USING PAGERANK AND SOCIAL NETWORK ANALYSIS TO SPECIFY MENTAL HEALTH FACTORS

Boodaghian Asl, Arsineh;
Raghothama, Jayanth;
Darwich, Adam;
Meijer, Sebastiaan

KTH Royal Institute of Technology

ABSTRACT

Various factors influence mental well-being, and span individual, social and familial levels. These factors are connected in many ways, forming a complex web of factors and providing pathways for developing programs to improve well-being and for further research. These factors can be studied individually using traditional methods and mapped together to be analyzed holistically from a complex system perspective. This study provides a novel approach using PageRank and social network analysis to understand such maps. The motives are: (1) to realize the most influential factors in such complex networks, (2) to understand factors that influence variations from different network aspects. A previously developed map for children's mental well-being was adopted to evaluate the approach. To achieve our motives, we have developed an approach using PageRank and Social Network Analysis. The results indicate that regardless of the network scale, two key factors called "Quantity and Quality of Relationships" and "Advocacy" can influence children's mental well-being significantly. Moreover, the divergence analysis reveals that one factor, "Recognition/Value Placed on well-being at School" causes a wide range of diffusion throughout the system.

Keywords: Process modelling, Complexity, Computational design methods, Social Network Analysis, Well-being

Contact:
Boodaghian Asl, Arsineh
KTH Royal Institute of Technology
Sweden
arsineh@kth.se

1 INTRODUCTION

Social Network Analysis (SNA) consists of methods to examine entities' emergent behavior and their interactions in a graph-like environment (Scott, 2011). SNA is applicable in different fields, such as information technology, politics, and health models (Cavallo et al., 2012). Models are a way to represent one or some part of the world to make it easier to understand how different parts function. There are various forms of models, such as physical, mathematical, graphical, etc. (Sturmberg and Martin, 2013). Healthcare systems are large, complex, grow irregularly, and are always evolving, making it challenging to analyze efficiently through mathematical or physical models. In healthcare, SNA is used to perceive network diffusion during an epidemic, understanding how the disease spreads in the body, and providing information about social behavior such as a source of causation. Mental health is a sector in healthcare triggered by various factors, which is complex to understand and analyze. To facilitate the understanding of complex mental models, the paper proposes an approach using PageRank and SNA.

Complex systems can be represented by graphical models consisting of nodes, where each node should have a link to at least one other node in the system. Links are either directional or bidirectional, determining the orientation of impact flow (Cavallo et al., 2012). Graphs are a way to represent different components of a network and their relations. Sometimes it is necessary to examine part of a graph in isolation; hence the isolated part of the graph is a sub-graph, consisting only of some components and connections from the network. As an example, a graph that consists of mental health factors can have multiple sub-graphs embracing only family or education related factors (Figure 2 and 3).

Many fields related to mental health deal with complex problems for precise management, decision-making, training, organizing, and understanding how their systems function (Cavallo et al., 2012). In fields such as sociology or psychology, one objective could be to realize how personality and behavior alter and what causes this alteration, given different circumstances such as environmental change, trauma, work pressure, and other incidents (Benham-Hutchins and Clancy, 2010). Each of these events can appear in different models. As an illustration, children's stress can be triggered at school, such as bullying or study difficulty, or home, such as abuse or family income. To examine the source of causation, a graph-like map is required consisting of both systems, family and home, with all these factors. This makes the model more complex to analyze efficiently due to growing number of factors and their interconnected links using methods such as Structural Equation Modelling (SEM) and Mediation Analysis (MA) (Letina et al., 2019).

Apart from the above example, other factors also influence children's mental well-being depending on mental conditions and environmental impact (Benham-Hutchins and Clancy, 2010). Some of these factors, given sufficient information, can be easily identified by professionals. Nonetheless, there can be underlying factors with higher impact. Each of these factors with specific strengths can have a harmful or beneficial impact on each other. As the number of factors grow, the connections among them increase, and thus, their influence varies, making it more complex to investigate the whole system.

SNA comprises methods to detect centrality according to the node position and the number of links associated with the node. Moreover, it can evaluate network density, rank nodes in a network, and return information regarding node and link properties (Oliveira and Gama, 2012). Centrality aims to detect the most influential nodes in a system through approaches that firmly rely on degrees, betweenness, closeness, accessibility of the nodes in a graph (Rodrigues, 2019), though PageRank aims to rank nodes. We propose an approach which has the capacity to provide policymakers and mental health professionals an in-depth and systemic view into mental health factors to investigate and develop programs. The purpose is to improve well-being by quantifying and breaking down the map into several sub-graphs using PageRank and SNA and analyzing them from different aspects; to evaluate the stability of most influential factors, regardless of the graph size. The approach has the potential to assist researchers in fields such as sociology and psychology to formulate and analyze their research questions or validate their hypotheses. Furthermore, the approach may accelerate the engineering design process by revealing the leverage points which demand attention. A key aspect for this analysis is to place, inspect and analyze mental health and well-being within a societal context, and not only from a clinical or care perspective.

The remainder of this paper discusses the background of SNA methods previously used to quantify and analyze complex health maps, and the purpose of adopting PageRank. Later, in the methodology section we described the newly proposed approach, and assess the approach in the results section. To that end, we specify the application and obstacles of the approach in discussion and conclusion sections.

2   BACKGROUND

The use of network theory in social structures dates back to 1930. Primarily, SNA was used only in the field of mathematics to study graph theory. Scott (2011) reviewed and provided insight into adopting SNA approach for both sociology and other areas. He argues that physicists' main concerns were the network's dynamics, where this network grew and changed consistently (Watts, 1999). Most of the focus in this area been static analysis of the network. In recent years the focus was also on network's dynamics; since a change, such as modifying link information, impacts the whole network behavior. Therefore, the focus on the dynamic analysis of the network is to understand the pattern variation (Scott, 2011).

Various researchers adopted SNA and proposed a new approach to analyze complex health systems. Zhao et al. (2016) concern was to analyze the user interactions in the online health community. They proposed a multi-relational SNA method using the centrality approach for smoking cessation to investigate the sub-graphs' connection. Their study created four different sub-graphs based on the data gathered from online health community users. Their approach compared the sub-graphs links and nodes via the node centrality. Moreover, the paper also analyzes each sub-graph evolving pattern. The research also concludes that centrality does not indicate smoking outcomes in sub-graphs (Zhao et al., 2016).

A similar study was performed by Schoen et al. (2014) using density, degree, and centrality approaches from SNA. The study aimed to measure the partnership to implement strategies in public health. The paper applied a different approach of centrality and density and did a comparative analysis for the networks. The paper claims that these approaches are suitable for analyzing the relationship, communication, and collaboration; though, the study acknowledged being unable to find the reason for network result variations (Schoen et al., 2014).

The use of SNA to study for complex systems was mostly focused on adopting the centrality methods such as betweenness, degree, closeness (Oliveira and Gama, 2012) as outlined above. In a research study by Bringmann et al. (2019), the researchers investigate the method's issues when applying for psychological networks analysis. The study mentioned different concerns of using centrality, which is: (1) primarily developed for the social network rather than for psychological networks, (2) unstable for cross-sectional and temporal networks and displays wide confidence interval and have inconsistency, (3) does not reveal the main symptom, (4) unclear of the methods underlying goal, (5) the methods considers only positive weights, which makes it challenging to apply it for psychological and sociological networks consisting of negative links (Bringmann et al., 2019).

Few studies on psychological networks have used other SNA methods than centrality. Letina et al. (2019) encourage using other SNA methods, which can extract more useful information than centrality alone provides. The paper used three different methods: Minimum Spanning Tree (MST), participation coefficient, and motif analysis. Nevertheless, the methods were used for a more precise measure of centrality (Letina et al., 2019). However, MST helps grouping nodes based on minimum sets of links in a graph, which increases the possibility of neglecting an important factor in health-related networks.

There are other methods of analysis such as SEM and MA. SEM is a framework used for factors and path analysis in complex systems, mainly by social and behavioral researchers (Hox and Bechger, 1999). On the other hand, MA is to understand why factors influence each other, commonly used by the SEM framework, which facilitates the performance (Gunzler et al., 2013). In both cases, data is required to quantify and perform analysis. To facilitate the analysis and quantification of large-scale systems with numerous factors, this paper proposes an approach using PageRank and SNA, quantifying and analyzing the factors. The approach provides information regarding the factors and their strength from different systems perspectives. Contrary to centrality, PageRank is capable of considering link information in ranking factors by disregarding the location of the factors. Moreover, the approach aims to realize whether the influential factors frequently emerge in different sub-graphs.

Most studies which adopted the SNA approach uses a qualitative method or a combination of qualitative and quantitative methods, using questionnaires to quantify individuals and groups in a network such as Zhao et al. (2016); Schoen et al. (2014). Edwards (2010) elaborated qualitative and quantitative methods and the mixed approach towards social sciences. In a research by Cheong et al. (2013) focused on multidisciplinary systems and the impact of patient health connection. Their research aimed to study the relations between individuals or groups in a network. Thus, researchers used a mixed method of an analytical and theoretical framework by mixing qualitative and quantitative methods in the SNA approach. The paper argued that the quantitative method helps map and measure the network property,

while the qualitative method helps analyze and better understand the network structure and the interaction between individuals and groups (Cheong et al., 2013). A few researchers used different approaches to quantify and measure the importance of nodes in complex networks, such as Piraveenan et al. (2013) proposed percolation centrality, which measures node influence according to topological connectivity (Piraveenan et al., 2013).

Hence, the approach proposed in this paper adopts the PageRank method to quantify graph nodes, which is adjustable by handling both positive and negative link values. Furthermore, to form the sub-graphs of different nodes, all the possible connecting paths are considered to avoid missing crucial health-related information. Thus, the approach facilitates finding the most influential factors in each sub-graph, and can provide sufficient information to sort and compare these factors in various sub-graphs of the map. Furthermore, in the case of dynamic analysis of the map, PageRank can continuously adopt new information about the link through the simulation and display the impact on all network nodes.

3 METHODOLOGY

3.1 The map

To develop and evaluate the proposed approach, we adopted a map (Figure 2 and 3) previously developed for children's and young people's mental well-being (Raghothama et al., a,b). Each node represents a factor that directly or indirectly may affect children's mental well-being. All the links have directions with positive and negative signs, determining how the factors affect each other. The nodes are categorized into core, education, skills, relationships, family, social, and work.

3.2 PageRank

PageRank is an algorithm primarily developed to rank the web pages in google search engine and display the search results according to the highest rank (Page et al., 1999). The purpose is to rank each page in a network according to its incoming and outgoing links. With the PageRank algorithm's advancement, other parameters such as damping-factor, link weight, error tolerance, number of iterations, and dangling nodes impact the ranking value. In a network consisting of web pages, links are all considered positive; however, complex health maps containing negative link values may cause a rank-sink. Other factors may cause a rank-sink, such as nodes with no outgoing links. Therefore, to avoid rank-sink caused by various reasons, it is necessary to adjust parameters such as damping-factor (Bressan and Peserico, 2010) and dangling nodes accurately. To validate PageRank accuracy, all the available ranks in each node list should add up to (1.0); otherwise, the node rank will be unreliable. Furthermore, the PageRank range should vary in the range of [0,1], with the involvement of negative weight, PageRank range [-1,1]. PageRank facilitates nodes quantification, where, in the case of SEM and MA, data gathering and sample size are vital. PageRank also has challenges, though highly dependent on the usage context. PageRank gives preference to incoming links information. From the search engine perspective, web page users can easily mislead the PageRank function to rank the untrustworthy websites higher. From the healthcare perspective, missing link or inaccurate link information can provide diversity in the outcome.

3.3 Procedure

In order to analyze the map, the following input variables are required: (1) list of available nodes, (2) list of available links, and (3) list of coefficients corresponding to each available link. The approach analyzes the map from two different perspectives: (1) global and (2) local. From the Global perspective, the algorithm considers the impact of all the external links and nodes linked to each sub-graph; however, the algorithm isolates the sub-graph from the external links and nodes from the local perspective. These two perspectives aim to analyze rank variation if the focus is only on some system segments. Eventually, Divergence is measured to evaluate if the rank fluctuates uniformly.

Figure 1 is the flow diagram developed to analyze the map. Followings are the approach procedure:
1. Construct a directed graph (*DiGraph*).
2. Initialize the directed graph by assigning nodes and corresponding edges *(N,E)*.
3. If link exists, assign correlation coefficient as link's weight.
4. Calculate the global PageRank (*PR*) by assigning the directed graph and dumping factor (α).
5. Iterating through the node list to begin a path (*U*) and to end a path (*V*).
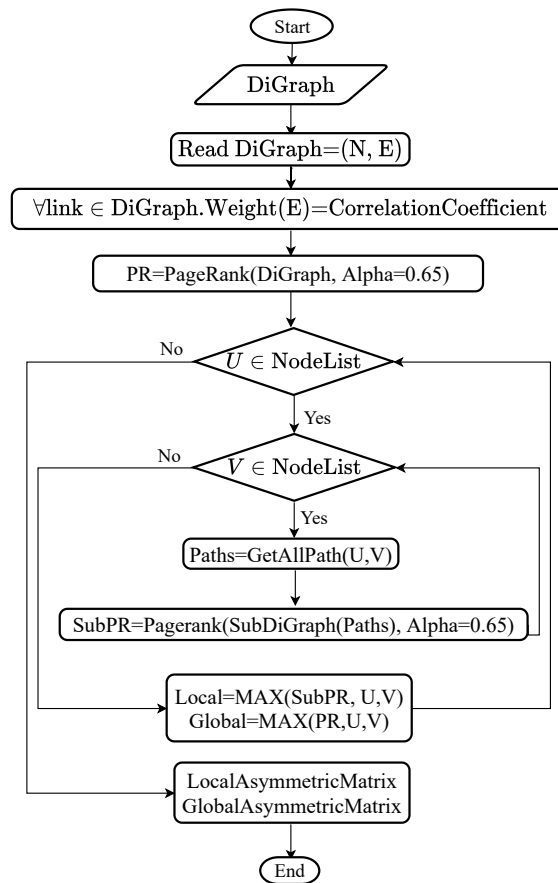
Figure 1. Flow diagram

6. Collect all the path in a single variable (*Paths*).
7. Calculate the local PageRank (*SubPR*) by assigning all the collected paths between two nodes and the correct dumping factor.
8. Find the maximum local and global PageRank and Assign it to corresponding asymmetric matrix.

The diagram first generates an empty directed graph where the nodes' impact-flow follows the links' directions. The graph can have both directional and bidirectional links. If the linked nodes have mutual impacts, then the link is bidirectional; otherwise, the link is directional. The graph initializes by importing nodes (*N*), links (*E*), and their properties from map. Each link has a coefficient, which indicates the strength of the connection. Link coefficients are assigned if two nodes are connected. Coefficients can be negative or positive according to the correlation type. Here, all the coefficients are set to (± 1.0). After graph initialization, PageRank (*PR*) is calculated for the whole graph by lowering the damping-factor *Alpha* to 0.65 (Bressan and Peserico, 2010). This adjustment is to avoid rank-sink due to the negative links. All the nodes in graph had outgoing links; thus, the rank-sink was not a result of dangling nodes. To reveal the most influential nodes in each sub-graph, the algorithm's conditional blocks iterate through all the available paths between each pair of nodes, with respect to link directions. *NodeList* is the list of all the available nodes in the graph, with their properties such as outgoing and incoming links and their coefficients. *U* annotation indicates the starting node of each path, and *V* annotation indicates the ending node of each path. Each path may consist of two to several nodes, where nodes do not repeat. Each pair of nodes can have one to several connecting paths according to link direction. The collection of paths between 2 nodes become a sub-graph. Then the PageRank (*SubPR*) is calculated for the sub-graph by isolating it from the connecting nodes and links. After finding the maximum ranks in each sub-graph, asymmetric matrices for both global and local ranks are generated for further study.

4 RESULTS

Outcomes are presented in two sections: (1) Highlight the highest ranking nodes in different sub-graphs using tables and map visualization, and (2) Display rank divergence between global to local paths by visualizing PageRank asymmetric matrix pattern fluctuation.

4.1 Most influential nodes

We used the data from the asymmetric matrices which was generated by examining the map. The complete map consists of 65 nodes and 168 links. Table 1 and 2 presents the most influential nodes and the corresponding ranks in some selective sub-graphs from a global and local perspectives. The top row indicates the nodes the path begins, and the first column indicates the nodes the path ends. Sub-graphs are selected purposely to display the most variation of the outcome. Empty cells imply no path between the two nodes. Each sub-graph has one influential node. If the sub-graph consists of 2 nodes, the local PageRank will be similar for both nodes (PR=0.5) due to sharing a common link and coefficient.

Followings are the list of selective nodes and their ranks from both local and global perspective:

1. Frequently most influential nodes (Global): *Quality and quantity of friendships*, *Advocacy*, and *Stress*.
2. Frequently most influential nodes (Local): *Quality and quantity of friendships*, *Advocacy*, *Value of wellbeing recognized in school culture*, and *Self-efficacy*.
3. Categories with frequently most influential node (Global): *Relationships* and *Core*.
4. Categories with frequently most influential node (Local): *Relationships*, *Core*, and *Education*.

Some nodes seldom emerge as the most influential node, such as *Support available for wellbeing within local community* (PR=0.0128) from *Education* category or *Use of illegal drugs and alcohol* (SubPR=0.3837) from *Family* category. This implies that nodes which frequently emerge are not listed in some paths due to their link direction, or the emerged nodes' rank is relatively higher. Figure 2 and 3 highlights the most influential nodes for all sub-graphs. In both figures *Quality and quantity of friendships* (PR=0.761, SubPR=0.08~0.11) from *Relationships* category, and *Advocacy* (PR=0.601, SubPR=0.07~0.13) from *Core* category frequently emerges as the most influential nodes (larger nodes). Even though nodes emergence vary in Figure 2 and 3, both display approximately the same impact diffusion over the network. For instance, *Work* category has no impact on the network. *Education* category delivers similar influential nodes from both local and global perspectives. However, *Relationships* and *Family* categories have denser amount of most influential nodes from local perspective. Hence, *Relationships* and *Core* categories frequently influenced other nodes in network. Figures comparison imply that some nodes continue to have the highest influencing power regardless of how the map evolves.

Table 1. Most influential nodes global path

Ends \ Begins	AESE	EL	ERB	FES	FWCDTT	LN	NWHE	PE	RRWSES	SA	STDD
ALDEY	QQR (0.0761)	QQR (0.0761)	QQR (0.0761)	QQR (0.0761)	QQR (0.0761)	ALDEY (0.0079)	QQR (0.0761)	QQR (0.0761)	QQR (0.0761)	QQR (0.0761)	ER (0.0116)
ACE	AESE (0.0257)	Ss (0.0658)	Ay (0.0601)	QQR (0.0761)	QQR (0.0761)	QQR (0.0761)	NWHE (0.0481)	QQR (0.0761)	QQR (0.0761)	NWHE (0.0481)	QQR (0.0761)
HC	N/A	N/A	N/A	SFS (0.0096)	N/A	N/A	N/A	N/A	SAWWLC (0.0128)	N/A	N/A
LN	QQR (0.0761)	QQR (0.0761)	QQR (0.0761)	QQR (0.0761)	QQR (0.0761)	N/A	QQR (0.0761)	QQR (0.0761)	QQR (0.0761)	QQR (0.0761)	ER (0.0116)
PES	QQR (0.0761)	Ss (0.0658)	Ay (0.0601)	QQR (0.0761)	QQR (0.0761)	QQR (0.0761)	PES (0.0547)	QQR (0.0761)	QQR (0.0761)	PES (0.0547)	QQR (0.0761)
PH	QQR (0.0761)	QQR (0.0761)	QQR (0.0761)	QQR (0.0761)	QQR (0.0761)	QQR (0.0761)	QQR (0.0761)	PH (0.0441)	QQR (0.0761)	QQR (0.0761)	QQR (0.0761)
RVPWS	QQR (0.0761)	Ss (0.0658)	Ay (0.0601)	QQR (0.0761)	RVPWS (0.0475)	QQR (0.0761)	VWRSC (0.0566)	QQR (0.0761)	QQR (0.0761)	VWRSC (0.0566)	QQR (0.0761)
RSA	AESE (0.0257)	Ss (0.0658)	Ay (0.0601)	QQR (0.0761)	QQR (0.0761)	QQR (0.0761)	NWHE (0.0481)	QQR (0.0761)	QQR (0.0761)	NWHE (0.0481)	QQR (0.0761)
TSWF	N/A	N/A	N/A	TSWF (0.0076)	N/A	N/A	N/A	N/A	N/A	N/A	N/A
UIDA	QQR (0.0761)	QQR (0.0761)	UIDA (0.0069)	QQR (0.0761)	QQR (0.0761)	QQR (0.0761)	QQR (0.0761)	QQR (0.0761)	QQR (0.0761)	QQR (0.0761)	QQR (0.0761)
Vy	QQR (0.0761)	Ss (0.0658)	Vy (0.0099)	QQR (0.0761)	QQR (0.0761)	QQR (0.0761)	QQR (0.0761)	QQR (0.0761)	QQR (0.0761)	QQR (0.0761)	QQR (0.0761)

Acronyms: AESE=Ability to Ensure Safe Environment, ALDEY=Access to Language Development in Early Years, ACE=Adverse Childhood Experience, Ay=Advocacy, EL=Emotional Literacy, ER=Emotional Regulation, ERB=Engaging in Risky Behaviors, FES=Family Economic Stability, FWCDTT=Focus on Wellbeing and Childhood Development in Teacher Training, HC=Housing Conditions, LN=Learning Need, NWHE=Nurture Within Home Environment, PES=Parental Engagement in School, PE=Physical Exercise, PH=Physical Health, QQR=Quality and Qantity of Relationships, RRWSES=recognition and reward for wellbeing support from education system, RVPWS=recognition/value placed on wellbeing at school, RSA=required sleep achieved, SA=Secure Attachment, SE=Self-Efficacy, SFS=Stability of Family Structures, Ss=Stress, STDD=Some Types of Development Disorder, SAWWLC=support available for wellbeing within local community, TSWF=Time to Spend With Family, UIDA=Use of Illegal Drugs and Alcohol, VWRSC=value of wellbeing recognised in school culture, Vy=Vulnerability

Table 2. Most influential nodes local path

Begins / Ends	AESE	EL	ERB	FES	FWCDTT	LN	NWHE	PE	RRWSES	SA	STDD
ALDEY	QQR (0.0920)	QQR (0.0933)	QQR (0.0933)	QQR (0.0863)	QQR (0.0896)	ALDEY (0.6226)	QQR (0.0920)	QQR (0.0949)	QQR (0.0849)	QQR (0.0920)	ALDEY (0.3320)
ACE	ACE (0.6226)	SE (0.0906)	Ay (0.1284)	QQR (0.0880)	QQR (0.0929)	QQR (0.1167)	ACE (0.4389)	Ay (0.0724)	QQR (0.0866)	NWHE (0.2500)	QQR (0.1146)
HC	N/A	N/A	N/A	HC (0.4389)	N/A	N/A	N/A	N/A	HC (0.3320)	N/A	N/A
LN	QQR (0.0960)	QQR (0.0973)	QQR (0.0973)	QQR (0.0899)	QQR (0.0935)	N/A	QQR (0.0960)	QQR (0.0990)	QQR (0.0885)	QQR (0.0960)	LN (0.4389)
PES	QQR (0.1089)	SE (0.0934)	Ay (0.1370)	QQR (0.1015)	QQR (0.0919)	QQR (0.1194)	PES NWHE (0.5000)	Ay (0.0743)	QQR (0.0842)	NWHE (0.4646)	QQR (0.1172)
PH	QQR (0.0920)	QQR (0.0933)	QQR (0.0933)	QQR (0.0863)	QQR (0.0896)	QQR (0.0933)	QQR (0.0920)	PH (0.6226)	QQR (0.0849)	QQR (0.0920)	QQR (0.0919)
RVPWS	VWRSC (0.2172)	VWRSC (0.5915)	VWRSC (0.1169)	VWRSC (0.2136)	RVPWS (0.6226)	VWRSC (0.3489)	VWRSC (0.3113)	VWRSC (0.4247)	VWRSC (0.2363)	PES (0.2942)	VWRSC (0.3462)
RSA	RSA (0.6226)	SE (0.0889)	Ay (0.1251)	QQR (0.0865)	QQR (0.0916)	QQR (0.1151)	RSA (0.4389)	Ay (0.0713)	QQR (0.0851)	RSA (0.3320)	QQR (0.1130)
TSWF	N/A	N/A	N/A	TSWF (0.6226)	N/A		N/A	N/A	N/A	N/A	N/A
UIDA	QQR (0.0920)	QQR (0.0933)	ERB UIDA (0.5000)	QQR (0.0863)	QQR (0.0896)	QQR (0.0933)	QQR (0.0920)	QQR (0.0919)	QQR (0.0849)	QQR (0.0920)	QQR (0.0962)
Vy	QQR (0.0920)	VWRSC (0.0855)	UIDA (0.3837)	QQR (0.0863)	QQR (0.0896)	QQR (0.0933)	QQR (0.0920)	VWRSC (0.0770)	QQR (0.0849)	QQR (0.0920)	QQR (0.0919)

Acronyms: AESE=Ability to Ensure Safe Environment, ALDEY=Access to Language Development in Early Years, ACE=Adverse Childhood Experience, Ay=Adovcacy, EL=Emotional Literacy, ERB=Engaging in Risky Behaviors, FES=Family Economic Stability, FWCDTT=Focus on Wellbeing and Childhood Development in Teacher Training, HC=Housing Conditions, LN=Learning Need, NWHE=Nurture Within Home Environment, PES=Parental Engagement in School, PE=Physical Exercise, PH=Physical Health, QQR=Quality and Qantity of Relationships, RRWSES=recognition and reward for wellbeing support from education system, RVPWS=recognition/value placed on wellbeing at school, RSA=required sleep achieved, SA=Secure Attachment, SE=Self-Efficacy, Ss=Stress, STDD=Some Types of Development Disorder, TSWF=Time to Spend With Family, UIDA=Use of Illegal Drugs and Alcohol, VWRSC=value of wellbeing recognised in school culture, Vy=Vulnerability
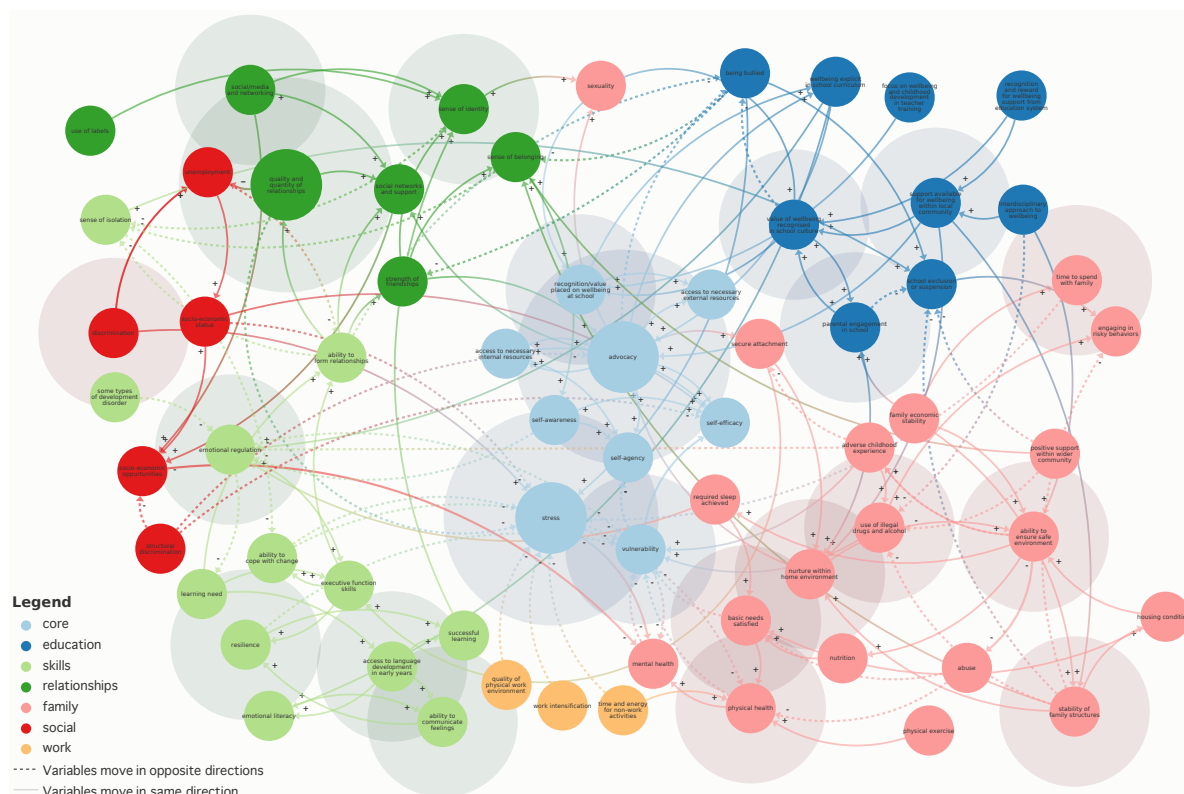

Figure 2. Children and young people well-being map with the global most influential nodes
(Raghothama et al., a,b)

4.2 Rank divergence

To evaluate the approach's reliability, global and local sub-graph nodes rank variation should follow approximately in a uniform pattern. Complex mental maps can evolve irregularly, therefore it is essential
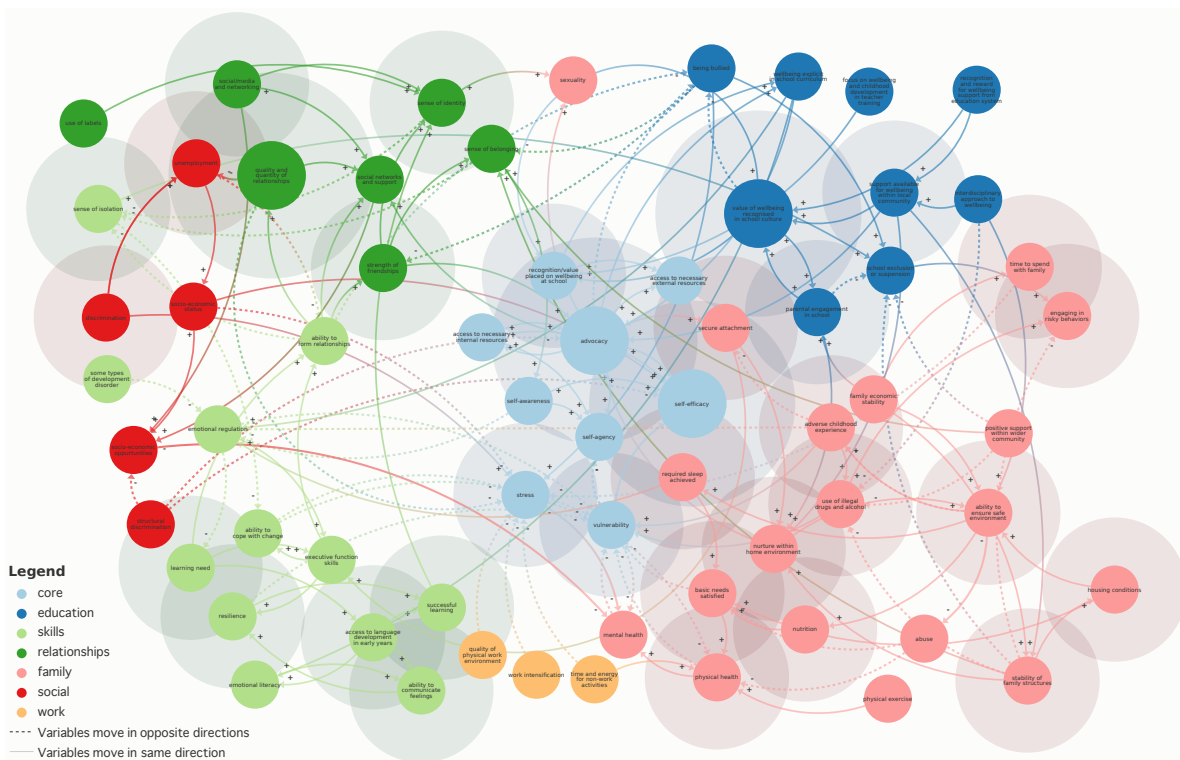
Figure 3. Children and young people well-being map with the local most influential nodes
(Raghothama et al., a,b)

to detect pattern fluctuations from two different perspectives. For this reason, the divergence of each sub-graph is measured to detect unexpected fluctuation.

Figure 4 displays how the PageRank varies from Global to Local sub-graph. The *x-axis* represents the nodes where the path begins, and the *y-axis* represents where the path ends, alphabetically. The divergence result indicates that rank variation lies in a steady range; however, for the sub-graphs consisting of the node *recognition-value placed on wellbeing at school* (PR=0.0475, SubPR=0.6226), the ranks fluctuation is higher than the uniform range and reach above the (0.3). The result implies that each sub-graph containing this node will have a high impact on other nodes. The fluctuation can be either due to a link coefficient, link type, or graph structure, and might be sensitive to rank-sink.

5 DISCUSSION AND CONCLUSION

In this paper we propose a new approach using PageRank and SNA to quantify and investigate health-related maps in more detail than centrality could provide by iterating over all the possible sub-graphs from both global and local perspectives. The approach provides information regarding factors that relatively have a higher impact on children's mental health. The result indicates that the approach has a beneficial impact from different field perspectives within health.

To policymakers and mental health professionals, detecting the influential nodes could potentially assist in investigating its impact from different aspects. One way is to focus on nodes' rank steady behavior in different sub-graphs; another way can be to ensure the node remains as the most influential node, regardless of how the network evolves. Another application of the approach is to focus on various influential nodes emerging in similar sub-graphs. As illustrated in Table 1 or 2, in each column there can be 3 or more influential nodes. If the variation is high, then the next target can be on the nodes' ranks. Additionally, the approach aids in detecting unexpected fluctuations of the patterns by comparing the asymmetric matrices from different maps (Figure 4). An example could be to measure the divergence of the most influential nodes between two maps but designed for the same purpose to observe the causes.

To researchers in fields such as sociology and psychology, one potential could be to use the approach for validating research hypotheses and formulate research questions, such as understanding the range of impact stress has on family. The proposed approach is capable of analyzing small to large-scale maps by breaking down the map into sub-graphs with different coefficients. This allows researchers to detect
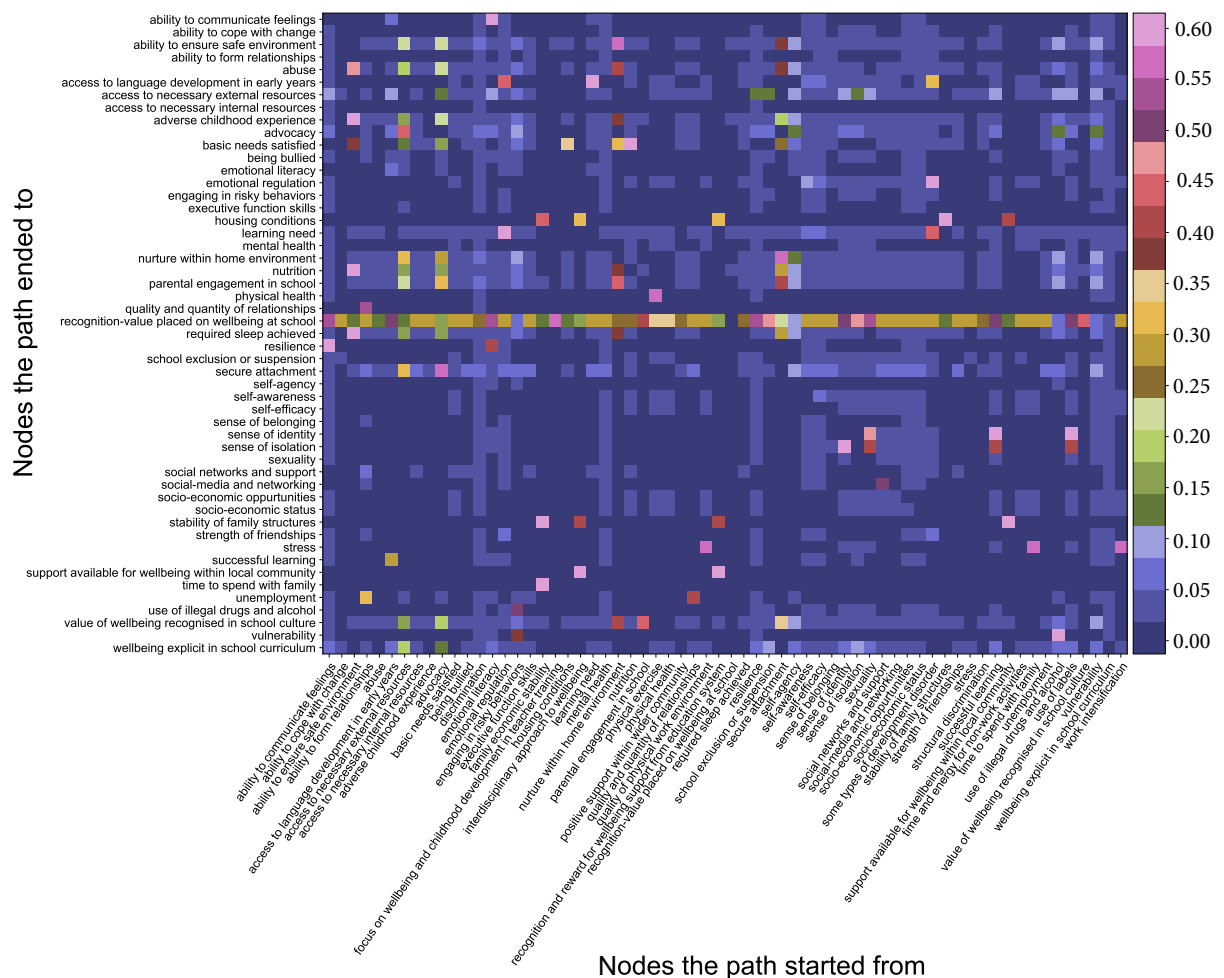
Figure 4. Divergence of PageRank between global and local sub-graphs

contradictions in their assumptions and inspect how the coefficients impact the results by observing if stress continuously has the highest rank in different sub-graphs such as local and global perspectives.

One of the obstacles of the approach is the PageRank sensitivity to the given input data, resulting in outcome variation. This might be caused by an incorrect coefficient, link types, or PageRank parameters; as Bressan and Peserico (2010) suggested, for applications other than a search engine, it is best to use damping-factor close to (0.5); however, Boldi et al. (2005) recommended to focus on adjusting the number of iteration instead of altering the damping-factor (Boldi et al., 2005). To further improve this approach, it is necessary to simulate the map with different coefficients and inspect the patterns among its sub-graphs. A simulation will provide a better insight into the source of causation.

There were alternative methods to analyze networks and social sciences. In a study by Vicsek et al. (2016), researchers compared two approaches, Actor-Network Theory (ANT) and SNA. ANT aims to investigate scientific and technological artefacts and if artefacts are separately functional. The paper suggests that ANT can be more effective if it adopts the SNA methods such as visualization and measurement, while, SNA should advance the approach from both theoretical and empirical perspectives to involve both non-human and human entities into the network (Vicsek et al., 2016). However, SNA is more reliable for large-scale networks than methods such as SEM and MA (Letina et al., 2019).

In general, the approach has the potential to be used for design processes and policies to identify leverage points due to the flexibility of PageRank function. The approach can be with the combination of simulation methods such as agent-based and system dynamic simulation to analyze complex systems from micro and macro level perspectives, compare different scenarios and observe the influence size.

REFERENCES

Benham-Hutchins, M. and Clancy, T.R. (2010), "Social networks as embedded complex adaptive systems", *Journal of Nursing Administration*, Vol. 40 No. 9, pp. 352–356. Doi: 10.1097/NNA.0b013e3181ee42bc.

Boldi, P., Santini, M. and Vigna, S. (2005), "Pagerank as a function of the damping factor", in: *Proceedings of the 14th International Conference on World Wide Web*, pp. 557–566. Doi: 10.1145/1060745.1060827.

Bressan, M. and Peserico, E. (2010), "Choose the damping , choose the ranking?", *Journal of Discrete Algorithms*, Vol. 8 No. 2, pp. 199–213. Doi: 10.1007/978-3-540-95995-3_7.

Bringmann, L.F., Elmer, T., Epskamp, S., Krause, R.W., Schoch, D., Wichers, M., Wigman, J.T. and Snippe, E. (2019), "What do centrality measures measure in psychological networks?", *Journal of Abnormal Psychology*. Doi: 10.1037/abn0000446.

Cavallo, P., Pagano, S. and Carpinelli, L. (2012), "The drug prescription process: A network medicine approach", in: *Handbook of Systems and Complexity in Health*, Springer New York, pp. 537–562. Doi =10.1007/978-1-4614-4998-0_32.

Cheong, L.H., Armour, C.L. and Bosnic-Anticevich, S.Z. (2013), "Primary health care teams and the patient perspective: A social network analysis", *Research in Social and Administrative Pharmacy*, Vol. 9 No. 6, pp. 741–757. Doi: 10.1016/j.sapharm.2012.12.003.

Edwards, G. (2010), "Mixed-method approaches to social network analysis", *ESRC national Centre for Research Methods*, pp. 1–30.

Gunzler, D., Chen, T., Wu, P. and Zhang, H. (2013), "Introduction to mediation analysis with structural equation modeling", *Shanghai archives of psychiatry*, Vol. 25, pp. 390–394. Doi: 10.3969/j.issn.1002-0829.2013.06.009.

Hox, J. and Bechger, T. (1999), "An introduction to structural equation modeling", *Family Science Review*, Vol. 11.

Letina, S., Blanken, T.F., Deserno, M.K. and Borsboom, D. (2019), "Expanding network analysis tools in psychological networks: Minimal spanning trees, participation coefficients, and motif analysis applied to a network of 26 psychological attributes", *Complexity*. Doi: 10.1155/2019/9424605.

Oliveira, M. and Gama, J. (2012), "An overview of social network analysis", *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, Vol. 2 No. 2, pp. 99–115. Doi: 10.1002/widm.1048.

Page, L., Brin, S., Motwani, R. and Winograd, T. (1999), *The PageRank Citation Ranking: Bringing Order to the Web*, Technical report, Standford InfoLab.

Piraveenan, M., Prokopenko, M. and Hossain, L. (2013), "Percolation centrality: Quantifying graph-theoretic impact of nodes during percolation in networks", *PLoS ONE*, Vol. 8 No. 1. Doi: 10.1371/journal.pone.0053095.

Raghothama, J., Darwich, A. and Meijer, S. (a), "Mapping well-being in children and young people - a participatory systems mapping approach", In Progress.

Raghothama, J., Darwich, A. and Meijer, S. (b), "Placing mental health and well-being in context through participatory mapping", In Progress.

Rodrigues, F.A. (2019), "Network centrality: An introduction", in: *A Mathematical Modeling Approach from Nonlinear Dynamics to Complex Systems*, Springer International Publishing, pp. 177–196. Doi: 10.1007/978-3-319-78512-7_10.

Schoen, M.W., Moreland-Russell, S., Prewitt, K. and Carothers, B.J. (2014), "Social network analysis of public health programs to measure partnership", *Social Science and Medicine*, Vol. 123, pp. 90–95. Doi: 10.1016/j.socscimed.2014.10.057.

Scott, J. (2011), "Social network analysis: Developments, advances, and prospects", *Social Network Analysis and Mining*, Vol. 1 No. 1, pp. 21–26. Doi: 10.1007/s13278-010-0012-6.

Sturmberg, J.P. and Martin, C.M. (2013), "Modelling", in: *Handbook of Systems and Complexity in Health*, Springer New York, pp. 81–111.

Vicsek, L., Király, G. and Kónya, H. (2016), "Networks in the social sciences: Comparing actor-network theory and social network analysis", *Corvinus Journal of Sociology and Social Policy*, Vol. 7 No. 2, pp. 77–102. Doi: 10.14267/CJSSP.2016.02.04.

Watts, D.J. (1999), "Networks , dynamics , and the small-world phenomenon", *American Journal of Sociology*, Vol. 105 No. 2, pp. 493–527. Doi: 10.1086/210318.

Zhao, K., Wang, X., Cha, S., Cohn, A.M., Papandonatos, G.D., Amato, M.S., Pearson, J.L. and Graham, A.L. (2016), "A multirelational social network analysis of an online health community for smoking cessation", *Journal of Medical Internet Research*, Vol. 18 No. 8. Doi: 10.2196/jmir.5985.