the multiplexity dimension. Multiplex networks representation consider the same set of nodes in all the layers of interactions, and constitute the most suitable network structure to understand such dynamical processes and their complex interdependence [12], [31]–[33]. The interplay between epidemic spreading and awareness dynamics allows to highlight the role of structures and human-related factors in the multiplex networks which influence the dynamical trend of both processes. Since the more the networked individuals are aware about a disease spreading, the more they may be able to adopt strategies with consciousness, permitting the rising of collective awareness. In this paper, we propose an analytical and data-driven rigorous investigation of what is the shape of response in terms of collective attention around a shocking and long-lasting event in which awareness and attention, in single and collective terms, jointly develop in an interdependent way. We investigate and measure which is the role and the weight of multiple and heterogeneous social ties, through the introduction of the weighted multiplex network, and the impact on the epidemic spreading. We consider two weighted multiplex networks, in which we evaluate both the complex dynamics of nodes belonging to a population and that one referred to keywords, mined from social networks, and linked to the topic COVID-19. As social ties between nodes may have different weights reflecting their intensity [34], we consider the two weighted multiplex networks providing new definitions of weights for both multiplexes. We include, in the weights of population-based multiplex, the homophily and the awareness difference between nodes, which in turn depends on the collective attention dynamics of the keywords-based multiplex, that considers the keywords co-adoption interactions, defined in our model. We quantify the impact of the attention and awareness on epidemic spreading, by introducing heterogeneity, both in terms of susceptibility and awareness, for optimally schedule effective crisis communications, facilitating timely crisis response planning, such as the decision of a time warning and quarantine. It is crucial to examine how the attention and the awareness arise and fade in different communities, affected in various times, impacting and influencing behaviours and decisions. Moreover, we define a data-driven approach applied in the case of COVID-19, referred to social networks and social media platforms user-generated data, as Twitter and Google Trends [1], [35]. We compare our findings with a null model, to investigate a time-dependant and complex relationship of attention and awareness centred on the ongoing epidemics of COVID-19.

A. CONTRIBUTIONS OF THIS PAPER

Once explained the motivations behind the proposed methodology, the main contributions of the paper are summarised below.

- We propose a modeling and data-driven approach to structure a systematic understanding of collective attention dynamics of social networks, during an extraordinary event for a better crisis response. We point out to unveil the dynamical interplay between collective attention and two co-evolving spreading processes of awareness and epidemics on weighted multiplex networks.
- We disclose the role of human-related factors, such as homophily, namely the tendency of interacts with similar users, network properties and heterogeneity through the introduction of two interdependent weighted multiplex networks which respectively include connections of networked individuals, based on theoretical and empirical interactions, and connections between keywords, based on the co-adoption relationships. The empirical interactions between users and the co-adoption interactions between keywords are mined from the user-generated data digitally traced from Twitter and Google Trends, during the COVID-19 epidemics.
- We define weights between users of the first weighted multiplex network by taking into account both the concept of homophily and awareness, pointing out how these weights influence the epidemics dynamics on the social multiplex network. For the second weighted multiplex network, we define the interactions weights based on the co-adoption of a pair of keywords, as the frequency of their combined use by users digitally traced through the data-driven approach.
- We explore and quantify the dynamical interplay of collective attention on COVID-19 with the co-evolving spreading processes, defining the awareness measures of the users, as a measure depending on their participation coefficient, and the entropy of the topics, traced and analyzed in a specific temporal window and mined from social and communications technologies.
- We quantify the reaction of collective attention, detecting a social network predictive marker of the awareness dynamics in social networks and its key role on a timely crisis response planning and management of the epidemics spreading of COVID-19.

II. RELATED WORKS

A. EPIDEMICS SPREADING AND AWARENESS AS SOCIAL CONTAGION IN MULTIPLEX NETWORKS

The complex systems are fully described by their connectedness and the network representation of the nodes, belonging to them, which interact to each other via multiple links. We can describe a network through a standard approach, consisting of the analysis of the aggregate graph, that includes all links between nodes but neglecting important information, resulting in a losing knowledge about the structural complexity and connectivity. In fact, the relationships and interactions

between nodes in many real-world systems can be different for relevance, context and meaning and can be characterized by a distance or weight [31], [32], [34]. To preserve the knowledge related to the different interactions in multiple layers we need to introduce the multiplex dimension of the network. It enables us to better quantify information encoded in terms of collective and social behaviours and in terms of spreading and diffusion. Multiplex networks, in which nodes can be adjacent to each other, through intra-layer edges or to its counterpart on another layer through the inter-layer ones [31], [32], represent the most suitable network structure for analyzing the emerging dynamical patterns of spreading phenomena, depending on the nature of social ties [17], [36]. The investigation of a multi-dimensional network representation through the multiplex networks enable us to fully characterize the behaviour of a complex system, unveiling interesting structural properties that helps to understand emerging phenomena such as cascading failures, super-diffusion, spreading and epidemic dynamics [17], [37]–[42]. These networks allow us to encompass multiple interactions exploring in what measure, the ties in the different layers, have an impact on the diffusion of social phenomena. The importance of this mathematical representation is such that the scientific interest around multiplex networks applications ranges between a multitude of different areas: biological, social and technological systems, social networks and relationships, epidemic and social contagions, air transportation networks and brain computing dynamics [11], [16], [17], [37]–[39], [41]. Complexity, in connections, can be explored in a deeper way in structural terms taking into considerations weighted multiplex networks [34], since in real-world context, the social ties between nodes may have different weights reflecting their intensity in the different layers. For that assumption, the links between nodes not only are distinguished by the kind of interaction linking the nodes, but also by the intensity reflecting the importance of these interactions. Since we are interested in capturing the impact that the dynamics of collective attention around a real event has on the co-evolution of epidemic spreading and awareness, and differently from the above cited scientific literature, we firstly investigate the role of weighted multiplex networks by defining specific weights of the ties. In addition, with regards to the complexity in emerging dynamics of the co-evolving and interdependent spreading processes, in the weighted multiplex networks, we include the analysis of social contagion. Social contagion is modelled as a spreading process of diseases [16], and it is based on epidemiological models [21], and it finds applications in network science, social behavioural analysis, misinformation diffusion, infectious disease and emotions spreading [11]. These phenomena spread inter-personally, and, thanks to a vast amount of literature that has investigated these complex dynamics, following classical epidemiological models [21]–[23], involving several research fields in network science [24]–[27], [43]–[46], we can state that the nature of social ties has a key role in a phenomenon diffusion on a social network [11], [14], [18], [28], [47], [48]. Several studies investigate the role of

the awareness in the disease spreading process, disclosing his effect on the networked individuals more likely to adopt strategies targeted at self-protecting [16], [17] to slow diffusion, if they are more aware about it. Many studies have explored the spreading and competition of both phenomena separating them into different layers of diffusion [9], [30], [49]–[51]. Otherwise, since the multiplex networks constitute the most fit network structure to study these processes, their interdependence and co-evolution [33], we follow the approach to do not separate each of the spreading process to only one layers, but as in [16], [17] we investigate and quantify the impact of the co-evolution of the two processes in all the layers of the weighted multiplex network, adding the collective attention influence on the awareness dynamics fitness during an extraordinary event as COVID-19 epidemics. Coherently with the real nature of multiplex networks [36], it has been taken into account heterogeneity and its impact along with awareness on the epidemic spreading [16], [52].

B. COLLECTIVE ATTENTION

A growing number of existing studies have investigated the concept of collective attention, especially when it comes from the user-generated data. Since the escalation of social and communication technologies provides everyday a chance to digitally observe the interests of networked individuals, there is a growing involvement by the researchers in leveraging these technologies with the aim at providing predictors useful in case of crisis response [35]. Social media, and major communication platforms is having a significant impact on broadening participation by social networked individual, facilitating the contribution of information, ideas, contents, opinions, that we can see as a marker of their interests [35], [53]. The heterogeneity in terms of information loaded in the communications platforms, in conjunction with the heterogeneity in terms of network structure, impact on the user attention, showing that if the collective attention is limited, this can leads to a low discriminating capacity for users in understanding the quality of information [53]. Other studies have investigated how the collective attention in the online social platforms can represent a digital fingerprint of the human behaviours. It is interesting to underline that, when a real-world event occurs, the activity, in terms of participation or content sharing, which represents the collective attention on social media, exhibits a burst-like increase and an irregular oscillation, otherwise it follows a regular circadian rhythms, as showed in [54]. Moreover, as investigated in [35], with a comparison between originally observed attention graphs and a baseline, it is notable that collective attention displays a heterogeneous structure, showing that the interests of networked individuals generally are polarized around clusters, with reference to different topics. When an emergency or shocking event occurs, there is a shifting on people's attention inducing a convergence into a more homogeneous structure. The events attract the attention of people impacting on the awareness of them, changing the social behaviour and, in some cases, it depends more from the content exposure and much less

by the event dynamics, such as in case of an epidemics [55]. Although the concept of collective attention remains not well defined and it is related to different quantities and parameters, such as burstiness or popularity in content sharing [51], [56], [57], in several works, authors decided to focus on the emergence of collective attention through the analysis of contents streams from Twitter communications, to detect the shape of the dynamics [35], [53], [54]. Other works are based on Google Trends, which enables to analyze the popularity of specific search terms [58], [59]. Recently, around the emergency linked to the COVID-19, in [1], such as in case of other health emergencies [60], the authors studies the overall search trends regarding the epidemics, which increase in the early time of the observation period, varying in each country considered. To systematically understand the digital traces of the collective attention released and the interplay with the individual awareness which in turns produces changes in behaviours, [11], [30], [45], we need to discover and quantify the dynamics of the interdependence, using both theoretical frameworks and empirical data. To this aim, we propose an analytical framework based on the heterogeneous weighted multiplex networks structure and the data-driven approach which allow us to integrate and collect data from Twitter communications and Google Trends. In this way we are able to detect a possible social network marker to forecast the trend of awareness dynamics and its impact on the epidemics, exploiting the infodemics of online social platforms and the connectedness of users.

III. MODEL

A. SCENARIO

We introduce a new framework, as showed in Fig. 1, to quantify the interplay between collective attention and awareness dynamics, modeling them in two co-evolving and interdependent weighted multiplex social networks. Centering the analysis around the COVID-19, we choose geographical countries, represented as “communities”, which have encountered the first case of COVID-19, within a fixed time window T . The T interval has been properly selected to extract the relevant spreading dynamics in different countries. Each first episode of COVID-19, officially reported in each selected geographical countries, here is stated as “event occurrence” E_{c_i} in T , with c_i , the country belonging to the set C of countries selected. Following the data-driven and sampling approaches, presented in the Sections IV-A, the users, belonging to the communities, constitute the nodes which populate the first multiplex social network M_1 , namely the population-based multiplex. Each layer of M_1 is referred to different kinds of interactions based on social networks relationships. For this reason, in M_1 , the interactions in the first layer between users, follow the theoretical scheme of a scale-free network [10], [61], with a power-law dependence of the degree distribution $P(k) \sim k^{-\gamma}$, with the exponent $\gamma = 2$ that typically satisfying values around $2 < \gamma < 3$. Instead, the second layer of M_1 follows the graph

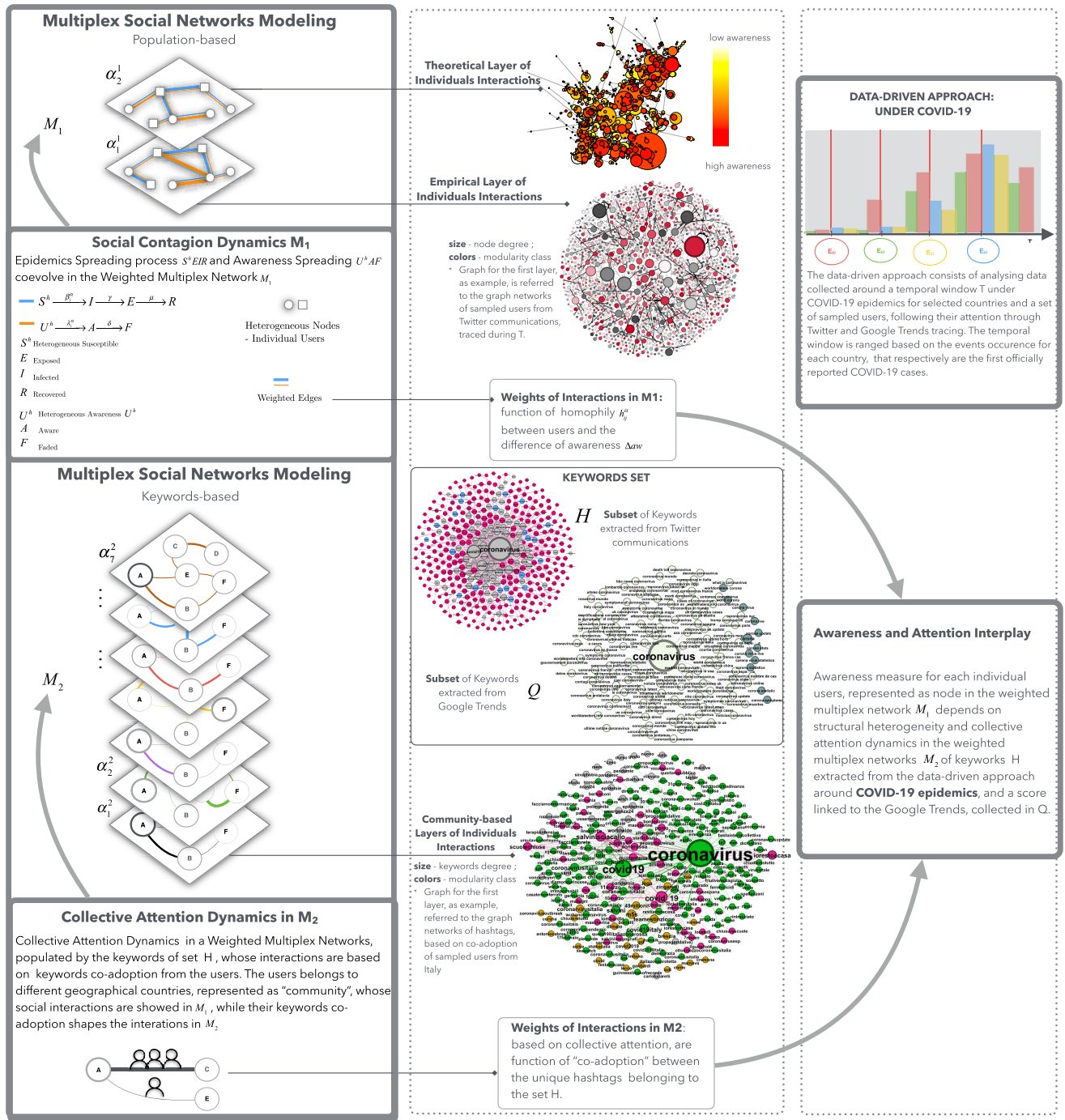


FIGURE 1. From Modeling to Data-driven approach. The figure schematically describes the various steps and key aspects starting from the modeling procedure to the data-driven approach. In the left side of the framework we highlight the modeling steps, in which we define two weighted multiplex social networks, population-based and keywords-based, with their dynamics, respectively, a social contagion dynamics of two co-evolving spreading processes, epidemics and awareness, and the collective attention. In the right panel we show the data-driven approach, specifically around the COVID-19 epidemics. The outcomes of the latter approach and analysis are showed in the central panel, in which some of the entries are empirically extracted from the Twitter communications and from Google Trends. Moreover, we underline the interplay and the interdependence between awareness and attention which link in a complex co-evolution the dynamics of the two weighted multiplex social networks, allowing us to disclose the joint impact of attention and awareness on epidemics, and helping us to forecast how the individuals react to the emergency.

network extracted from the data-driven approach, showing virtual mined relationships for the same set of users. As detailed in the Section IV-A, we integrate and analyze large social networks communication by using Twitter datasets and

Google Trends. We construct a set of keywords K , mining a set H of unique hashtags from Twitter, related to the sampled relevant users (see IV-A3), and the set Q of the most popular terms of searches from Google Trends, in the fixed temporal

window T . The elements of the subset H of K , represent the nodes of the population of the second weighted multiplex networks M_2 , namely the keywords-based multiplex network. In M_2 , each layer is a community-based level, and it is referred to a kind of interaction, that we define as “co-adoption” relationships, based on the combined use of keywords by any users of that community representing the layer, whose social interactions are explored in the multiplex network M_1 . Following, for each weighted multiplex social networks, we detail the mathematical representation, the definitions of weights, by highlighting statistical estimators and the metrics, which impact on the co-evolution of collective awareness and attention, and the impact on the epidemic spreading.

B. MULTIPLEX SOCIAL NETWORKS MODELING

1) POPULATION-BASED

With regards to the first weighted multiplex social network, we highlight the importance of including multiple relationships between users, representing people of different communities as explained in the section III-A. The first multiplex network M_1 , allowing to investigate how different network structure impact on the contagion dynamics, consists of two layers of interactions of the same set of users, a scale free-networks and a graph extracted from the data-driven approach. This makes us able to show a theoretical and an empirical interactions scheme to represent and investigate two types of virtual relationships for the same set of users. Let us consider the first weighted multiplex social network of M_1 layers $\alpha_1 = \{1, \dots, M_1\}$ and N_1 nodes $i = \{1, \dots, N_1\}$. The multiplex network is a set of M_1 weighted networks (or layers) defined as a set of graphs G , vertices V and edges E , $G_\alpha = (V, E_\alpha)$ (see Fig 1). The set of nodes N_1 is the same for each layer, whereas the set of links E_1 changes according to the layer [34]. Each network G_α is described by the adjacency matrix, denoted by A^α with elements a_{ij}^α , where $a_{ij}^\alpha = w_{ij}^\alpha > 0$, if there is a link between i and j with a weight w_{ij} , otherwise $a_{ij}^\alpha = 0$. We include a definition of weights as function of the discrepancy of users awareness about the referred topic (Δaw) of COVID-19 and homophily (h_{ij}^α), namely the tendency of interacts with similar users.

Definition 1 (Weights of Interactions in M_1): Referred to the first multiplex network, the weights denoted as w_{ij}^α are given by:

$$w_{ij}^\alpha|_{M_1} = h_{ij}^\alpha \cdot |\Delta aw| + 1 \quad (1)$$

where $h_{ij} = 1/1 + \delta_{ij}$ is the homophily between nodes i and j . Homophily is defined as the tendency to associate and interact more with similar people [36], where δ_{ij}^α is the measure of the homophily difference between nodes i and j as expressed in [36] and Δaw is equal to the absolute difference of awareness, $|aw_i - aw_j|$, between nodes i and j (see III-B3). Homophily represent the principle that similarity breeds connections, enabling the structural understanding of how social connections are forged over time, [36]. It stands for an overall value which takes into account different dimension of similarity. Following the representation of that different dimension as

different homophily values, we choose to indicate with δ_{ij} the homophily difference between a pair of nodes and with σ its standard deviation.

2) KEYWORDS-BASED

With reference to the second multiplex network, we underline the role of “keywords” as broker for users’ involved topics. We characterize the second weighted multiplex social networks M_2 , populating it with keywords as nodes, and the multiple interactions, different for each layer, between nodes as the “co-adoption” relationships. We consider each layer of this multiplex network referred to each geographical country and the coexistence of several types of interactions among keywords based on the collective attention of individuals mined from social networks. Let us consider the second multiplex network of M_2 layers $\alpha_2 = \{1, \dots, M_2\}$ and N_2 nodes $i = \{1, \dots, N_2\}$. The set of nodes is the subset H of h hashtags, linked to the main topic under investigation. As explained below, in section IV-A, we construct that set, mining a set of unique hashtags from Twitter, used by any users in the fixed temporal window T . In T , based on a data-driven approach, we analyze a large corpus of datasets and we digitally investigate and measure the collective attention. As underlined for M_1 , we can also describe M_2 by the adjacency matrix, denoted by A^α with elements a_{ij}^α , where $a_{ij}^\alpha = w_{ij}^\alpha > 0$, if there is a link between i and j with a weight w_{ij} , otherwise $a_{ij}^\alpha = 0$.

Definition 2 (Weights of Interactions in M_2): Regarding to the second multiplex network, the weights are denoted as $w_{ij}^\alpha|_{M_2}$. To mathematically define the interaction weights, since they are based on the “co-adoption” relationships, as explained before, we introduce the hashtags “co-adoption” probabilities as relative frequencies, that are: $p_{h_i h_j}^\alpha$ which denotes the relative frequency of using both hashtags by users belonging to the community of the layer α of M_2 , and $p_{h_i}^\alpha$ as the relative frequency of using the hashtags h_i^α , by users belonging to the same community. In accordance to the latter, the weights of interactions in M_2 are given by:

$$w_{h_i h_j}^\alpha|_{M_2} = p_{h_i h_j} \cdot (p_{h_i} - p_{h_j}) \quad (2)$$

where, h_s with s in $\{1, 2, \dots, H\}$ are hashtags elements belonging to a subset of hashtags H , linked to COVID-19, the main topic under investigation.

3) AWARENESS AND ATTENTION INTERPLAY

The awareness represents the acquired knowledge as a result of reasoning and understanding and, eventually, experiencing the epidemics. To capture the linkage between awareness and attention dynamics, which in turns impacts on the epidemics patterns, in this paper, we assume that the “awareness” measure, aw_i , for each node i belonging to users population of weighted multiplex network M_1 , depends on structural heterogeneity, collective attention dynamics on M_2 , and node properties [31], [34], [62]. Moreover, we shed light on that awareness measure has not change according to the layers, since awareness once acquired is the same in the different

layers in which a user is involved. It can be influenced by the fading of interests on acquiring additional and correlated awareness, or verifying if it is fact-checked or misinformation based [17], [63]. The awareness gap between two interacting nodes in M_1 , impacts on the weights of links as detailed in section III-B1, in conjunction with the homophily value h_{ij} . In fact it is crucial to include the tendency to interact with similar people, through the introduction of homophily measure since strongly homogeneous groups tend to prefer contents that confirm their shared beliefs, polarizing rumours or misinformation. This phenomenon is defined as “echo chambers” effect which have a strict interplay with the spread of misinformation [63]. In presence of misinformation and highly homophilic clusters of users in social networks, it is very likely having fake news masqueraded as fact-checked contents [64].

Definition 3 (Awareness Measure in M_1): Referred to the first multiplex network M_1 , the awareness measure for each node i , is denoted as aw_i , and it is given by:

$$aw_{i=T} = (P_i|_{M_1} \cdot \sum_{h \in H} H_h \cdot \sum_{q \in Q} \eta_q) + aw_{i=t_0} \quad (3)$$

where P_i is the multiplex participation coefficient which enable us to include the heterogeneity of the number of neighbours of node i across the layers in M_1 [31]. The H_h represents the Shannon entropy of the hashtags h , mined from Twitter communications, belonging to the set H , which represents the population of nodes in M_2 [31], [34], and η_q is a score, represented by the relative search volume (RSV), associated with the Google search popularity of terms q , elements of the subset Q extracted through Google Trends (see section IV-A).

C. SOCIAL CONTAGION DYNAMICS

1) DISCUSSING THE APPLICABILITY OF SEIR MODEL IN CO-EVOLUTION WITH THE UAF MODEL FOR COVID-19

We study the spreading of two processes in M_1 in co-evolution in the same multiplex network, referred to the epidemic and the awareness spreading [16]. Deciding to not disjoint the two processes we are able to explore the emerging complex dynamics in order to highlight the effect of the awareness dynamics, which is marked by cognitive limitations, on epidemic spreading, underlining also its impact on how we are likely to accept the unprecedented restrictive measures due to the COVID-19. In fact, in response to the epidemic outbreak [2], [65]–[67] both local and national governments have taken non-pharmaceutical physical distancing interventions, strategies to reduce its impact, extended school closures and workplace distancing along with emergency spreading in time [68]. This complex dynamics investigation can shed light on the role of awareness in this emergency event, unveiling its interdependence with collective attention, keeping a watchful eyes on the fact-checked news, that can distance the natural memory leaks or fading of interest by human brain on epidemics. It is nevertheless worth mentioning that COVID-19, as it is clearly visible observing the timeline of pointed key occurred events [5], differently

from shocking peaked events, like they could be earthquakes or mass shooting [35], it still observed it is an unfolded emergency in time. This raises the challenge of quantifying the impact of awareness, influenced by the dynamics of people interests, in terms of searching for information and content adoption, that is revealed by the analysis of the growing number of users-generated data [35], [54], [56], [69]. To this aim, we explore how dynamics could change according to network structure heterogeneity and awareness, which in turns depends on the attention dynamics explored in M_2 . To detect a marker of the complex social dynamics, we adopt a data-driven approach which exploits large social media platform communications and searches as Twitter [35], [56] and Google Trends [59]. A rising number of studies explored the model to fit mathematically the transmission of COVID-19, taking into consideration measures and strategies which reduces social mixing, modifying both the pattern within the population and the trajectory of the epidemics. Since the transmission is mostly driven by who interacts with whom, which can vary by age, location, contact, the protective measures of distancing have been designed to shift the social mixing patterns [2]. Given that, most of these limitations transform population mixing as well as human habits and the real social connectedness, consequently it becomes crucial the understanding to what extent the attention and the awareness in social networks can lead to behaviours in accordance with the provided reduction strategies. For that reason, currently, the increasing virtual connectivity [1], [64], [70], expressed in multiple layers, can marker the realization of a collective consciousness and impacting on the COVID-19 epidemic progression. In accordance with the recent epidemiological studies on parameters of the outbreak and propagation [71], [72], [72]–[75], we consider the susceptible-exposed-infected-recovered (SEIR) model [2], [21], [72], as the epidemiological model for COVID-19, analyzing it in conjunction and co-evolution with the awareness social contagion process *UAF* [16], in M_1 . The SEIR model is generally used to model influence-like-illness, and it has been used to numerically analyze the evolution of the Severe Acute Respiratory Syndrome (SARS) in different social settings [21]. As observed in [71], the SEIR model fits mathematically with the nature of COVID-19 transmission. This model is a variation of the SIR model, with the inclusion of exposed state (E) [21]. Moreover, in line with [2], [21], [73], we assume that, Exposed state (E) is a state in which individuals have been infected by the disease but cannot yet transmit it. In [72], the authors consider a SEIR model with different assumptions, since they indicate as E the latent state in which the population is asymptomatic but infectious, and I referred to the symptomatic and infectious population. The incubation rate σ is described as the rate by which the E individual develops symptoms. In [74], instead, the authors named the proposed model as SIDARTHE discriminating between detected and undetected cases of infection and between different severity of illness (SOI), non-life-threatening cases (asymptomatic and pauci-symptomatic; minor and moderate infection) and

potentially life-threatening cases (major and extreme). In our work, we have assumed that asymptomatic cases represent a minor proportion of infectiousness compared with symptomatic cases, following the hypothesis in accordance with [2], [73]. In fact, although the contributions of asymptomatic or sub-clinical cases represent a pivotal point, and the question of whether such individuals are able to transmit infection remains unresolved at the time of writing, we do not consider that cases in this work, in accordance with the choice of temporal window that was aim at tracing the initial phase of rising the interest of collective attention around COVID-19. To analyze the co-evolving dynamics of both epidemics and awareness spreading on M_1 multiplex network, we exploit the Dynamic Microscopic Markov Chain Approach (MMCA) (see section III-C3). Following our assumptions, we do not consider all demographic changes, age, structure features and other relevant compartments in population [21].

2) CO-EVOLVING SPREADING PROCESSES

In accordance with what we have defined in the previous section, here we start from the key assumption that each node in M_1 has a different awareness on COVID-19 as underlined in section III-B3. As a consequence, each node, will be heterogeneously susceptible to the epidemics spreading, since the awareness acquired has an influence on the behaviours of people, improving their knowledge about epidemics. Heterogeneity and awareness are introduced in order to describe the spreading scenario and extricate the complex co-evolution of the interdependence, from one hand, between epidemics and awareness and, from the other hand, between awareness and attention (see section III-B3). We analyze two co-evolving spreading processes in the first weighted multiplex network M_1 (see section III-B1). The first process of epidemic spreading is indicated by S^hEIR , where S^h indicates the heterogeneous susceptible state, E the exposed state and I and R , respectively the infection and the recovered ones [16], [21]. The second is the awareness spreading process, coexisting and co-evolving with the first one, indicated as U^hAF , that stands for “Unaware-Aware-Faded” model. The awareness spreading process is thought as a SIR-like, where U indicates the condition of heterogeneous unawareness state, A is the aware state, while the F state, is the faded state in which the nodes, tend to decrease the interest to the referred topic, over time up to the point that it completely vanishes. The Awareness state is referred to the case of those who is aware on topic, which means that a user acquires knowledge about epidemics, as a result of a collective attention dynamics in social networks. This can leads to a discriminating capacity in understanding quality of good information on epidemics. In our model, the awareness state is not directly correlated to the ensuring that users follow quarantine or other physical strategies. Although an increasing of awareness can infers a good behaviour in line with that strategical measures, influencing people to follow preventive measures as distancing interventions [16], [17], [71], in our work it has been intended as a knowledge and a consciousness that

can contrast for example the misinformation diffusion about COVID-19. The node, in the F state, maintain the same awareness measure, but it has no interest in increasing its acquired knowledge on the phenomenon linked to the referred topic. Starting from the model presented in [16], we extend it introducing, other than the heterogeneity in terms of nodes’ susceptibility and awareness, in the layers of M_1 multiplex network, also the linkage to the expressed collective attention which feeds and drives the shape of awareness measure. The meaning of the various states of the two spreading processes are explained in Table 1, and a complete list of symbols with its fair meaning is summarized in Table 2, as follows:

TABLE 1. MMCA - States. We include the definitions of the different states in the MMCA, adding more details about the interdependence between epidemics, awareness and attention (see section III-C).

Status	Description
S^h	Heterogeneous Susceptible - Those who can contract the infection with a different infection rate. The nodes’ heterogeneity, influenced by the awareness rate, impacts on the probability to have a transition in exposed status.
E	Exposed - Those who have been infected but cannot yet transmit it.
I	Infected - Those who contracted the infection and are contagious.
R	Recovered - Those who recovered from the disease.
U^h	Heterogeneous Unaware - Those who is unaware of the epidemic, or with minimum or wrong awareness. The probability to have a transition in Aware state is different for each node, depending on the impact of the collective attention, whom dynamics is analyzed in M_2 , on the weights of interactions of the multiplex network M_1 .
A	Aware - Those who is aware of epidemic, or who develop an awareness measure based on fact-checked information. The awareness measure depends on the collective attention on epidemics (see section III-B3)
F	Faded - Those who tend to decrease the attention on epidemics, or the interest to improve knowledge. The more susceptible are nodes that reach the faded state, the more vulnerable they become.

α : HETEROGENEOUS SEIR

The S^hEIR spreading process is diagrammatically expressed, in terms of reaction-diffusion equations, as follows:

$$S^hEIR \Rightarrow S^h \xrightarrow{\beta_i^\alpha} E \xrightarrow{\gamma} I \xrightarrow{\mu} R \quad (4)$$

with β_i^α , the infection rate for each node i at each layer α in the multiplex network M_1 . The probability of an exposed individual becoming infectious is γ , while the probability of becoming recovered after infection is equals to μ . The heterogeneous infection rate represents the probability of a susceptible-infected contact results in a new exposure, as following expressed:

$$\beta_i^\alpha = \frac{1}{1 + \lambda_i^\alpha} \cdot \frac{1}{Y_i^\alpha} \quad (5)$$

It depends on the network structure heterogeneity, through the measures of inverse participation ratio Y_i^α [34] (see Supplementary Information) and the rate of awareness λ_i^α as expressed in Eq. 7.

TABLE 2. List of symbols.

Symbol	Description
T	Observed Time Interval.
t_i	Time sub-intervals considered.
DB_t	Datasets collected.
C	Set of locations c_i considered
E_{c_i}	Event occurrence with c_i the country belonging to C .
H	Set of unique hashtags h from Twitter.
Q	Set of the most popular terms of searches q from Google Trends.
K	Set of keywords $K = H \cup Q$.
U	Set of all users u .
N	Set of Sampled users.
A^α	Adjacency Matrix with elements a_{ij}^α .
M_1	Population-based Multiplex Network.
M_2	Keywords-based Multiplex Network.
aw_i	Awareness of a node i in the Multiplex Network M_1
$w_{i,j}^\alpha _{M_1}$	Weights of interactions between two nodes i and j in a layer α of the Multiplex Network M_1
$w_{h_i,h_j}^\alpha _{M_2}$	Weights of interactions between two hashtags h_i and h_j in a layer α of the Multiplex Network M_2
$h_{i,j}$	Homophily between node i and node j .
ph_i	Relative frequency to use the hashtag h_i .
ph_{i,h_j}	Relative frequency to use both hashtags h_i, h_j .
P_i	Multiplex Participation Coefficient of node i in M_1 .
H_h	Shannon entropy of hashtag h in M_2 .
η_q	Score associated with the RSV from GTrends.
Y_i^α	Inverse Participation Ratio in the layer α .
s_i^α	Strength of node i in the layer α .
o_i	Overlapping degree of a node i .
Z_{o_i}	Z-score of a node i .
β_i^α	Heterogeneous Infection rate.
λ_i^α	Heterogeneous Awareness rate.
δ	Fading rate.
γ	Probability to transit from state E to state I
μ	Probability to transit from state I to state R
$q_i(t)$	Probability not being infected at step t .
$r_i(t)$	Probability not being aware at step t
β_c	Contagion Threshold.
ρ	Density of Infected nodes.

b: HETEROGENEOUS UAF

The second spreading process is diagrammatically expressed as follows:

$$U^h A F \Rightarrow U \xrightarrow{\lambda_i} A \xrightarrow{\delta} F \quad (6)$$

with λ_i^α and δ , defined as transition rates, respectively represent the rate of awareness for each node i at each layer α of the multiplex M_1 and the fading rate. The rate of awareness is expressed as follows:

$$\lambda_i^\alpha = \frac{s_i^\alpha}{1 + s_i^\alpha} \cdot \lambda \quad (7)$$

with s_i^α the strength of node i in the layer α of the multiplex networks M_1 [34] (see Supplementary Information). The Dynamic Microscopic Markov Chain Approach (MMCA), enables us to explore the spreading dynamics of co-evolution on the weighted multiplex network M_1 , depending also from the collective attention dynamics in M_2 . In the next subsection we give a more detailed explanation of the MMCA used in our model.

3) DYNAMIC MICROSCOPIC MARKOV CHAIN

We start to assign to each node a state probability to be in one of the initial state, since at the beginning it can occupy only one of the following: susceptible and unaware (SU), susceptible and aware (SA), infected and aware (IA), exposed and unaware (EU) (see Fig. 2). Some states are not reachable or do not exist, such as infected and unaware (IU) and recovered and unaware (RU), due to the assumption that if a node has been infected or recovered has experienced the disease, developing a measure of awareness (see Fig. 2). At time step t each node i can occupy one of the initial states, with probabilities $p_i^{SU}(t)$, $p_i^{SA}(t)$, $p_i^{IA}(t)$, and $p_i^{EU}(t)$ respectively. We consider, as in [17], $q_i(t)$ as the probability of node i not being infected at time step t and $r_i(t)$ the probability of an unaware node i staying unaware at time step t , as follows:

$$q_i(t) = (1 - \bar{\beta}_i) \prod_j [1 - a_{ji} p_j^I(t) \bar{\beta}_i] \quad (8)$$

$$r_i(t) = (1 - \bar{\lambda}_i) \prod_j [1 - a_{ji} p_j^A(t) \bar{\lambda}_i] \quad (9)$$

where a_{ij} are the elements of the adjacency matrix of each layer of the weighted multiplex network M_1 . $\bar{\beta}_i$ and $\bar{\lambda}_i$ respectively are the mean values of the heterogeneous infection and awareness rate. The following MMCA equations represent the probability of each node of being in one of the states at time step $t + 1$, as showed in Fig. 2:

$$\begin{aligned} p_i^{SA}(t+1) &= q_i(t) p_i^{SA}(t) + (1 - r_i(t))(1 - \delta) p_i^{SU}(t); \\ p_i^{SU}(t+1) &= r_i(t) p_i^{SU}(t); \\ p_i^{SF}(t+1) &= \delta(1 - r_i(t)) p_i^{SU}(t); \\ p_i^{IA}(t+1) &= \gamma(1 - q_i(t))(1 - \mu) p_i^{SA}(t) \\ &\quad + (1 - \mu) p_i^{IA}(t); \\ p_i^{EU}(t+1) &= r_i(t) p_i^{EU}(t); \\ p_i^{EA}(t+1) &= (1 - \delta)(1 - r_i(t)) p_i^{EU}(t) \\ &\quad + (1 - \gamma)(1 - q_i(t)) p_i^{SA}(t); \\ p_i^{EF}(t+1) &= \delta(1 - r_i(t)) p_i^{EU}(t); \\ p_i^{RA}(t+1) &= \mu(1 - \delta) p_i^{IA}(t) \\ &\quad + \gamma \mu(1 - q_i(t))(1 - \delta) p_i^{SA}(t); \\ p_i^{RF}(t+1) &= \mu \delta p_i^{IA}(t) + \delta \gamma \mu(1 - q_i(t)) p_i^{SA}(t); \end{aligned} \quad (10)$$

To obtain the contagion threshold, we investigate the steady state solution of the system constituted by the previous equations. When time $t \rightarrow +\infty$, there exists a contagion threshold β_c for the two co-evolving processes, so that the contagion can outbreak only if $\beta \geq \beta_c$. Following the mathematical approach presented in [16], the contagion threshold is given by the order parameter ρ_i and it is defined as follows:

$$\rho^I = \frac{1}{N} \sum_{i=1}^N p_i^I = \frac{1}{N} \sum_{i=1}^N p_i^{IA} \quad (11)$$

Thus, starting from equation $p_i^{IA}(t+1)$ (see Eq. (10)), at steady state we have:

$$p_i^{IA} = \gamma(1 - \mu)(1 - q_i) p_i^{SA} \quad (12)$$

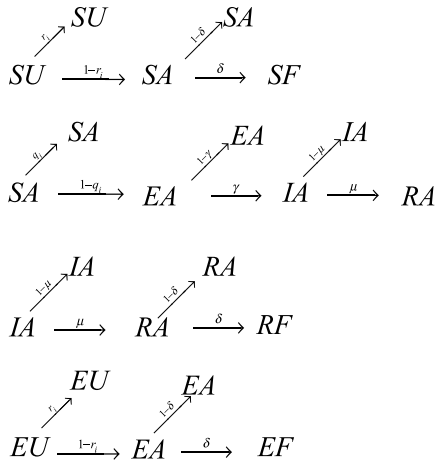


FIGURE 2. Probability tree. We show the probability tree, linked to the MMCA method, representing the states and the transitions, at each time step. Roots represent the initial states, SA, SU, IA and EA at time step t , while leaves are all the possible states at the subsequent time step. Arrows are labelled with the corresponding transition probabilities.

Since around the contagion threshold β_c , the informed probability is close to zero ($p_i^A = \eta_i \ll 1$), the probabilities of being informed can be approximated as follows:

$$q_i = (1 - \bar{\beta}_i)[1 - \bar{\beta}_j \sum_j a_{ij}\eta_j] = (1 - \bar{\beta}_i)(1 - \omega_i) \quad (13)$$

where:

$$\omega_i = \bar{\beta}_j \sum_j a_{ij}\eta_j \quad (14)$$

Furthermore, close to the contagion onset we have that the fading rate is approximately close to zero ($\delta \simeq 0$). Considering this approximation into Eq. 12 and omitting higher order items, Eq. 12 is reduced to the following form:

$$\mu\eta_i \simeq \gamma(1 - \mu)p_i^{SA} \bar{\beta}_i \bar{\beta}_j \sum_j a_{ij}\eta_j \quad (15)$$

The contagion threshold is obtained starting from the following condition:

$$\sum_j \left| \gamma(1 - \mu)\bar{\beta}_i p_i^{SA} a_{ij} - \frac{\mu}{\bar{\beta}_j} t_{ji} \right| \eta_j = 0 \quad (16)$$

where t_{ji} are the elements of the Identity matrix. By defining the matrix H whose elements are given by: $h_{ij} = [\gamma(1 - \mu)\bar{\beta}_i p_i^{SA}] a_{ij}$, the contagion threshold β_c is the value corresponding to the largest eigenvalue of the matrix H , which is given by $\Lambda_{\max}(H) = \gamma\mu/\bar{\beta}_j$, so finally we get: $\beta_c = \gamma\mu/\Lambda_{\max}(H)$. The onset of the epidemic is the minimum value satisfying the Eq. 16. By denoting with $\Lambda_{\max}(H)$ the largest eigenvalue of H , we obtain the critical point of β_c which depends explicitly on the co-evolving dynamics. Even if we consider the critical value referred to the onset of the awareness spreading dynamics λ_c as a simple spreading process, when decoupled from the epidemic, the (β_c, λ_c) defines

a sort of meta-critical point for the spreading dynamics. In our work, the threshold model depends on the complex dynamical interplay, since the values of β change also accordingly with the awareness rate, the network structure of the weighted multiplex network, the collective attention and the double heterogeneity in terms of susceptibility and awareness.

IV. EXPERIMENT

A. DATA-DRIVEN APPROACH

1) CHARACTERIZING COLLECTIVE ATTENTION AND AWARENESS INTERPLAY, UNDER COVID-19

In our model, we propose a data-driven approach for evaluating the complex dynamics of co-evolving epidemics and awareness spreading, in function of the collective attention. To this aim, in the weighted multiplex network M_2 , as showed in section III-B2, we analyze user-generated data and searches, respectively, by using a large corpus of Twitter communications datasets as listed in Table 3, and the most popular Google search terms, under COVID-19. The vast communications streams and searches, which still going on, enables us to monitor collective attention and, through the proposed framework in this paper, understand how it manifests itself under a real-world emergency event. In the whole temporal window considered T , we include small ranges Δt_{E_c} created as a consequence of relevant events happened, referred to the time in which the event has been officially reported in the c_i geographical countries considered, as detailed in section III-A. In particular, we take into account seven short sub-intervals, as listed in Table 4, in order to monitor attention patterns around the first officially reported cases of COVID-19 in each geographical country considered. Due to the observation constraints of Twitter API, in order to evaluate a large corpus of data in a long time window we exploited data collected in [76], [77], whose statistics are shown in Table 3. These datasets were created searching for users who have applied hashtags related to COVID-19 such as: #coronavirus, #coronavirusoutbreak, #coronavirusPandemic, #covid19, #covid_19 [76], [77]. Each tweet in the datasets includes textual content, the author id and nickname, the creation time, if it was in reply to another tweet, whether it is a retweet and additional metadata. To identify all topic, we extracted from tweets' text all hashtags adopted and we considered both original tweets and retweets. In order to have a representation of the social network where the hashtags' diffusion takes place and to estimate the activity of users around COVID-19 emergency, we traced back users, through the Twitter REST API, and we collected additional information such as locations and number of Followers (see Supplementary Information for further details). As showed in section III-A, we take into consideration two interdependent weighted multiplex networks to disclose the interplay between epidemics and awareness spreading, based on collective attention dynamics. We mined the "Retweet-Mention-Reply" graph of a group of relevant users, that through the sampling approach, becomes the sampled set N of unique

TABLE 3. Datasets used in this study.

Dataset	Period	Users	Tweets Volume	Hashtag Volume
[76]	01-Dec-2019 - 28-Feb-2020	41,914	60,160	10,986
[77]	01-Mar-2020 - 11-Mar-2020	22,1928	526,791	135,664
[77]	12-Mar-2020	223,187	497,299	110,999
[77]	13-Mar-2020	569,099	994,611	238,367
[77]	14-Mar-2020	146,960	465,506	145,612
[77]	15-Mar-2020	245,218	464,396	142,924
[77]	16-Mar-2020	251,134	612,481	190,967
[77]	17-Mar-2020	356,262	834,944	249,083
[77]	18-Mar-2020	220,564	626,206	209,202
[77]	19-Mar-2020	284,981	855,845	267,057
[77]	20-Mar-2020	403,761	758,621	253,082
[77]	21-Mar-2020	393,600	881,955	256,816
[77]	22-Mar-2020	284,692	700,932	216,159
[77]	23-Mar-2020	255,576	658,225	229,724
[77]	24-Mar-2020	297,203	620,860	226,176
[77]	25-Mar-2020	278,293	684,556	245,815
[77]	26-Mar-2020	195,235	881,278	209,368
[77]	27-Mar-2020	302,559	664,120	239,108
[77]	28-Mar-2020	271,370	582,055	211,808
[77]	29-Mar-2020	227,820	564,141	204,568
[77]	30-Mar-2020	285,320	586,262	224,554

TABLE 4. Time windows.

t1	12-Dec-2019 - 31-Dec-2019
t2	01-Jan-2020 - 15-Jan-2020
t3	16-Jan-2020 - 31-Jan-2020
t4	01-Feb-2020 - 15-Feb-2020
t5	16-Feb-2020 - 29-Feb-2020
t6	01-Mar-2020 - 15-Mar-2020
t7	16-Mar-2020 - 30-Mar-2020

users, population of the weighted multiplex network M_1 . We also construct the subset of the most relevant hashtags H which represent the population of the weighted multiplex network M_2 , restricting our data to the unique hashtags used by users in N in all states under observation (see Supplementary Information). For these two weighted multiplex networks, we analyze the collective dynamics and interactions of the population of sampled users N for M_1 , and hashtags belonging to the subset H for M_2 . In order to examine the attention dynamics, we evaluate also the set Q mined from Google Trends considering the top 25 search keywords, having the relative search volume (RSV) score greater than 0, and in particular, among these, the related queries of searches about “COVID-19” in each geographical country for each time interval t_i in T . The pseudo code of the sampling and modeling approach and the social network marker detection is shown in Algorithm 1.

2) SAMPLING APPROACH

To construct the networks structure of the M_1 weighted multiplex network (see section III-B1), we consider a population of users, interacting through a scale-free network in the

Algorithm 1 The Interplay of Collective Dynamics in Multiplex Social Networks

Input: Event in T (as COVID-19); Datasets collected (DB_t); $\delta_{ij}, \mu, \delta, \gamma$.

Output: $M_1, M_2, aw_i, \beta_c, \rho_I$ in T .

1: Phase 1: Sampling Approach.

Sampled Set of Users $N \leftarrow \emptyset$

Sampled Set of Hashtag $H \leftarrow \emptyset$

Sampled Set of Queries $Q \leftarrow \emptyset$

2: Set $t_i \in T$ filtering DB_t by Tweets’ creation time.

3: $\forall t_i \in T, \forall u \in U$ collect data about users through Twitter API.

4: $\forall t_i \in T, \forall c$ collect queries through GTrends API.

5: $\forall t_i \in T$ and $\forall u \in U$ calculate “activeness” and “connectedness” of users, considering the number of tweets n_{T_u} of a user u and a threshold T_T , the number of followers n_{F_u} of a user u and a threshold T_F .

if $\forall t_i \in T, n_{T_u} > T_T$ **AND** $n_{F_u} > T_F$ **then** $N = N \cup u$.

6: $\forall t_i \in T$ and $\forall u \in N$ retrieve Hashtags from Tweets in DB_t .

7: Filter N by location, c , and consider H_c set of retrieved Hashtags.

8: Create $H = H_1 \cap H_2 \dots \cap H_c$, and $Q = Q_1 \cap Q_2 \dots \cap Q_c$ from the most frequent queries $\forall t_i \in T$.

9: Mining of the “Retweet-Mention-Reply” graph of the users in N and the c “Hashtags co-adoption” graphs of the hashtags in H .

10: Calculate the relative frequency of the hashtags in H , ph .

11: Phase 2: Multiplexity - Social Contagion.

12: Set N as N_1 population of M_1 .

13: Set the “Retweet-Mention-Reply” graph of the users in N as $G_{\alpha=1}$ and a Scale-free networks as $G_{\alpha=2}$, with G_{α} graphs of the M_1 with $\alpha = 1, \dots, M_1$ layers.

14: Set H as N_2 population of M_2 multiplex network and the “Hashtags co-adoption” graphs of the users in N as $G_{\alpha=c}$, with G_{α} graphs of the M_2 with $\alpha = 1, \dots, M_2$ layers.

15: $\forall i \in N_1$ in M_1 calculate k_i, o_i, P_i, Z_{o_i} .

16: $\forall h \in H$, in M_2 , calculate calculate k_h, o_h, H_h and $\forall q \in Q$ a score η for each c location.

17: $\forall i \in N_1$, in M_1 , calculate $aw_i, w_{ij}, s_i^\alpha, Y_i^\alpha$

18: Calculate $w_{h,hj}$ in M_2 .

19: $\forall i \in N_1$, in M_1 , assign to i one of the initial states SU - SA - IA.

20: at time step t , calculate $\lambda_i^\alpha, \beta_i^\alpha, q_i(t), r_i(t)$.

21: MMCA method.

22: Calculate β_C, ρ_I in T .

23: Phase 3: Social Network Marker.

24: detect the emergence of the first event case E_c^I , and $\forall c$, detect E_c, R_c^I, SM_c^I , and calculate the delays $DR_c^I E_c^I, D_{cm_c^I} E_c^I$ and $DR_c^I E_c^I$.

25: $\forall c$, calculate aw_r , and, its growth rate $\pm aw_r(\%)$ as the social network marker impact.

first layer and a sampled weighted graph network in the second layer, mined from a selection of a set of interacting nodes on Twitter and based on the datasets analyzed [76], [77]. Although a straightforward way for an understanding of attention dynamics is to gather data from as many users as possible, we assume that it is fundamental to apply a sampling approach, to identify the most influential users and their relationships, preserving the topological properties of the original structure. This set of sampled users, are that more likely can trigger awareness diffusion giving an impact on epidemic spreading. An ideal sample set should consider users with the following features [35]:

- *Activeness*: We examine in the sampled set, users who tend to tweet with hashtags linked to COVID-19, at a relatively high frequency during the time of interest T . In this way we include who have the tendency to maintain a high level of interest on topic through time, presuming that they are more likely to be active in the spreading dynamics. A user is included in the sampled set N if has posted a number of tweets greater than a threshold T_T in each time window t_i .
- *Connectedness*: We examine and include in sampled set, users who tend to be actively connected with other users showing their capability to cover a broader set of users, namely those who actively join into the common interests of many other users. We define the threshold T_F of users' number of followers and we add to the set N of sampled users who has a "followers_count" value greater than T_F .

Evaluating both the activeness and the connectedness we select the most popular users that are at the same time, the most active within the time period considered, avoiding missing content or activity gaps over time.

3) COMPARISON WITH A NULL MODEL

An interesting pivotal point is bringing together the nature of the extraordinary event, such as the one represented by COVID-19, with the usual emerging dynamical patterns from social networks. Without any doubt, this epidemic is not a common health emergency. In terms of collective attention, the interest is spread over time and continues to involve the population as a result of strategical measures and an uninterrupted updates on the data and the evolution of spreading itself [68], [70]. To evaluate if the emerging dynamical patterns is caused by an interplay between polarized attention and awareness on epidemics, or by a change on individual interests dynamics, rippling with a similar shape in other cases, we compare the observed dynamics with that one based on a null model. The null model is created for a fixed set of users, comparable to those considered in our model, in a period of time, prior to COVID-19, and comparable with our time interval T , during which we analyze the same set of data required to match the dynamical patterns. In this way, we can observe how and in what measure the observed dynamics differs from a baseline.

B. SIMULATION RESULTS AND DISCUSSION

Simulations have been carried out considering two weighted multiplex networks, respectively M_1 and M_2 as explained in Sections III-B1 and III-B2. Firstly, with regards to the M_1 weighted multiplex social network, we consider the key role of multiple relationships between users, which represent people of different communities, referred to the seven states considered in the data-driven approach. In M_1 , we take into consideration the population of $N = 1461$ users, representing the users of the sample extracted as showed in Section IV-A. In the first layer the interactions are based on the graph network extracted from the data-driven approach, showing virtual relationships for the sampled set of users, while in the second one the interactions follow the theoretical scheme of a scale-free network [10], [61].

From Fig. 3, we can observe how the structural heterogeneity of interactions and the heterogeneous distribution of the awareness characterize our model, depending on the parameters of the multiplex network, considering the features of each node regarding the two co-evolving spreading processes around epidemics and awareness dynamics. Each node belongs to the sampled users set extracted from the data-driven approach, and it owns a distinct awareness measure, such a consciousness that allows to react differently to the epidemics, producing a different susceptibility. Following the assumption of our proposed model and the data mined, in panel (a) of Fig. 3 we highlight the heterogeneity distribution in network structures for the case of awareness measures based on multiplex parameters and user-generated data mined from Twitter and Google Trends (see details in both Sections III-B3 and IV-A), in comparison with the awareness measures based on baseline statistics, as showed in panel (b). It is notable how the multiplex network exhibits a heterogeneous distribution of awareness measures in panel (b), as a result of a global attention distributed in various clusters of different topics. Instead, consequently to the occurrence of an emergency event, as COVID-19, it acts as a shaking force that polarizes the collective attention shaping the awareness distribution in multiplex network M_1 into a more homogenized structure since the nature of the event rules the choice on what to pay attention [35]. This effect produces a decreasing in randomness and an increasing of a collective process. In accordance with the assumption of our model, the awareness measures distribution acts on the interactions weights in conjunction with the homophily measure, impacting on the heterogeneity in terms of node susceptibility against the epidemics, and in terms of infection and awareness rate, shaping also the structural heterogeneity of the degree distribution in the M_1 weighted multiplex network.

In each plot of Fig. 4, the curves correspond to the different values of the distribution of P_i , the participation coefficient of node i in the weighted multiplex network M_1 in the range $[0, 1]$, that gives information about the distribution of the edges across the layers for a node i . Firstly, we can observe that there is a quite broad distribution suggesting the presence of various levels of node's participation to each of the two

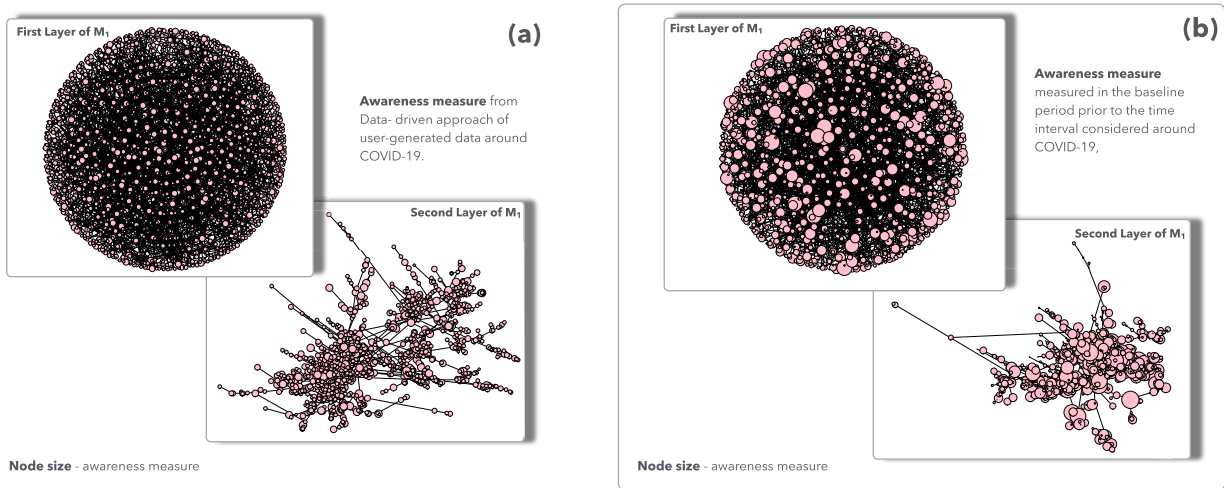


FIGURE 3. Awareness distribution on Multiplex Network. We show the heterogeneous distribution of the awareness on the two layers in the weighted multiplex network M_1 . In (a) and (b) we illustrate the two layers of the weighted multiplex network M_1 (as explained in section III-B1). In the top-left and bottom-left panels, in (a), we respectively show the heterogeneous distribution of the awareness measures, in the first and in the second layer, extracted from the data-driven approach of user-generated data around COVID-19. In the top-right and bottom-right panels, in (b), we respectively shed light on the heterogeneous distribution of the awareness measures, in the first and in the second layer, in the baseline case. See section IV-A for details about baseline and data-driven comparison method.

layers in M_1 . We consider this variation in function of the Z-score of the overlapping degree of the node, representing its overall importance in terms of number of edges, with the aim at introducing a classification of nodes in terms of its properties into the multiplex network, highlighting also the awareness distribution, showed as size of nodes. Moreover, the panels (a)-(c)-(e) are referred to the data-driven awareness measures, the panels (b)-(d)-(f) to the baseline statistics, and in both cases from the top to bottom panels we decrease the homophily measures among nodes [31], [78]. We identify in each plot different categories of nodes, hubs, for which $o_i \geq 2$, regular nodes, for which $o_i < 2$. Representing each node as a point in the plane $P_i - Z_{o_i}$, by considering the multiplex participation coefficient and the variation in function of Z_{o_i} , we pick out the six classes of nodes as highlight in Fig.4, varying homophily measures, showing the awareness distribution. While we observe the same trend in both cases, it is observable the greatest randomness effect, in terms of awareness distribution, in plots referred to the baseline statistics, (b)-(d)-(f). In the case of (a)-(c)-(e) plots, the more participatory the nodes are, the more they acquire a greatest value of awareness, especially in the case of high homophily among nodes (as showed in plot (a)), disclosing this trend especially for regular mixed and multiplex aware nodes. Decreasing homophily values, from (c)-(d) to (e)-(f), respectively referred to awareness data-driven and that one based on baseline statistics, we highlight, with an increasing of P_i values, a heterogeneous distribution of awareness in the plane $P_i - Z_{o_i}$, and a higher density of hubs nodes, in focused roles with higher awareness values. This trend is more evident in the case of awareness based on data and it means that, decreasing the homophily, it is possible to find out a higher presence of focused hubs, more aware in the case

of awareness based on data than in baseline statistics case, since the latter suffers the effect of attention randomness. As a result of a collective process around an extraordinary event, in (e) plot we depict this effect, showing how the most aware nodes are either more participatory or more central.

In Fig. 5 we display the phase diagram in the plane $\rho - \beta - \lambda$, according to the co-evolving processes of epidemics and awareness in the weighted multiplex network M_1 , and its interdependence with the collective attention dynamics in M_2 . In both panels we show how the double heterogeneity, in terms of infection and awareness rates, allows delaying the contagion outbreak and reducing the density of infected nodes. An increasing of awareness rate results in a decreasing of infection rate up to a specific value of awareness rate, as showed in (a) plot, in which the awareness measures for each node derived from the baseline statistics in a scale-free structure. This means that the remarkable shape is due to the dependence of the heterogeneous susceptibility on the awareness and, when the probability of acquiring more awareness exceed a threshold, this leads to a more probable shift in attention in correlated topics, resulting in an increasing of ρ , the density of infection. Differently, in the (b) plot of Fig. 5, which is referred to the case of awareness measures derived from data, that effect vanishes due to more homogenized trend of awareness values around topic linked to the extraordinary event occurrence, and with an increasing of awareness rate we have a decreasing of the epidemics trend. In Fig. 6 we show the trend of infection rate in function of awareness rate, varying the homophily value. The homophily, representing a human-related factor of our modeling approach enables the structural investigation of how the connections are forged in social networks. For that reason, we vary its standard deviation associated to traits, that define the similarity between a

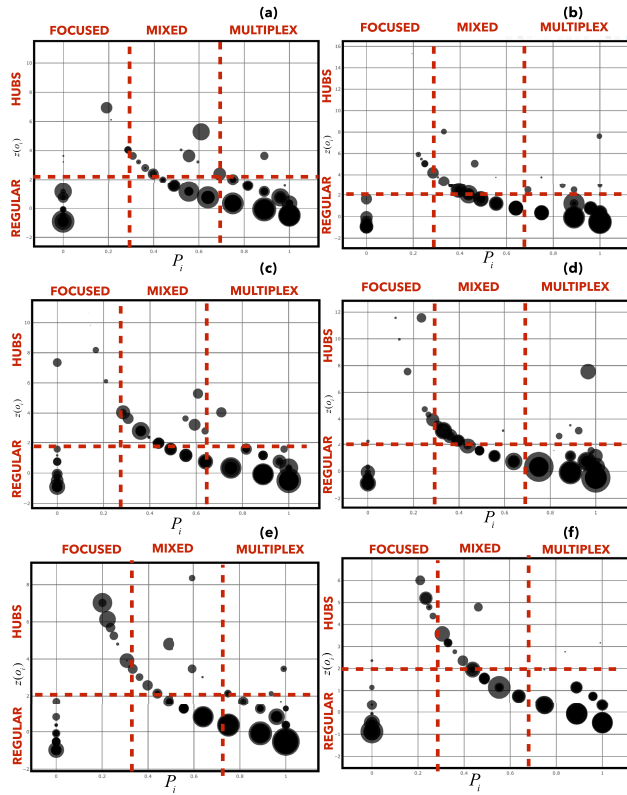


FIGURE 4. Roles of nodes. We show a cartography of the roles of the nodes in the multiplex network M_1 obtained by plotting, for each node, the multiplex participation coefficient P_i versus the Z-score of the total overlapping degree $z(o_i)$. Even if two nodes have exactly the same value of $z(o_i)$, they can have different roles, as indicated in figure, according to the value of the multiplex participation coefficient. We compare the plots (a)-(c)-(e), which are referred to the case of data-driven awareness measures, with (b)-(d)-(f), which represent the roles of nodes in the baseline case as explained in Section IV-A. We can identify, as put in evidence in the plots, the focused, mixed, multiplex, regular or hubs nodes. The size of each node indicates the awareness measure aw_i and its distribution, depending on the collective attention, structurally and dynamically analyzed in M_2 . In both cases, in (a)-(c)-(e) and in (b)-(d)-(f), from top to bottom panels, we decrease the homophily measures among nodes.

pair of nodes showing the resulting trend in the plane $\lambda - \beta$. We take into consideration the two cases of the baseline and the data-driven in reference to the awareness of the users, varying the homophily standard deviation. In the plot (a), referred to the baseline case, we find out that by decreasing the homophily (from green curve to red curve), up to a threshold, the increasing of awareness rate have a minor impact on the infection rate since the population interests are distributed around different topics and the collective attention is shaped as a more heterogenous structure. Differently, in (b), that is based on data mined from Twitter communications under COVID-19, we figure out that by increasing the awareness rate, we can observe a decreasing in infection rate. By decreasing the homophily, the awareness impact is lower than the baseline case, unveiling that the collective dynamics of attention and awareness is more homogenous around the COVID-19.

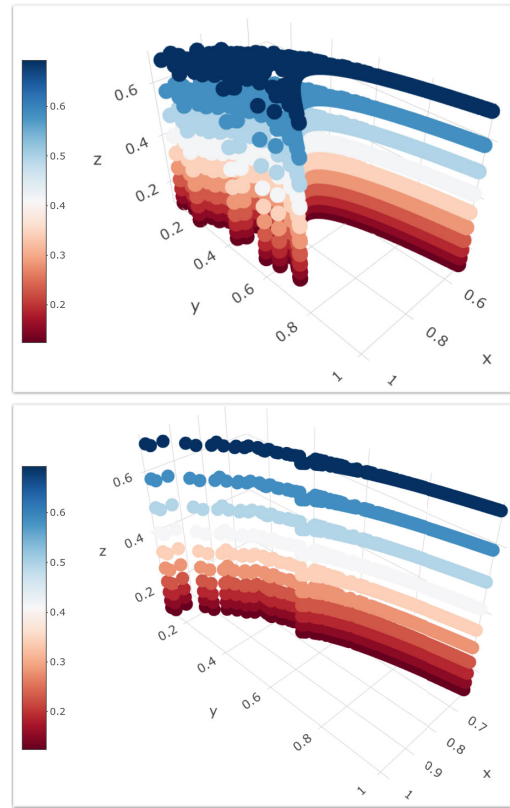


FIGURE 5. Phase Diagrams $\lambda - \beta - \rho$. Plots, both illustrate in the panel (a) and (b), show the density of infected nodes ρ , ranging from red to blue, in function of awareness rate λ , and the investment fraction rate β , as a result of the dynamical co-evolution of the two spreading processes, epidemics and awareness, in the multiplex network M_1 . We compare the cases of awareness measures aw_i , extracted from a null model or baseline, as in (a) plot, and the extracted from data-driven approach ones, as showed in (b) plot. Plots are obtained using MMCA method and MC simulations as detailed in Section III-C3.

In Fig.7, we exploit an heatmap to summarize the collective attention dynamics, during temporal window T , divided in seven-time intervals, as detailed in Table 4, around COVID-19. The heatmap is based on user-generated data from Twitter communications and queries from Google Trends (see Section IV-A). This plot allows for comparing different network metrics, such as tweets volume, networks size, collective attention from Twitter and Google Trends, ranging the different normalized values into coloured bands. We depict the different colour shades in the bands for each geographical state considered, in reference to each time interval to underline the changes in time, in correlation with the red points, that indicates in which time interval falls the first case of COVID-19 officially, reported to WHO. The grey bands point out the lack of values on collective attention, in the range that covers t_1 and t_2 . Following the assumption of our model, this puts in evidence that in the sampled users set we take into consideration users which have a role in the rising of collective attention and consequently in the awareness and epidemics dynamics. For that reason, although there are collective attention data on COVID-19 in t_1 and t_2 , these are

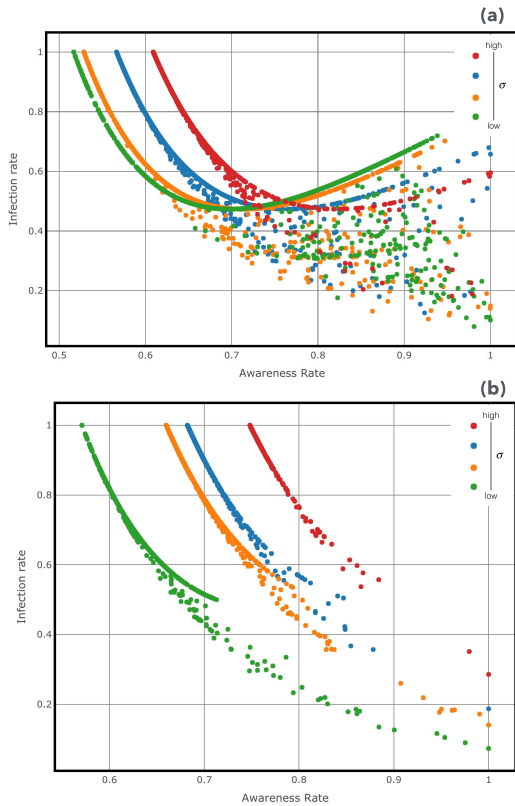


FIGURE 6. Diagrams $\lambda - \beta$. The plots show, in the plane $\lambda - \beta$, the trend of infection rate in function of awareness rate, respectively in both cases of awareness measures of nodes (a), the baseline case and (b), the data-driven. In both cases we set the following parameters $\beta = 0.8, \lambda = 0.4, \delta = 0.2, \gamma = 0.5$, and the standard deviation of homophily value is ranging from a low ($\sigma = 0.2$) to high ($\sigma = 30$) values (curves from green to red) in both (a) and (b) plots.

not referred to relevant nodes in dissemination. Moreover, we highlight, through the red points, the time in which each state goes from being observer to affected community. Thus, considering the tweets volume and the network size metrics, we shed light on a shifting in trend after t_6 , detecting a delay of two time intervals from the time when the majority of events occurred. Focusing on the collective attention metrics, stated for sampled set of users, we deep inside the patterns of collective dynamics zooming the shape of reaction in response to the events occurred. This shows a growing interest, that although it is not free from delay, follows the timing of COVID-19 appearance. In Fig. 8 we illustrate for each state the evolution of the collective attention, mined from the analysis of user-generated data from Twitter communications, in different crucial time intervals, t_3, t_5 and t_7 , that fall within the temporal window T . We exhibit the representation of the weighted graph networks captured by the use of different co-adopted hashtags, filtered following the data-driven approach as underlined in Section IV-A. The Fig. 8 shows, looked at it from top to down, the dynamical evolution of co-adoption around the topic, reflecting the scattered attention among various co-adopted hashtags linked to COVID-19 (see Supplementary Information). It highlights a

quantitative measure of users’ attended topics, capturing it close to the events in t_3 , after the events happened in t_5 , and finally, when the collective attention get up to speed up to t_7 . The graph networks populate the layers of weighted multiplex networks M_2 , following the assumption of our model (see Section III-B2), introducing quantitative dynamical statistical parameters referred to the collective attention impacting on the dynamical behaviours of the users in the weighted multiplex network M_1 . The Fig. 8 exhibits the rising of the digital traces allowing for exploring the users’ interest into co-adopted various hashtags about COVID-19, representing the willingness to expose their attention to as many people as possible, thus, increasing the likelihood to give a boost to a collective process. The Table 5 lists various information about the response time, in terms of both strategical measures and social collective attention, the occurrence of COVID-19 cases and the awareness reactivity to the emergency occurrence, also with the impact of a social network marker. For the purpose of comparing the public response to the topic around COVID-19 across countries selected, we compare the time of the cases reported to WHO for each country and the peak response in terms of attention, from Twitter and Google Trends, as a result of the analysis of user-generated data. We also consider the “Delay from China reported first case”, which represents the delay of the first response of the collective attention from the China reported first case, the “Starting lockdown-quarantine measures”, which is the time when the countries decide to start the strategical measures, and the “Delay of lockdown measures from the reported internal cases for each country”. Moreover, we define the awareness reactivity awr_i , for a country i as a statistical parameter which quantify a measure of responsiveness based on information entropy H , computed on the basis of data extracted from social media platforms considered, as Twitter and Google Trends, over the elapsed time from first reported case of each country. This parameter weighs the speed at which a country knowingly took notice of the emergency, deciding to start strategical measures, to assess the optimal policies while minimizing the output costs of the protective strategies. High values of awareness reactivity awr_i means that there was a fast response to the emergency, that matches high entropy of content shared in online social media platforms and collective attention dynamics, and a high alertness impacting the timing of strategical safety measures. Moreover, taking into consideration the elapsed time between the first peak of the collective attention from the first reported case for each country, we can calculate the awareness reactivity by subtracting this delay from the other delays considered. As indicated in the last column, the resulting value is expressed as an awareness reactivity growth rate, and we find out in most of the cases a percentage of increase of awareness reactivity. In case of China, we see a percentage in decrease of awareness reactivity since it is the country that has the first peak of collective attention response after its first case reported in WHO, differently from the other countries selected. This value represent the impact on the awareness reactivity, finding out

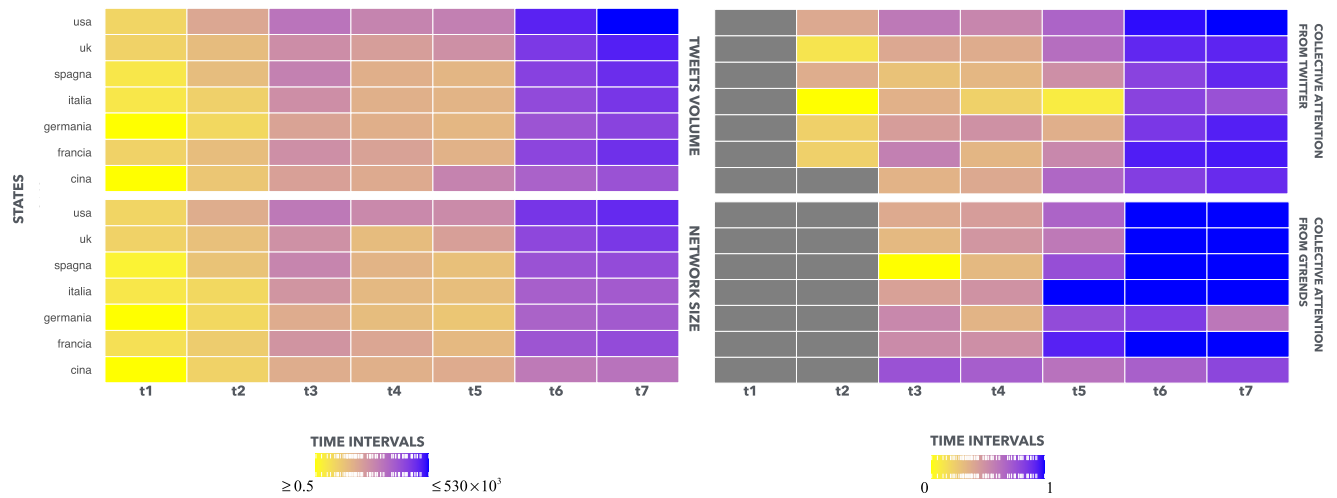


FIGURE 7. The changes of collective attention from user-generated data during time intervals around COVID-19. For each considered state, on time intervals basis, from t_1 to t_7 , as argued in section IV-A, we display the heatmap, with colours ranging from yellow to blue, and in accordance with the two referred deviation bars of network metrics values considered in the data-driven approach, as tweets volume, network size, collective attention from Twitter and collective attention from Google Trends. The colour shades of the tweets volume and network size are referred to the entire population observed, differently from the other two metrics, collective attention from Twitter and Google Trends in which we consider the sampled users population. See section IV-A for details.

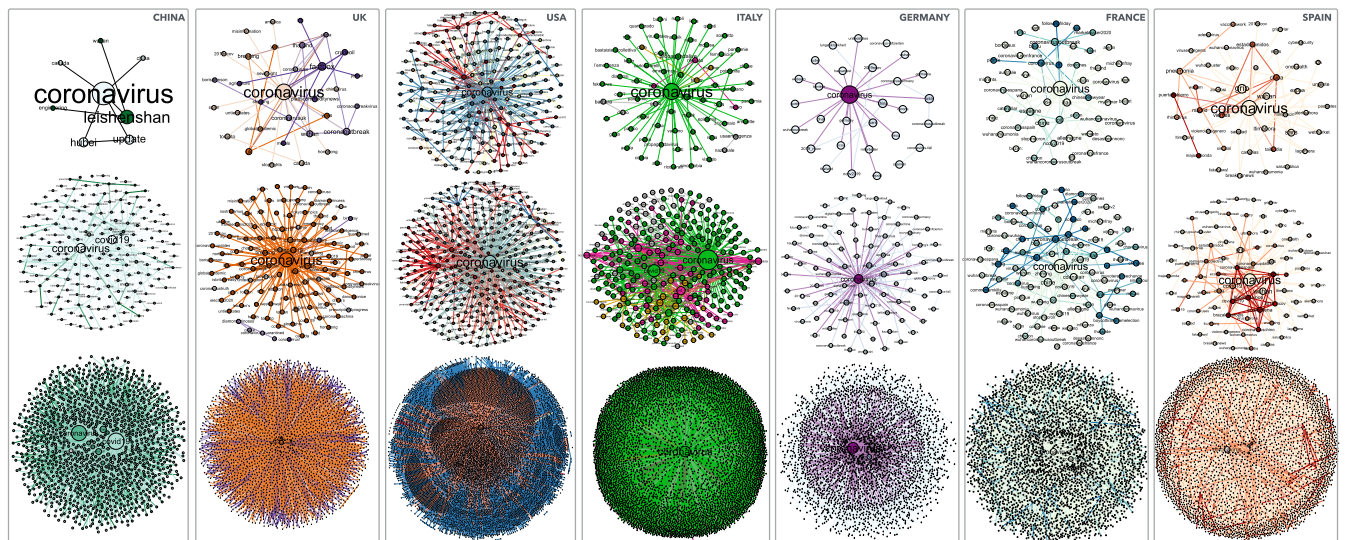


FIGURE 8. Collective attention and Hashtags co-adoption shifting in time intervals around COVID-19. We display three temporal changes, referred to time intervals $t_3 - t_5 - t_7$, respectively corresponding to graphs in top, middle, bottom, for each state, of the weighted multiplex M_2 , composed by seven layers, represented by the co-adoption network graph, extracted from Twitter user-generated data, from sampled users come from each state considered in the model (CHINA, UK, USA, ITALY, GERMANY, FRANCE, SPAIN).

the speeding up of the preparedness and responsiveness that would have had if it had considered the effects of the collective attention and awareness dynamics, that is a rising social measure of the public interest around an emergency that would soon have arrived. The figure 9 shows the awr_i changes and the social marker impact on its growth percentage on the different time windows considered (see Table 4) in function of the collective attention extracted from the data-driven approach. The figure graphically displays the variation of the awr_i with regards to COVID-19 based on awareness

dynamics, considering the impact of the social marker (black line) and not (red line). In particular, the figure represents a prediction of how the awr_i for each state would increase or decrease, graphically representing the impact of the social marker, if the attention peak would have been in different time intervals. As we can see, in the cases b), c), d), e), f), g) in the time intervals previous respect to the first reported case, when the states are in the condition of observer the social marker's effect produces an increase in awr_i and when the states become affected causes a decrease over time. The only

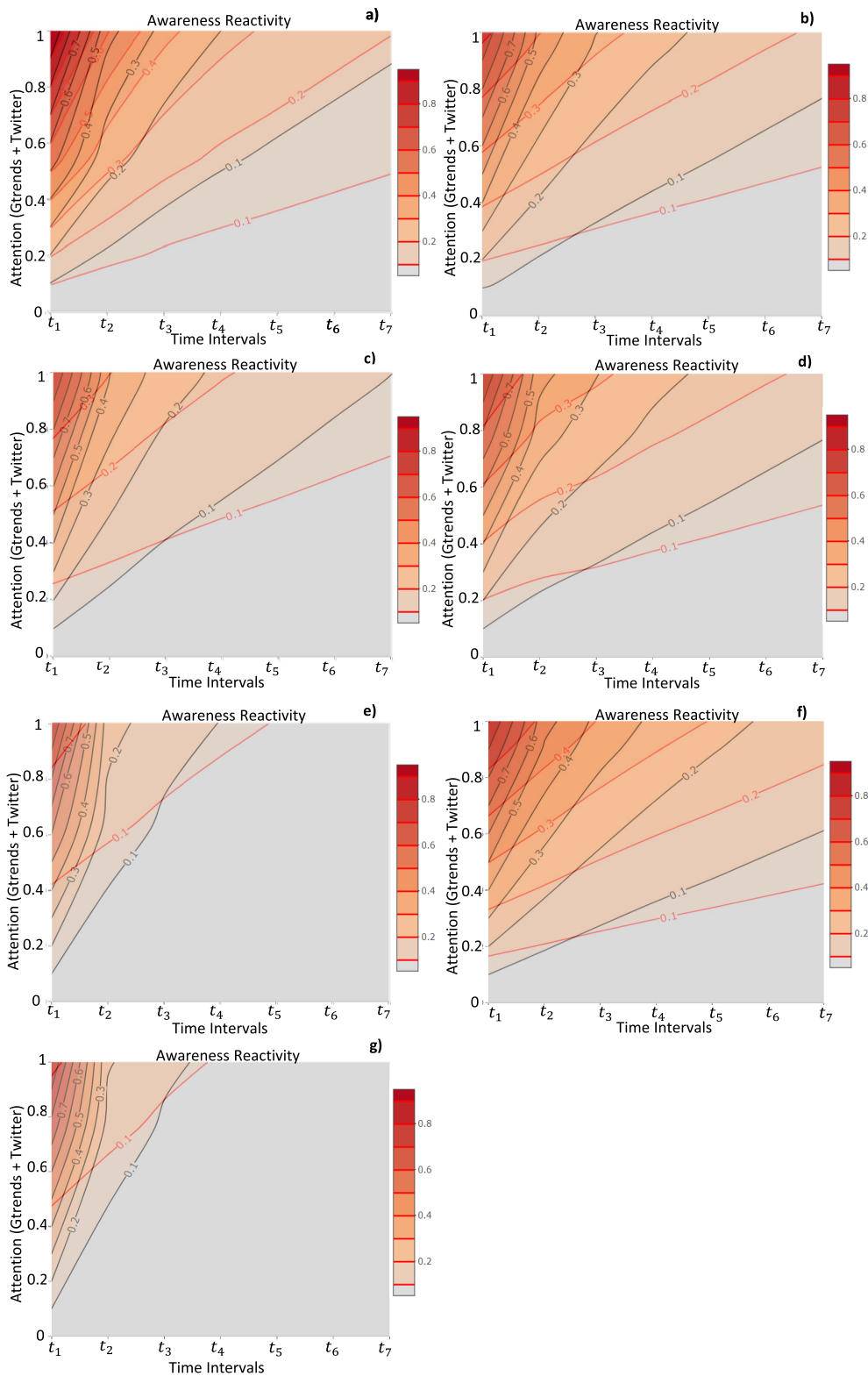


FIGURE 9. Social marker impact on the Awareness Reactivity Growth Rate. The contour plots show the awareness reactivity changes and the social marker impact on the different time windows in function of the collective attention extracted through the data retrieved from Twitter and Google Trends. Each plot is referred to the different states under investigation ((a)China, b) France, c) UK, d) Germany, e) Spain, f) USA, g) Italy), the red lines represents the values of awareness reactivity and the black lines the values of awareness reactivity taking into consideration the impact of the social marker. The area enclosed between them is its variation, in terms of growth rate or decrease.

TABLE 5. Response time and awareness reaction time.

States	China Reported to WHO	Internal cases reported to WHO	First Response (from Twitter)	Peak Response (from Google Trends)	Delay from China Reported First Case	Start Lockdown - Quarantine Measures	Delay of lock-down/quarantine measures from first internal case	Awareness Reactivity	Awareness Reactivity(+/-) social marker impact (growth rate in %)
China	31-Dec-2019	-	13-Jan-2020	17-Jan-2020	13 days	23-Jan-2020	23 days	0.93	-26%
USA	-	23-Jan-2020	04-Jan-2020	19-Jan-2020	4 days	(19-24)-Mar-2020	(56-61) days	(0.70 - 0.65)	+46% - +41%
UK	-	01-Feb-2020	13-Jan-2020	20-Jan-2020	13 days	23-Mar-2020	51 days	0.73	+39%
Italy	-	31-Jan-2020	13-Jan-2020	21-Jan-2020	13 days	9-Mar-2020	38 days	0.57	+49%
Germany	-	28-Jan-2020	13-Jan-2020	19-Jan-2020	13 days	23-Mar-2020	55 days	0.45	+28%
France	-	25-Jan-2020	13-Jan-2020	20-Jan-2020	13 days	17-Mar-2020	52 days	1	+22%
Spain	-	01-Feb-2020	01-Jan-2020	20-Jan-2020	01 days	14-Mar-2020	42 days	0.86	+258%

exception is represented by China (a) which is affected in the whole-time interval under investigation and has a peak of attention that is always lagging behind the first reported case to the WHO. In this case, taking into consideration the effect of the social marker, the awr_i decreases from the first time interval.

V. CONCLUSION

Our work has investigated and modelled a systematic methodology, resulting from a complex modeling approach. Exploiting the social networks and communications technologies, we have proposed the understanding of interplay between the collective attention dynamics and the two co-evolving spreading processes as awareness and epidemics, in two interdependent and heterogeneous weighted multiplex networks, under emergency situations in consequence of the occurrence of an extraordinary event as the ongoing COVID-19 epidemics. The findings demonstrate how the proposed modeling and data-driven approaches, and the mathematical framework represent a complex digital observatory to detect a social network marker exploiting the digital traces of behaviours and connections of people, depicting the possible collective social response. We identify from dynamical collective patterns, social networks markers to optimally schedule effective crisis communications, facilitating timely crisis response planning, such as the decision of a time warning, based on the trends of attention and awareness. The rising of the collective attention in case of shocking event, as a terrorist attacks, mass shootings or earthquakes takes place immediately during or at the end of the event, and the fast awareness acquired on the circumstances could have the strength to limit the consequences, but does not change the event. Differently, a case similar to a long-lasting event as COVID-19 epidemics, the acquisition of awareness over time, conveyed by the attention and its maintenance, could change the flow of epidemics. This can improve the future preparedness plans, risk factors assessment and actionable strategies for delaying or stopping the spreading. In our work, the non-physical distancing factors,

as attention and awareness play an important part in mitigating the spreading, and especially when physical distancing measures are relaxed, the dynamics of collective behaviours in social networks can produce a flow of information that can lead to a shape of attention to be allocated in an efficient way for the affected communities. Physical social distancing strategies, although we do not assume in our work, represent a key point for our future research that could be directed towards the understanding of the role of the location-specific physical distancing policies and the joint impact of real and virtual collective dynamics on awareness. As future work, we plan to conduct a more comprehensive study on collective attention network, comparing different types of events, and the co-occurrence of some of them. This offers more insights on human behaviours and interests leading to an optimally management of the dissemination and the discrimination of the contents, by including mechanisms to improve users' cognitive ability. Moreover, weighing the dynamics of real and virtual conditions, considering physical strategies, as the social distancing, and non-physical strategies, we will aim at detecting the configurations of social dilemma and network structures that conduct to the emergence and sustainability of human cooperation. This can result in an increasing of discriminating capacity for users to identify better information in comparison with low-quality ones or misinformation. We envision to consider clustering methods and community detection algorithms, in order to introduce cluster-specific preparedness plans.

CODE AVAILABILITY

To build the model, do computation and obtain our results we used the programming language R and the IDE RStudio. The figures were generated thanks to the package Plotly and the software Gephi [79]–[82]. The developed methodology in terms of multiplex representation, spreading processing modeling and data-driven approach applied to the user-generated data extracted from Twitter communications and Google Trends, under COVID-19 epidemics, is available to the

scientific community at the following repository: <https://github.com/UnictScata/Collective-Attention-Awareness-and-Epidemics-Spreading-in-the-Multiplex-Social-Network>.

REFERENCES

- [1] D. Hu, X. Lou, Z. Xu, N. Meng, Q. Xie, M. Zhang, Y. Zou, J. Liu, G. Sun, and F. Wang, "More effective strategies are required to strengthen public awareness of COVID-19: Evidence from Google Trends," *J. Global Health*, vol. 10, no. 1, 2020, Art. no. 011003, doi: [10.2139/ssrn.3550008](https://doi.org/10.2139/ssrn.3550008).
- [2] K. Prem, Y. Liu, T. W. Russell, A. J. Kucharski, R. M. Eggo, N. Davies, S. Flasche, S. Clifford, C. A. Pearson, and J. D. Munday, "The effect of control strategies to reduce social mixing on outcomes of the COVID-19 epidemic in Wuhan, China: A modelling study," *The Lancet Public Health*, vol. 5, no. 5, pp. e261–e270, 2020, doi: [10.1016/S2468-2667\(20\)30073-6](https://doi.org/10.1016/S2468-2667(20)30073-6).
- [3] K. Sun, J. Chen, and C. Viboud, "Early epidemiological analysis of the coronavirus disease 2019 outbreak based on crowdsourced data: A population-level observational study," *Lancet Digit. Health*, vol. 2, no. 4, pp. e201–e208, 2020, doi: [10.1016/S2589-7500\(20\)30026-1](https://doi.org/10.1016/S2589-7500(20)30026-1).
- [4] W. H. O. of Library and H. L. Services. *Pneumonia of Unknown Cause China Emergencies Preparedness, Response, Disease Outbreak News*. Accessed: Oct. 1, 2020. [Online]. Available: <https://www.who.int/csr/don/05-january-2020-pneumonia-of-unknown-cause-china/en/>
- [5] C. for Systems Science and E. C. at Johns Hopkins University. *COVID-19 Dashboard*. Accessed: Oct. 1, 2020. [Online]. Available: <https://gisanddata.maps.arcgis.com/apps/opsdashboard/index.html#/bda7594740fd40299423467b48e9ecf6>
- [6] N. Kandel, S. Chungong, A. Omaar, and J. Xing, "Health security capacities in the context of COVID-19 outbreak: An analysis of international health regulations annual report data from 182 countries," *Lancet*, vol. 395, no. 10229, pp. 1047–1053, Mar. 2020, doi: [10.1016/S0140-6736\(20\)30553-5](https://doi.org/10.1016/S0140-6736(20)30553-5)
- [7] W. H. O. of Library and H. L. Services. *Global Preparedness Monitoring Board. Annual Report on Global Preparedness for Health Emergencies*. Accessed: Oct. 1, 2020. [Online]. Available: <https://apps.who.int/gpmb/assets/annual-report/GPMB-annualreport-2019>
- [8] S. P. Layne, J. M. Hyman, D. M. Morens, and J. K. Taubenberger, "New coronavirus outbreak: Framing questions for pandemic prevention," *Sci. Transl. Med.*, vol. 12, no. 534, Mar. 2020, Art. no. eabb1469, doi: [10.1126/scitranslmed.abb1469](https://doi.org/10.1126/scitranslmed.abb1469).
- [9] S. Funk, E. Gilad, and V. A. A. Jansen, "Endemic disease, awareness, and local behavioural response," *J. Theor. Biol.*, vol. 264, no. 2, pp. 501–509, May 2010.
- [10] A.-L. Barabási, "Scale-free networks: A decade and beyond," *Science*, vol. 325, no. 5939, pp. 412–413, Jul. 2009.
- [11] N. A. Christakis and J. H. Fowler, "Social contagion theory: Examining dynamic social networks and human behavior," *Statist. Med.*, vol. 32, no. 4, pp. 556–577, Feb. 2013.
- [12] L. Yang, Z. Yu, M. A. El-Meligy, A. M. El-Sherbeeny, and N. Wu, "On multiplexity-aware influence spread in social networks," *IEEE Access*, vol. 8, pp. 106705–106713, 2020.
- [13] M. D. Vicario, W. Quattrociocchi, A. Scala, and F. Zollo, "Polarization and fake news: Early warning of potential misinformation targets," *ACM Trans. Web*, vol. 13, no. 2, p. 10, 2019.
- [14] M. Del Vicario, A. Bessi, F. Zollo, F. Petroni, A. Scala, G. Caldarelli, H. E. Stanley, and W. Quattrociocchi, "The spreading of misinformation online," *Proc. Nat. Acad. Sci. USA*, vol. 113, no. 3, pp. 554–559, 2016.
- [15] E. M. Hill, F. E. Griffiths, and T. House, "Spreading of healthy mood in adolescent social networks," *Proc. Roy. Soc. B, Biol. Sci.*, vol. 282, no. 1813, Aug. 2015, Art. no. 20151180.
- [16] M. Scatà, A. Di Stefano, P. Liò, and A. La Corte, "The impact of heterogeneity and awareness in modeling epidemic spreading on multiplex networks," *Sci. Rep.*, vol. 6, no. 1, p. 37105, Dec. 2016.
- [17] M. Scatà, A. Di Stefano, A. La Corte, and P. Liò, "Quantifying the propagation of distress and mental disorders in social networks," *Sci. Rep.*, vol. 8, no. 1, p. 5005, Dec. 2018.
- [18] J. H. Fowler and N. A. Christakis, "Dynamic spread of happiness in a large social network: Longitudinal analysis over 20 years in the Framingham Heart Study," *Bmj*, vol. 337, p. a2338, Dec. 2008, doi: [10.1136/bmj.a2338](https://doi.org/10.1136/bmj.a2338).
- [19] L.-X. Yang, P. Li, X. Yang, Y. Xiang, and Y. Y. Tang, "Simultaneous benefit maximization of conflicting opinions: Modeling and analysis," *IEEE Syst. J.*, vol. 14, no. 2, pp. 1623–1634, Jun. 2020.
- [20] A. Dadlani, M. S. Kumar, M. G. Maddi, and K. Kim, "Mean-field dynamics of inter-switching memes competing over multiplex social networks," *IEEE Commun. Lett.*, vol. 21, no. 5, pp. 967–970, May 2017.
- [21] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani, "Epidemic processes in complex networks," *Rev. Mod. Phys.*, vol. 87, no. 3, p. 925, 2015.
- [22] R. M. Anderson, R. M. May, and B. Anderson, *Infectious Diseases of Humans: Dynamics and Control*, vol. 28. Hoboken, NJ, USA: Wiley, 1992.
- [23] F. Brauer, C. Castillo-Chavez, and C. Castillo-Chavez, *Mathematical Models in Population Biology and Epidemiology* (Text in Applied Mathematics), vol. 40. New York, NY, USA: Springer, 2012, p. 508.
- [24] A. Vespignani, "Modelling dynamical processes in complex socio-technical systems," *Nature Phys.*, vol. 8, no. 1, pp. 32–39, Jan. 2012.
- [25] M. Zheng, C. Wang, J. Zhou, M. Zhao, S. Guan, Y. Zou, and Z. Liu, "Non-periodic outbreaks of recurrent epidemics and its network modelling," *Sci. Rep.*, vol. 5, no. 1, p. 16010, Dec. 2015.
- [26] M. E. J. Newman and C. R. Ferrario, "Interacting epidemics and co-infection on contact networks," *PLoS ONE*, vol. 8, no. 8, Aug. 2013, Art. no. e71321.
- [27] S. N. Dorogovtsev and J. F. Mendes, *Evolution of Networks: From Biological Nets to the Internet and WWW*. Oxford, U.K.: Oxford Univ. Press, 2013, doi: [10.1093/acprof:oso/9780198515906.001.0001](https://doi.org/10.1093/acprof:oso/9780198515906.001.0001).
- [28] E. Campbell and M. Salathé, "Complex social contagion makes networks more vulnerable to disease outbreaks," *Sci. Rep.*, vol. 3, no. 1, p. 1905, Dec. 2013.
- [29] C. Granell, S. Gómez, and A. Arenas, "Dynamical interplay between awareness and epidemic spreading in multiplex networks," *Phys. Rev. Lett.*, vol. 111, no. 12, Sep. 2013, Art. no. 128701.
- [30] C. Granell, S. Gómez, and A. Arenas, "Competing spreading processes on multiplex networks: Awareness and epidemics," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 90, no. 1, Jul. 2014, Art. no. 012808.
- [31] F. Battiston, V. Nicosia, and V. Latora, "The new challenges of multiplex networks: Measures and models," *Eur. Phys. J. Special Topics*, vol. 226, no. 3, pp. 401–416, Feb. 2017.
- [32] S. Boccaletti, G. Bianconi, R. Criado, C. I. del Genio, J. Gómez-Gardeñes, M. Romance, I. Sendiña-Nadal, Z. Wang, and M. Zanin, "The structure and dynamics of multilayer networks," *Phys. Rep.*, vol. 544, no. 1, pp. 1–122, Nov. 2014.
- [33] D. Zhao, L. Li, H. Peng, Q. Luo, and Y. Yang, "Multiple routes transmitted epidemics on multiplex networks," *Phys. Lett. A*, vol. 378, no. 10, pp. 770–776, Feb. 2014.
- [34] G. Menichetti, D. Remondini, P. Panzarasa, R. J. Mondragón, and G. Bianconi, "Weighted multiplex networks," *PLoS ONE*, vol. 9, no. 6, Jun. 2014, Art. no. e97857.
- [35] X. He and Y.-R. Lin, "Measuring and monitoring collective attention during shocking events," *EPJ Data Sci.*, vol. 6, no. 1, p. 30, Dec. 2017.
- [36] A. Di Stefano, M. Scatà, A. La Corte, P. Liò, E. Catania, E. Guardo, and S. Pagano, "Quantifying the role of homophily in human cooperation using multiplex evolutionary game theory," *PLoS ONE*, vol. 10, no. 10, 2015, Art. no. e0140646.
- [37] A. Santoro and V. Nicosia, "Algorithmic complexity of multiplex networks," *Phys. Rev. X*, vol. 10, no. 2, Jun. 2020, Art. no. 021069.
- [38] V. Nicosia, P. S. Skardal, A. Arenas, and V. Latora, "Collective phenomena emerging from the interactions between dynamical processes in multiplex networks," *Phys. Rev. Lett.*, vol. 118, no. 13, Mar. 2017, Art. no. 138302.
- [39] F. Battiston, M. Perc, and V. Latora, "Determinants of public cooperation in multiplex networks," *New J. Phys.*, vol. 19, no. 7, Jul. 2017, Art. no. 073017.
- [40] G. Bianconi, *Multilayer Networks: Structure and Function*. Oxford, U.K.: Oxford Univ. Press, 2018, doi: [10.1093/oso/9780198753919.001.0001](https://doi.org/10.1093/oso/9780198753919.001.0001).
- [41] M. De Domenico, C. Granell, M. A. Porter, and A. Arenas, "The physics of spreading processes in multilayer networks," *Nature Phys.*, vol. 12, no. 10, pp. 901–906, Oct. 2016.
- [42] P. Brodka, K. Musial, and J. Jankowski, "Interacting spreading processes in multilayer networks: A systematic review," *IEEE Access*, vol. 8, pp. 10316–10341, 2020.
- [43] M. Pósfai, J. Gao, S. P. Cornelius, A.-L. Barabási, and R. M. D'Souza, "Controllability of multiplex, multi-time-scale networks," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 94, no. 3, Sep. 2016, Art. no. 032316.
- [44] P. Crépey, F. P. Alvarez, and M. Barthélemy, "Epidemic variability in complex networks," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 73, no. 4, Apr. 2006, Art. no. 046131.

- [45] W. Wang, M. Tang, H. Yang, Y. Do, Y.-C. Lai, and G. Lee, "Asymmetrically interacting spreading dynamics on complex layered networks," *Sci. Rep.*, vol. 4, no. 1, p. 5097, May 2015.
- [46] G. Ren and X. Wang, "Epidemic spreading in time-varying community networks," *Chaos: Interdiscipl. J. Nonlinear Sci.*, vol. 24, no. 2, Jun. 2014, Art. no. 023116.
- [47] H. Huang, Z. Yan, Y. Chen, and F. Liu, "A social contagious model of the obesity epidemic," *Sci. Rep.*, vol. 6, no. 1, p. 37961, Dec. 2016.
- [48] A. L. Hill, D. G. Rand, M. A. Nowak, and N. A. Christakis, "Infectious disease modeling of social contagion in networks," *PLoS Comput. Biol.*, vol. 6, no. 11, Nov. 2010, Art. no. e1000968.
- [49] Q. Guo, Y. Lei, X. Jiang, Y. Ma, G. Huo, and Z. Zheng, "Epidemic spreading with activity-driven awareness diffusion on multiplex network," *Chaos: Interdiscipl. J. Nonlinear Sci.*, vol. 26, no. 4, Apr. 2016, Art. no. 043110.
- [50] Q. Wu, X. Fu, M. Small, and X.-J. Xu, "The impact of awareness on epidemic spreading in networks," *Chaos: Interdiscipl. J. Nonlinear Sci.*, vol. 22, no. 1, Mar. 2012, Art. no. 013101.
- [51] W. Wang, Q.-H. Liu, S.-M. Cai, M. Tang, L. A. Braunstein, and H. E. Stanley, "Suppressing disease spreading by using information diffusion on multiplex networks," *Sci. Rep.*, vol. 6, no. 1, p. 29259, Jul. 2016.
- [52] H. Yang, M. Tang, and T. Gross, "Large epidemic thresholds emerge in heterogeneous networks of heterogeneous nodes," *Sci. Rep.*, vol. 5, no. 1, p. 13122, Oct. 2015.
- [53] X. Qiu, D. F. M. Oliveira, A. Sahami Shirazi, A. Flammini, and F. Menczer, "Limited individual attention and online virality of low-quality information," *Nature Hum. Behav.*, vol. 1, no. 7, p. 0132, Jul. 2017.
- [54] K. Sasahara, Y. Hirata, M. Toyoda, M. Kitsuregawa, and K. Aihara, "Quantifying collective attention from tweet stream," *PLoS ONE*, vol. 8, no. 4, Apr. 2013, Art. no. e61823.
- [55] M. Tizzoni, A. Panisson, D. Paolotti, and C. Cattuto, "The impact of news exposure on collective attention in the united states during the 2016 zika epidemic," *PLOS Comput. Biol.*, vol. 16, no. 3, Mar. 2020, Art. no. e1007633.
- [56] J. Lehmann, B. Gonçalves, J. J. Ramasco, and C. Cattuto, "Dynamical classes of collective attention in Twitter," in *Proc. 21st Int. Conf. World Wide Web (WWW)*, 2012, pp. 251–260.
- [57] S. D. Roy, G. Lotan, and W. Zeng, "The attention automaton: Sensing collective user interests in social network communities," *IEEE Trans. Netw. Sci. Eng.*, vol. 2, no. 1, pp. 40–52, Jan. 2015.
- [58] Z. Da, J. Engelberg, and P. Gao, "In search of attention," *J. Finance*, vol. 66, no. 5, pp. 1461–1499, Oct. 2011.
- [59] H. Choi and H. Varian, "Predicting the present with Google trends," *Econ. Rec.*, vol. 88, pp. 2–9, Jun. 2012.
- [60] A. Mavragani and G. Ochoa, "Forecasting AIDS prevalence in the united states using online search traffic data," *J. Big Data*, vol. 5, no. 1, p. 17, Dec. 2018.
- [61] G. Caldarelli, *Scale-Free Networks: Complex Webs in Nature and Technology*. Oxford, U.K.: Oxford Univ. Press, 2007, doi: 10.1093/acprof:oso/9780199211517.001.0001.
- [62] G. Menichetti, D. Remondini, and G. Bianconi, "Correlations between weights and overlap in ensembles of weighted multiplex networks," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 90, no. 6, Dec. 2014, Art. no. 062817.
- [63] P. Törnberg, "Echo chambers and viral misinformation: Modeling fake news as complex contagion," *PLoS ONE*, vol. 13, no. 9, Sep. 2018, Art. no. e0203958.
- [64] M. Cinelli, W. Quattrociochi, A. Galeazzi, C. M. Valensise, E. Brugnoli, A. L. Schmidt, P. Zola, F. Zollo, and A. Scala, "The COVID-19 social media infodemic," 2020, *arXiv:2003.05004*. [Online]. Available: <http://arxiv.org/abs/2003.05004>
- [65] S. Chen, J. Yang, W. Yang, C. Wang, and T. Bärnighausen, "COVID-19 control in China during mass population movements at new year," *Lancet*, vol. 395, no. 10226, pp. 764–766, Mar. 2020.
- [66] Y. Bai, L. Yao, T. Wei, F. Tian, D.-Y. Jin, L. Chen, and M. Wang, "Presumed asymptomatic carrier transmission of COVID-19," *Jama*, vol. 323, no. 14, pp. 1406–1407, 2020.
- [67] E. Mahase, "China coronavirus: WHO declares international emergency as death toll exceeds 200," *BMJ*, vol. 368, p. m408, Jan. 2020.
- [68] D. B. Taylor, "How the coronavirus pandemic unfolded: A timeline," *The New York Times*, Accessed: Oct. 1, 2020. [Online]. Available: <http://www.nytimes.com>
- [69] Y.-R. Lin, B. Keegan, D. Margolin, and D. Lazer, "Rising tides or rising stars?: Dynamics of shared attention on Twitter during media events," *PLoS ONE*, vol. 9, no. 5, May 2014, Art. no. e94093.
- [70] E. Koeze and N. Popper, "The virus changed the way we Internet," *The New York Times*, Apr. 2020.
- [71] Q. Li, X. Guan, P. Wu, X. Wang, L. Zhou, Y. Tong, R. Ren, K. S. Leung, E. H. Lau, and J. Y. Wong, "Early transmission dynamics in Wuhan, China, of novel coronavirus-infected pneumonia," *New England J. Med.*, vol. 382, no. 13, pp. 1199–1207, 2020.
- [72] Z. Yang, Z. Zeng, K. Wang, S.-S. Wong, W. Liang, M. Zanin, P. Liu, X. Cao, Z. Gao, and Z. Mai, "Modified SEIR and AI prediction of the epidemics trend of COVID-19 in China under public health interventions," *J. Thoracic Disease*, vol. 12, no. 3, p. 165, 2020.
- [73] A. J. Kucharski, T. W. Russell, C. Diamond, Y. Liu, J. Edmunds, S. Funk, R. M. Eggo, F. Sun, M. Jit, and J. D. Munday, "Early dynamics of transmission and control of COVID-19: A mathematical modelling study," *lancet Infectious Diseases*, vol. 20, no. 5, pp. 553–558, 2020.
- [74] G. Giordano, F. Blanchini, R. Bruno, P. Colaneri, A. Di Filippo, A. Di Matteo, and M. Colaneri, "Modelling the COVID-19 epidemic and implementation of population-wide interventions in Italy," *Nature Med.*, vol. 26, no. 6, pp. 855–860, Jun. 2020.
- [75] J. Ye, P. Ji, and M. Barthelemy, "Scenarios for a post-COVID-19 world airline network," 2020, *arXiv:2007.02109*. [Online]. Available: <http://arxiv.org/abs/2007.02109>
- [76] Trackmyhashtag. *Corona Virus (COVID-19) Tweet Metadata Compilation 2020*. Accessed: Oct. 1, 2020. [Online]. Available: <https://www.trackmyhashtag.com/blog/twitter-datasets-free/>
- [77] Kaggle. *Coronavirus (COVID19) Tweets*. Accessed: Oct. 1, 2020. [Online]. Available: <https://www.kaggle.com/smid80/coronavirus-covid19-tweets>
- [78] F. Battiston, V. Nicosia, and V. Latora, "Structural measures for multiplex networks," *Phys. Rev. E, Stat. Phys. Plasmas Fluids Relat. Interdiscip. Top.*, vol. 89, no. 3, Mar. 2014, Art. no. 032804.
- [79] R. C. Team, *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria, 2019. [Online]. Available: <https://www.R-project.org/>
- [80] R. Team, *Rstudio: Integrated Development for R*, RStudio, Inc., Boston, MA, USA, vol. 42, p. 14, 2015. [Online]. Available: <http://www.rstudio.com>
- [81] C. Sievert, "Plotly for R," R Package Version 4.0, Tech. Rep., 2018. [Online]. Available: <http://www.plotly-r.com>
- [82] M. Bastian, S. Heymann, and M. Jacomy, "Gephi: An open source software for exploring and manipulating networks," in *Proc. ICWSM*, vol. 8, 2009, pp. 361–362.



MARIALISA SCATÁ (Member, IEEE) received the B.S. and M.S. degrees in telecommunication engineering from the Department of Electrical, Electronic and Computer Engineering (DIEEI), University of Catania, Italy, in 2005 and 2010, respectively, and the Ph.D. degree in computer science and telecommunication engineering from the University of Catania, in 2012, under the supervision of Prof. A. L. Corte, with whom she has been collaborating since 2009. During the Ph.D., she attended several schools and conferences. She made internships at the Computer Laboratory, Department of Computer Science, University of Cambridge, U.K., under the supervision of Prof. P. Liò, with whom she has also been collaborating since 2011. She is currently working as a Postdoctoral Researcher with the DIEEI, University of Catania. She also serves as a reviewer for several top-tier journals, among them IEEE journals, *Scientific Reports Nature*, and *PLoS One*. She has an interdisciplinary approach to research and her interests include bio-inspired models, telecommunications, social networks, complex networks, multi-layer networks, social contagion and epidemic spreading, evolutionary game theory, data mining, machine learning, healthcare applications, and ICT.



game theory, multi-layer networks, data science, and the IoT.

BARBARA ATTANASIO (Member, IEEE) received the B.S. degree in computer engineering and the M.S. degree in telecommunications engineering from the Dipartimento di Ingegneria Elettrica, Elettronica e Informatica (DIEEI), University of Catania, Italy, in 2016 and 2018, respectively, where she is currently pursuing the Ph.D. degree in systems, energy, computer and telecommunications engineering. Her research interests include multi access edge computing, evolutionary



GRAZIA VERONICA AIOSA received the B.S. degree in computer engineering and the M.S. degree in telecommunication engineering from the Dipartimento di Ingegneria Elettrica, Elettronica e Informatica (DIEEI), University of Catania, Italy, in 2016 and 2019, respectively. She is currently an Early Stage Researcher with the DIEEI, University of Catania. Her research interests include multi-layer networks, social contagion, and collective attention.



activities at the University of Catania, he has been involved in coordinating, designing, and developing networks and ICT services. His research interests include telecommunication systems and integrated services, innovative ICT services, and technological innovation.

AURELIO LA CORTE received the degree in electrical engineering, in 1988, and the Ph.D. degree in electronic engineering and computer science, in 1994. He is currently an Associate Professor of telecommunication engineering with the University of Catania, Italy. He has more than 25 years of experience in scientific and teaching activity and specific experience in designing telecommunications networks and systems and in managing complex projects. Responsible for various ICT-related

...