Network properties of Informal Support Networks in Mental Health

Author: Oriol Martínez Pérez
Facultat de Física, Universitat de Barcelona, Diagonal 645, 08028 Barcelona, Spain.

Advisor: Franziska Peter

Abstract: We analyze a dataset of citizens' relations to a set of situations related to mental health. The dataset was collected by a digital tool created in the framework of Citizen Social Science. We generate two networks from it and compute different metrics to study them. In particular, we answer two particular questions: *Do the experiences of young people in mental health differ from those of old people?* and *Are young people and old people in two different bubbles, i.e. are the mental health experiences they know from their respective surroundings different from each other?* Their answers were determined by computing Newman's assortativity coefficient but generalized to weighted networks.

All code is available on https://github.com/oriolmp/TFG_Fisica.

I. INTRODUCTION

Complex network analysis [1–3] is a powerful method to extract knowledge when a large amount of data with non explicit relationships is given. This tool has been widely used among Biology and Social Science during the last years, and it is becoming an important tool in Physics as well, even being the central technique of the the Nobel Prize winners in Physics 2021. In this project, we combine this powerful tool with a new way of doing science, Citizen Science, to study connections among different people related to a variety of mental health situations.

Citizen Science can be described as the involvement of the general public to the scientific research [4]. Typically, science has been delegated only to specialized researchers who do their research independently, with the main part of the population not being able to take an active role in science. Citizen science tries to make science in an open and inclusive way, allowing non-scientific people to be part of the projects, from performing simple tasks such as generating simple data to specific roles in the investigation.

Citizen Science has evolved a lot since about ten years ago when citizens were merely carrying out easy repetitive tasks (like classifying, counting, collecting) for a research project. Citizens are now an active part of the research team and take decisions on the research question, design, and analysis. In the sense that they contribute their highly valuable personal experiences and knowledge during the entire research process, they act as experts in the field. That increases both the range of perspectives that shape the research during all of its stages, as also the practical relevance of the scientific results. Moreover, it increases its impact on society, since it is easier to reach political entities and associations, who may take into consideration the research results. In this way, scientific knowledge is increased and democratized.

Specifically, we focus on Citizen *Social* Science, where non-scientific people take a major role in the research. In the Horizon2020 project *CoActuem per la Salut Mental*

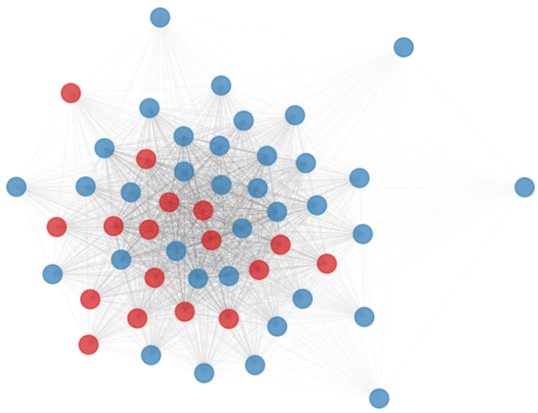

FIG. 1: Network generated with answers to question Q_1 of 68 micro-stories, drawn with the Fruchterman-Reingold force-directed algorithm [5]. The red nodes represent 17 participants in the age bracket 18-44 years, blue dots represent 35 participants in the age bracket 45 - more than 65 years.

30 people directly affected by mental health problems (of themselves or of family members) acted as co-researchers in the project. In the first phase of the project, i.e. the research design phase, they shared their experiences with the aim to improve informal support networks in mental health. With their knowledge, together with scientific and other professionals such as psychologists and writers, they generated a number of short stories describing situations related to mental health.

In the same research phase, a Telegram chatbot was co-created with the purpose of sharing these stories to any citizen who wants to take part in the project, and ask them questions about it. The chatbot runs since July 2021 on a server of the University, collecting the citizens' answers around the clock. In February 2023, a first data release was published on [6]. A limited data analysis and interpretation was performed with the co-researchers in summer 2022. The co-researchers used the results to underpin their political demands that they presented to the Catalonian political authorities in mental health in the first Mental Health Community Assembly [7]. We present here a more in-depth analysis of the dataset, gaining further insight on relations and connections of different actors in informal support networks of mental health.

In this case, complex network analysis is our way to go. Essentially, a network is a collection of elements, the nodes, joined together in pairs by connections: the edges [1]. In mathematical terms, a graph. This representation is well suited to describe real world systems with a large amount of data, allowing us to extract patterns and behaviours from them. For example, one could represent humanity using a network. Each human would be a node which is connected to another human if they know each other.

There are different types of networks, being characterized by how the edges are considered. We can quantize the strength of the connections by assigning a weight to each edge. Following the latter example, we could assign to an edge the score 1 if both humans are family, 2/3 if they are friends and 1/3 if they are acquaintances. We call this type of network a weighted one. Moreover, we can also take into account if the connections are reciprocal or not, and assign a direction to each edge, obtaining a network named as directed.

In our case, an undirected weighted network is generated, where each node is a participant and its edges are weighted depending on their answers to the same stories. Furthermore, each node is enriched with node attributes, which provide social-demographic information of the participants. In Sect. II A, some networks metrics are introduced, which allowed us to study the connectivity in our network. We analyze these metrics to get insights about how people perceive the mental health situations related to the stories. Specifically, we focus on how the age of the participants is reflected on their answers, trying to answer whether different age groups of people behave differently.

II. EXPERIMENTAL

A. Theoretical background

In our work, the following network concepts have been the basis we have worked with to analyse our data.

1. *Common network metrics*

Given the nodes of the network, $\{x_i\}_{1\leq i\leq n}$, and their edge weights among all pairs, $\{w_{ij}\}_{i\neq j}$, the **degree** of a node x_i is defined as [1]

$$d_i = \sum_{j\neq i} w_{ij}$$

The degree quantizes how much a node is connected with the others, being its distribution a power-law in many real world networks.

A **path** between two nodes is defined as a sequence of unique edges that connect them. The **shortest path** that connects two nodes is the path that minimizes the sum of their edge weights, which is called the **path length**. The **diameter** is defined as the longest among all pairwise shortest paths in a network [1]. In this sense, the diameter expresses how compact a network is.

2. *Assortativity*

One may want to analyze whether there is a tendency for nodes in a network to connect with other nodes which are similar to them, i.e. their attributes are alike. Introduced by [8], **assortativity**, also known as homophily in terms of social networks, is a metric which measures this behaviour.

Given a node attribute which can take m different **discrete** options, a_i, $1 \leq i \leq m$, we denote e_{ij} as the fraction of edges that connect a node with attribute a_i to a node with attribute a_j, which satisfy

$$\sum_{ij} e_{ij} = 1 \tag{1}$$

This defines the matrix

$$\mathbf{e} = \begin{pmatrix} e_{11} & e_{12} & \dots & e_{1m} \\ e_{21} & e_{22} & \dots & e_{2m} \\ \vdots & & & \vdots \\ e_{m1} & e_{m2} & \dots & e_{mm} \end{pmatrix} \tag{2}$$

The **assortativity coefficient** is defined as

$$r = \frac{\mathrm{Tr}(\mathbf{e}) - ||\mathbf{e}^2||}{1 - ||\mathbf{e}^2||} \tag{3}$$

where $||\cdot||$ denotes the sum of all the elements of a matrix and $\mathrm{Tr}(\mathbf{e})$ is the sum of all diagonal elements, i.e. the trace of the matrix.

Notice that if all nodes only connect with other nodes that have the same attribute, a situation denoted as **perfect assortativity**, we have a diagonal matrix in (2). Combined with property (1), we have that $\mathrm{Tr}(\mathbf{e}) = 1$ and therefore $r = 1$. Since e_{ij} are fractions, $||\mathrm{Tr}(e)|| < 1$, and the maximum value of r is 1.

$$r_{max} = \frac{1 - ||\mathbf{e}^2||}{1 - ||\mathbf{e}^2||} = 1 \tag{4}$$

On the other hand, if all nodes are connected with nodes with different attribute, we have a hollow matrix, i.e. all its diagonal elements are zero. Therefore, $\mathrm{Tr}(\mathbf{e}) = 0$. Hence $r < 0$, and it is the minimum value of r that can be achieved

$$r_{min} = -\frac{||\mathbf{e}^2||}{1 - ||\mathbf{e}^2||} \tag{5}$$

Such network is called **perfectly diassortative**.

A metric is achieved which quantizes assortativity, laying in the range $r_{min} \leq r \leq 1$. Notice that when $r = 0$, one have a neutral situation where assortativity does not affect the network.

Specifically, for a weighted network, we adapt the assortativity coefficient to work with weights. To do so, we define

$$W_{ij} \equiv \{w_{ij} \text{ s.t. node } x_i \text{ has attribute } a_i \text{ and } x_j \text{ has attribute } a_j\} \tag{6}$$

which allows us to define a weighted fraction:

$$e_{ij} = \frac{\sum_{w \in W_{ij}} w}{2 \sum_{ij} w_{ij}} \tag{7}$$

Considering a matrix as (2) with these weighted fractions, the assortativity coefficient is defined with the same equation (3) as before.

B. Data

All data analyzed here come from the answers given by the chatbot participants. Once a participant joins the chatbot, a socio-demographic survey is presented to them, where they can decide not to answer in any of the questions. Afterwards, short stories related to mental health situations are sent on a daily basis. All participants receive the same content, i.e. the same set of micro-stories in the same order, and are asked to answer two questions: Q_1*: Have you had the same experience?* and Q_2*: And those around you... Has anybody had the same experience?* Each question has the same three possible answers: *Yes (A)*, *Not exactly (B)* and *No (C)*. The number of participants is currently $N_{tot} = 748$ and the number of stories is $n_{tot} = 130$.

In order to homogenize the dataset, a subset is generated selecting a reduced group of participants who all answered to the same stories. As known from other digital Citizen science experiments [9, 10], few chatbot participants answer to all questions. In the experiment, a fixed order of contents was set to ensure a considerable number of answers at least for an intermediate number of stories. For the analysis, we pick a subset of $n = 68$ micro-stories that each were answered by the same $N = 52$ participants.

With the aim of focusing on participants age, we generate a new node attribute derived from their age. The dataset contemplates 6 age ranges, which we group in two main categories, **young** and **old**, as represented in table I. Besides, this grouping allow us to increase our group statistics.

Age range (years)	18-24	25-34	35-44	45-54	55-65	+65
Category	young			old		

TABLE I: Age grouping

C. Network creation

From the answers to Q_1 and Q_2 we generate two networks, G_1 and G_2, that have as nodes the N participants. See Fig. 1. Each pair of nodes i and j is connected by an edge with a weight defined by the following formula:

$$w_{ij}^q = \frac{\#_s AA + \frac{2}{3}\#_s AB + \frac{2}{3}\#_s BA + \frac{1}{3}\#_s BB}{n(N-1)} \tag{8}$$

where $q \in \{Q_1, Q_2\}$ indicates the question that is being considered, and $\#_s AB$ is the number of stories to which participant i answered A and participant j has answered B, and so forth. This formula gives the highest importance to the connection of participants who answered affirmatively to the same questions, which means that both participants share the same experience. It also takes into account when either or both of them answered B, meaning that somehow they experienced something similar. In case that any of the two participants answered C, the added weight for this story will be considered 0.

D. Graph analysis

To get a general insight of our two networks, G_1 and G_2, we explore the above mentioned network measures and collect them in a network card [11], see Tab. II.

TABLE II: Network Card.

Name	Mental Health	
Kind	Undirected, weighted	
Nodes are	Participants of the chatbot	
Links are	Participants' answers similarity	
Link weights are	Defined in Sect. II C	
Graph	G_1	G_2
Number of nodes N	52	52
Number of links $\frac{N(N-1)}{2}$	1326	1326
Degree 25 quantile	7,57	13,82
Degree mean	10,42	16,15
Degree 75 quantile	9,08	10,64
Degree median	10,38	17,56
Connected	Yes	Yes
Diameter	0,54	0,36
Assortativity (degree)	-0.02	-0,02
Node metadata	Social-demographic information	
Link metadata	Zenodo Dataset Link	
Date of creation	2023	
Data generating process	Public Telegram Chatbot	
Ethics	Coactuem Informed Consent	
Funding	European Union's Horizon 2020	
Citation	[6]	

We first compute their node degree and diameter. We observe that G_2 has higher values for both metrics, which

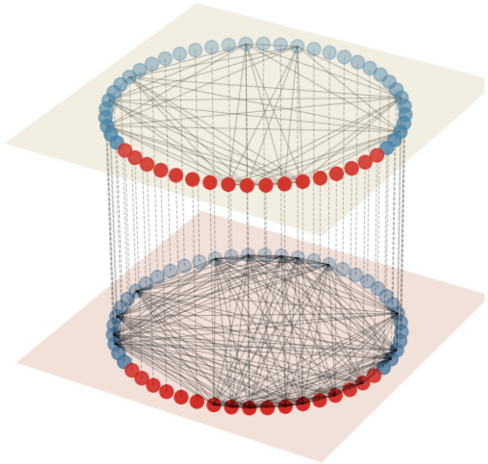

FIG. 2: Nodes connection for a specific story, "Compartir" [12], regarding Q_1 answers (top) and Q_2 answers (bottom). Color coding as Fig. 1.

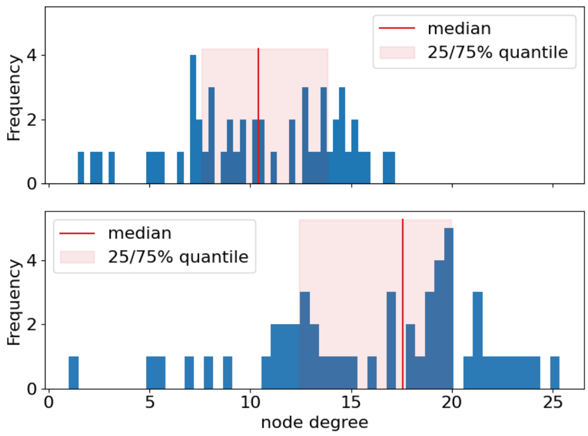

FIG. 3: Node degree distribution of G_1 answers (top) and G_2 answers (bottom), with medians and quantiles, respectively.

is a consequence of people answering more positively to Q_2 question. See Fig. 3. This fact can be interpreted as people getting in touch with different mental health situations more frequently by hearing them from their surroundings rather than experiencing them on their own. Even though individually one can feel more intensively a mental health situation, one usually is affected by one or only a few specific diseases. The micro-stories cover a large variety of mental health problems. Hence, it is reasonable to think that a more compact network is derived from Q_2, where the answers are associated to different people and the situations spectre can be larger.

Regarding age, we are interested in answering the following questions: *Do the experiences of young people in mental health differ from those of old people?* and *Are young people and old people in two different bubbles, i.e. are the mental health experiences they know from their respective surroundings different from each other?* The motivation behind these questions is to study whether the way people experience mental health situations differ between generations. The questions arise naturally since society has experienced significant changes during the last years, such as the raise of social media, and they may have an impact on people mental health. To answer both questions, we compute the assortativity coefficient described at Sect. II A 2, considering as our node attribute the age group defined at Tab. I. In this context, a higher assortativity coefficient would indicate that participants from the same age group experience similar situations (G_1) or know similar experiences from their surroundings (G_2) as other participants from the same age group, differentiating from the other group. If we can differentiate a behaviour among the age groups for the stories Q_1 and Q_2, that would indicate a positive answer to our two hypothesis respectively.

The assortativity coefficient for both G_1 and G_2 is near zero, meaning that there is non assortative mixing. Therefore, no differentiation among the defined age groups is detected for neither of the two type of stories and our answers to the suggested questions would be **no**. Nonetheless, this result should not be considered absolute, since our dataset is reduced. Besides, our dataset is biased. 70% of the participants are affected by mental health situations and the rest have been in contact with people suffering from mental health diseases. Hence, our participants consciousness about mental health may be independent of their age, being highly influenced by their personal experiences.

III. CONCLUSIONS

Using a dataset created by the means of Citizen Science, we used network analysis to study people perception of several mental health situations. Containing two different type of stories, Q_1 and Q_2, and their respective participants' answers, we created two networks derived from them, G_1 and G_2.

We presented all necessary theoretical background of networks to analyze these graphs. The main concepts and metrics were described, together with a modified assortativity coefficient inspired by [8]. We used them to analyze our graphs and conclude that the network associated to Q_2 is more compact than Q_1, meaning that it is more common to know about different mental health situations from our surroundings rather than to experience them individually. Also, we answered our hypothesis questions of whether young people mental health experiences differ from older people ones with no, since the assortativity coefficient for both graphs was near zero.

However, further research should be done to fully answer the question, since our dataset has a small size and it is not representative enough. We encourage to do a larger analysis of this situation fully exploiting the dataset by considering also other numbers of participants and stories, and furthermore extending the collection radius of the datasets, incorporating a higher ratio of participants not directly in touch with mental health situations.

Acknowledgments

I wanted to express my gratitude to my advisor Franziska Peter, for her constant help during all the project. Special thanks to my parents for always being by my side.

[1] M. Newman, *Networks: An Introduction*, (Oxford University Press, 2010)

[2] G. Bianconi, et. al., *Complex systems in the spotlight: next steps after the 2021 Nobel Prize in Physics*, Journal of Physics: Complexity **4**: 010201 (2023)

[3] D. Zinoviev, *Complex Network Analysis in Python. Recognize → Construct → Visualize → Analyze → Interpret*, (The Pragmatic Bookshelf, 2017)

[4] K. Vohland, et. al., *The Science of Citizen Science*, (Springer Cham, 2021)

[5] T. M. J. Fruchterman, E. M. Reingold, *Graph drawing by force-directed placement*, Software: Practice and Experience **21**: 1129-1164 (1991)

[6] F. Peter, I. Bonhoure & J. Perelló, *Data from: CoAct Citizen Science chatbot explores social support networks in mental health based on lived experiences*, Zenodo: **url**

[7] Universitat de Barcelona, *First citizen science chatbot opens up research on mental health*, **url**

[8] M. Newman, *Mixing patterns in networks*, Phys. Rev. E. **67**: 026126 (2003)

[9] H. Sauermann, C. Franzoni, *Crowd science user contribution patterns and their implications*, Proc. Natl. Acad. Sci. 304 **112**: 679-684 (2015)

[10] H. Spiers, et. al., *Everyone counts? Design considerations in online citizen science*, J. Sci. Commun. **18** (2019)

[11] J. P. Bagrow & Y. Ahn, *Network cards: concise, readable summaries of network data*, Applied Network Science **7**: 1-17 (2022)

[12] **Micro-story "Compartir"**: Somebody says to me, "Hello. I'm... Have you got a moment?" I was at the bar having a coffee. We'd met at the cultural centre. She'd heard I had a son with mental health problems like hers. I was afraid to open up. But gradually we shared experiences and information about resources. I got up feeling as if I'd had a good meal. (orig. language: Catalan)