Master's Thesis

A Network Analysis on Mental Health Symptoms to Identify Possible Intervention Points in a University Environment

Sevde Ucak

Department of Biomedical Engineering
(Human Factors Engineering)

Ulsan National Institute of Science and Technology

2021

A Network Analysis on Mental Health Symptoms to Identify Possible Intervention Points in a University Environment

Sevde Ucak

Department of Biomedical Engineering
(Human Factors Engineering)

Ulsan National Institute of Science and Technology

A Network Analysis on Mental Health Symptoms to Identify Possible Intervention Points in a University Environment

A thesis/dissertation submitted to
Ulsan National Institute of Science and Technology
in partial fulfillment of the
requirements for the degree of
Master of Science

Sevde Ucak

07.13.2021 of submission

Approved by

Advisor
Dooyoung Jung

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

A Network Analysis on Mental Health Symptoms to Identify Possible Intervention Points in a University Environment

Sevde Ucak

This certifies that the thesis/dissertation of Sevde Ucak is approved.

07.13.2021 of submission

Signature

Advisor: Dooyoung Jung

Signature

권오상

Oh-Sang Kwon

Signature

Dongil Chung

Dongil Chung

Abstract

This thesis unravels the symptomatic structure behind the two commonly observed mental health disorders namely depression and anxiety. Additionally, structural effects of worry and meta-worry mechanisms on these disorders are explored by estimating their network models. From mental health data collected over the years, best-fitting models are created, analyzed, and interpreted. This is the first study to deep dive into the symptomatic relations within depression and anxiety disorders and to estimate the intricate relationship of the depression-worry & meta-worry and anxiety-worry & meta-worry mechanisms with psychological networks. Since learning of networks is a high-dimensional statistical problem, I employ regularization to learn sparse networks and thereby control over-fitting. In particular, I used the popular graphical Lasso model which is an implementation of the undirected Gaussian Graphical Model (GGM). Centrality and edge weight stabilities are computed for all the generated networks prior to interpretation. Furthermore, I used Bayesian networks (directed acyclic graphs–DAGs) using the hill-climbing algorithm provided by the R package bnlearn including the correlation structures in order to reveal terminal nodes as crucial intervention and prevention targets. As a result, I revealed central symptoms and symptom initiators for depression and anxiety by means of undirected regularized partial correlation network and directed acyclic networks respectively. Comorbid symptoms are investigated through bridge network analysis and the causality of each bridge item is further explored through directed networks. For depression, "Self-dislike", "Loss of energy", "Worthlessness" and "Tiredness or fatigue" symptoms emerged as the most central symptoms. Further analysis by directed acyclic networks showed "Self-dislike", "Worthlessness" along with "Past failure" as the initiator of remaining symptoms. For anxiety, "Shaky / unsteady", "Hearth pounding / racing", "Nervous" and "Scared" appear as the strongest nodes in the undirected network. "Nervous" and "Scared" together with "Fear of worst happening" comes up as triggering symptoms of anxiety in the directed network. "Need to control thoughts" subscale only showed comorbidity with depression symptoms. While "Lack of cognitive confidence", "Negative beliefs about ncontrollability and danger" sub-scales of the meta-cognition questionnaire, and "My worries overwhelm me" item from the PSWQ(worry) showed comorbidity with both depression and anxiety symptoms. The causality of each bridge item is explored through a directed network further.

Contents

List of Figures

I Introduction

Recent years have seen increasing use of network modeling for exploratory studies of psychological behavior as an alternative to latent variable modeling [1, 2]. In simple terms, Network is a set of nodes connected by set of edges. In social networks nodes can represent entities like people and cities while edges (links connecting two nodes) represent friendship, contacts or distance. Exploration is based upon well-defined edges between certain nodes. In so-called psychological networks [3], nodes represent psychological variables such as symptoms and psychological constructs, while edges represent unknown statistical relationships such as comorbidity, causality, interaction that required to be estimated. Therefore, this class of network models is quite different from social networks, where edges between nodes are well defined [4].

Indeed, psychological networks differ from networks of many other disciplines in one very fundamental aspect. Edges are not observed and need to be estimated and this definitely paves the way of novel problems for statistical inference. Recently, psychological variables are perceived just like interacted particles so as directly effecting each other rather than utterly causing an unobserved latent entity. For instance, considering depression as a latent variable, "common cause model" tells us depression is the cause of all symptoms, on the other hand, "network model" tells that symptoms are essentially dynamical systems that cause depression. So, there is a certain need of focusing on analyzing psychological data.

There are number of research papers published dealing with the estimation of network models [5–8]. However, those studies rarely focus on analyzing psychological data rather they are mostly conceptual and require deep technical background. In the last five years, network research has gained attention in psychological sciences [1, 9]. Representing edges as statistical relationships has gained substantial footing and has been used in various different fields of psychology, such as clinical psychology [10–13], psychiatry [14–16], personality research [17], social psychology [18]), and quality of life research [19].

Statistical relationships between edges can be interpreted when drawn as a network structure. Edges indicate pathways on which nodes can affect each other. The edges can differ in strength of connection, that we call edge weights in network theory, indicating if a relationship is strong (thick edges) or weak (thin, less saturated edges) and positive (blue edges) or negative (red edges) as common conventions. After a network structure is estimated, the visualization of the graph itself tells us a detailed story of the multivariate dependencies in the dataset. Furthermore, many inference methods from graph theory can be used to assess which nodes are the most important in the network, termed the most central nodes. In network theory, some nodes are more important and special than others in different ways that we defined some of them as strength, betweenness, closeness, and expected influence.

This thesis study, mostly focuses on two centrality measures:i) "Strength" as the sum of the absolute edge weights between a focal node and all other nodes to which it is connected in the network, ii) "Expected Influenced" as sum of edge weights (accounts for negative edges). In networks lacking of red edges, expected influence and strength measures give similar results. Betweenness and closeness centralities were computed and analysed when it is meaningful to interpret. Betweenness looks at how many of the shortest paths between two nodes go through the node in question; the higher the betweenness, the more important a node is in connecting other nodes. Closeness takes the inverse of the sum of all the shortest paths between one node and all other nodes in the network. Therefore investigates how strongly a node is indirectly connected to other nodes in the network.

An important question is, what type of findings can be obtained throughout the usage of this new psychological network analysis? Following is a list of what can be potentially learned via psychological network analysis. Symptoms (nodes) are active causal elements, they can directly lead to one another. Symptoms connecting two disorders can be found. Thus comorbidity can be revealed. Symptoms are correlated due to dynamic interactions. Therefore, detection of the most important symptom is crucial and can be computed via centrality metrics. How are symptoms uniquely related to one another, in short, finding the interactions among symptoms constitutes the disorder. How symptoms level relationships inform interventions, because targeted treatments will be based on symptoms interactions. After describing the necessity and importance of this recently emerged field, and in the light of the given psychometric network terminology the purpose of the conducted study and research findings will be given in organized fashion at following sections.

Psychometric network modelling first emerged as an alternative to the latent variable modeling. However, it is proven that if a latent variable model is a good fit for the data, it appears as a strong clustering in the network model as well. If both models were driven by using the same data set, it is also possible to show mathematical equivalence between the latent variable model and the psychometric network model. Most common comparison is between Confirmatory factor analysis and Gaussian graphical model which will be discussed in this study [20–23].

II Objective of the Research

This study is fully data-driven. Instead of testing pre-determined hypotheses, data is explored through pre-determined research aims/goals. Through out the study, depression and anxiety is addressed independently from each other while undergoing the same route of research structure.

1. First goal of this study is to explore depression and anxiety symptoms to detect central symptoms and strongly connected symptoms of each. This is to find most efficient ways to intervene the symptoms.

2. Second goal is to establish the causal-link between depression and anxiety symptoms separately. Identifying initiator symptoms can enable and help to design preventive methods.

3. Third goal is to find out how worry and meta-worry concepts interact with depression and anxiety symptoms. This is to pinpoint what kind of worrying minds have alerting status for one's mental health.

4. Fourth goal is to observe where worry and meta-worry elements fit in the causal-link between symptoms of depression and anxiety. This observation will lead to finding out potential detrimental worry and meta-worry concepts to one's mental health.

I considered some arguments which helped me during hypothesis generation. Although depression and anxiety are closely related concepts, it has been stated that anticipation of fewer positive outcomes appears to be one characteristic that may distinguish depression from anxiety. [24] Pessimism and maybe other future-event related concepts within the transparent borders of depression can be key concepts to do research. To paraphrase and emphasize, worry and meta-worry items, particularly future related ones, can make clear cuts or dissect the combined entity of depression and anxiety. If that would be achieved, one could extend the concept of generalized anxiety syndrome (GAD) to mixed anxiety/depression (MAD) by using the worry measures.

To achieve the research goals above; I explore the mental-health data of university students collected over a long period of time through the university web page. I have conducted network analysis to explore three fold interactions between depression-worry-meta worry, and anxiety-worry-meta worry. After revealing the importance of each feature of depression and anxiety measures, correlations of those individual entities with worry and meta worry items are analysed. To do so, first I construct the networks with Gaussian Graphical Model (GGM) as undirected network, estimate a statistical model on data, where parameters can be represented as weighted network between observed variables (symptoms). Second, I apply Graphical Lasso bootstraps for accuracy and stability tests. Subsequently, I check the existence of bridge symptoms, and interpret them. Finally, I apply directed acyclic graph (DAG) model based on the findings of previous analysis.

III Methods

3.1 Datasets

Mental health surveys are self reported measures to assess one's mental health state at a given time. If their reliability and validity are proven, they can also be used as a clinical tool for diagnosing certain mental health disorders. There are considerate number of clinically approved mental health related surveys available. In this study, focus of attention is given to the clinical surveys which are conjugates or closely associated in theory level.Among the data, there are most widely used psychometric tests such as Beck Depression and Anxiety Inventories. The complete set of the surveys which are used in this study is given in Figure 1. The data set was filtered by removing incomplete values and the number of pairwise intersections who responded the same surveys were computed.

	BDI-II	BAI	MCQ	PSWQ
Beck Depression Inventory (BDI-II)	5015			
Beck Anxiety Inventory (BAI)	2157	5430		
Metacognition Questionnaire (MCQ)	3584	1500	3590	
Penn State Worry Questionnaire (PSWQ)	1761	1762	**743**	1762

Figure 1: Original participant sizes & Pairwise overlaps.

3.2 Measures

For the purpose of this study, node name abbreviations are all presented in the Figure 3 in respect to the item order in the original survey starting from A1 to A21 for anxiety, from D1 to D21 for depression, sub-scale node names from M1 to M5 for Meta-cognition and certain selected items for Worry scale. Summary of the item numbers, scoring, and the selected items are summarized in Figure 2.

Scales	Number of Items	Scoring	Selected Items
BDI-II	21	0-3	All
BAI	21	0-3	All
PSWQ	16	1-5	8 Items (PSWQ-Short)
MCQ	30	1-4	5 Sub Scales

Figure 2: Summary of chosen scales, their scoring and number of selected items from each.

Beck Depression Inventory (BDI-II)

The BDI-II is a self reporting tool to assess one's depression symptoms (Beck et al., 1996 [25]). 21 item scored based on a 4-point Likert scale from 0 to 3 to check psychosomatic and cognitive symptoms of depression. The Korean version of the survey showed good reliability and validity in Korean adult population [26] and university student population both online [27] and offline [28]. All of the 21 items of the BDI-II measure were used as nodes whenever depression is included in a network estimation.

Beck Anxiety Inventory (BAI)

BAI have 21 items to self-assess one's psychosomatic and cognitive anxiety symptoms [29]. Symptoms are rated on a 4-point Likert scale from 0 (Not at all) to 3 (Severely-it bothered me a lot). The Korean version of the survey was used which showed good reliability and validity in Korean adult sample [30]. All of the 21 items of the BAI measure were used as nodes whenever anxiety is included in a network estimation.

Meta-cognition Questionnaire (MCQ-30)

The MCQ-30 questionnaire consists of 30 questions to evaluate one's meta-cognitive beliefs and processes by self reporting. Each item is scored based on a 4-point Likert scale from 1 (Do not agree) to 4 (Agree very much). It is divided into five sub-scales that have equal number of questions making their total score minimum 6 and maximum 24. Subscales are named as following: "(Lack of) cognitive confidence", "Positive Beliefs about worry", "Cognitive self-consciousness", "Negative beliefs about uncontrollability and danger", "Need to control thoughts". Korean version of the scale were used [31]. Instead of each item of the questionnaire, only the sub-scale scores were included in the network analysis.

Penn State Worry Questionnaire (PSWQ)

The PSWQ is a self reporting tool to assess one's excessive or uncontrollable worries [32]. It has 16 items rated on a 5-point Likert scale from 1 (not at all typical of me) to 5 (very typical of me). Korean version of the PSWQ showed good validity and reliability [33]. To avoid items measuring the same thing in the network only 8 items were included in the network analysis. To assure theoretical support, item selection was done based on the short version of the PSWQ. Selected items were named in respect to the item order in the original survey being: W2, W4, W5, W6, W7, W9, W12, W13.

Node Names	BDI-II items	Node Names	BAI items	Node Names	MCQ subscales
D1	Sadness	A1	A1-Numbness or tingling	M1	(Lack of) cognitive confidence
D2	Pessimism	A2	A2-Feeling hot	M2	Positive Beliefs about worry
D3	Past failure	A3	A3-Wobbliness in legs	M3	Cognitive self-consciousness
D4	Loss of pleasure	A4	A4-Unable to relax	M4	Negative beliefs about uncontrollability and danger
D5	Guilty feelings	A5	A5-Fear of worst happening	M5	Need to control thoughts
D6	Punisment feelings	A6	A6-Dizzy or lighthe aded		PSWQ items
D7	Self-dislike	A7	A7-Hearth pounding / racing	W2	My worries overwhelm me
D8	Self-criticalness	A8	A8-Unsteady	W4	Many situations make me worry
D9	Suicidal thoughts or wishes	A9	A9-Terrified or afraid	W5	I know I should not worry about things, but I just cannot help it
D10	Crying	A10	A10-Nervous	W6	When I am under pressure I worry a lot.
D11	Agitation	A11	A11-Feeling of choking	W7	I am always worrying about something
D12	Loss of interest	A12	A12-Hands trembling	W9	As soon as I finish one task, I start to worry about everything else I have to do.
D13	Indecisiveness	A13	A13-Shaky / unsteady	W12	I have been a worrier all my life
D14	Worthlessness	A14	A14-Fear of losing control	W13	I notice that I have been worrying about things
D15	Loss of energy	A15	A15-Difficulty in breathing		
D16	Changes in sleeping pattern	A16	A16-Fear of dying		
D17	Irritability	A17	A17-Scared		
D18	Changes in appetite	A18	A18-Indigestion		
D19	Concentration difficulty	A19	A19-Faint / lightheaded		
D20	Tiredness or fatigue	A20	A20-Face flushed		
D21	Loss of interest in sex	A21	A21-Hot / cold sweats		

Figure 3: Abbreviations of questionnaire items used as network nodes

3.3 Participants

The mental-health data including BDI-II, BAI, PSWQ and MCQ-30 measure scores of Korean university students from the same school was collected online from 2008 to April 2011. Students were able to selectively attend to each measure. Among them, 5015 students participated to BDI-II, while 5430 students participated to BAI, 3590 students to MCQ-30 and finally 1762 students to PSWQ measures. Only 743 students attended to all 4 measures which was also the intersection set between PSWQ and MCQ-30 pairs participants.Table 4 3 data-sets were chosen to be used for different network analysis. For networks which includes depression symptoms only, all 5015 (2379 female, 2636 male) participants for BDI-II measure were used aging from 17 to 65 with mean 23,01 and standard deviation 2,35. For networks which includes anxiety symptoms only, all 5430 (2694 female, 2736 male) participants for BAI measure were used aging from 18 to 47 with mean 22,86 and standard deviation 2,46. For all the other networks, 743 participants who attended to all 4 measures were used. Among them 366 were female, 377 were male aging from 19 to 31 with mean 23,12 and standard deviation 2,13.

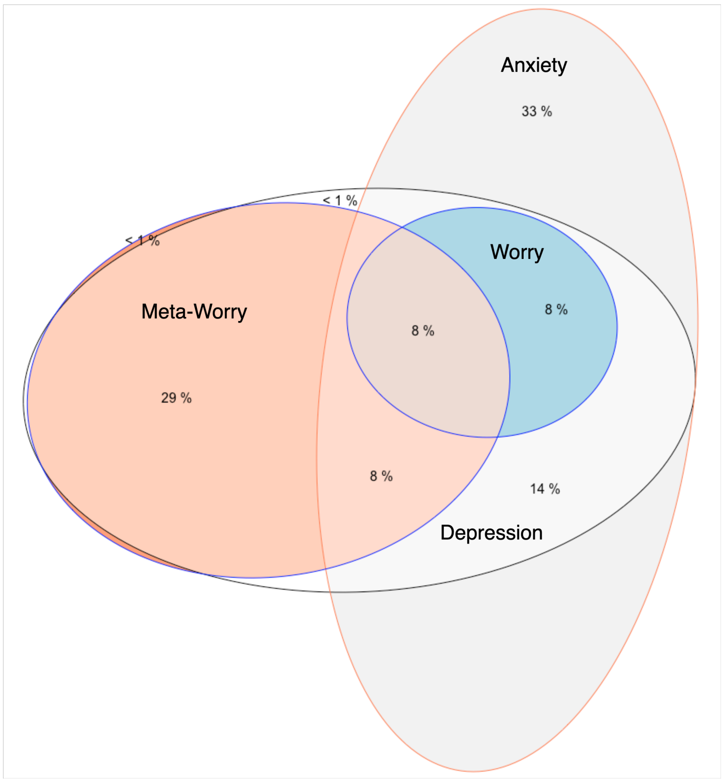


Figure 4: Area proportional Euler Diagram which shows the set properties of curated data collected over the years.

3.4 Networks

Regularized partial correlation network(GLASSO)

In a psychometric network, each node represents an item or a symptom of a measure and each edge represents a relationship between nodes. Blue edges shows positive correlations and red edges accounts for negative correlations. The relationship between two nodes can be assessed in two ways. First, we can directly assess the edge weight. If edges are weighted, thicker edges represents stronger relationships. Edge weight of zero indicate there is no edge, however most of the time edges are nonzero. The sign of the edge weight indicates the type of interaction while the absolute value indicates the strength of the effect. As such, the length of an edge is defined as the inverse of the edge strength. The distance between two nodes is equal to the sum of the lengths of all edges on the shortest path between two nodes. In a partial correlation network, correlations between each node are calculated by controlling the influence of all the other nodes. This allows the network to be sparse, meaning, representing most variance with the fewest number of edges possible.

There are several methods for computing partial correlation coefficients as described by Cohen et al [34]. Here we focus on two widely used approaches to obtain partial correlations, namely standardizing the precision matrix or performing node-wise regressions. In essence, both approaches leads to the exact same estimate. At first method, we can obtain the partial correlations directly from the inverse of a variance–covariance matrix. Let **y** represents a set of item responses, and Σ denote a variance–covariance matrix. By assuming **y** follows a multivariate normal distribution:

$$\boldsymbol{y} \sim N(\boldsymbol{O}, \boldsymbol{\Sigma}).$$

The precision matrix, denoted by the letter **K**, can then be defined as the inverse of Σ

$$\boldsymbol{K} = \Sigma^{-1}$$

we can standardize the elements of the precision matrix between variable y_i and y_j, after conditioning on all other variables in $\boldsymbol{y}_i$, $\boldsymbol{y}_{-(i,j)}$ [35].

$$Cor(y_i, y_j | \boldsymbol{y}_{-(i,j)}) = -\frac{\kappa_{ij}}{\sqrt{\kappa_{ii}}\sqrt{\kappa_{jj}}}.$$

Second method is to use node-wise regressions [36] to obtain partial correlation coefficients. If y_i is predicted from all other variables in a multiple regression, then:

$$y_1 = \beta_{10} + \beta_{12} y_2 + \beta_{13} y_3 + \cdots + \varepsilon_1,$$

followed by a similar regression model for y_2:

$$y_2 = \beta_{20} + \beta_{21} y_1 + \beta_{23} y_3 + \cdots + \varepsilon_2,$$

Likewise, we can compute the above formula for y_3, y_4, etc. The regression slope predicting y_i from y_j ory_j from y_i is proportional to same partial correlation coefficient between y_i and y_j [37].

$$Cor(y_i, y_j | \mathbf{y}_{-(i,j)}) = \frac{\beta_{ij} SD(\varepsilon_j)}{SD(\varepsilon_i)} = \frac{\beta_{ji} SD(\varepsilon_i)}{SD(\varepsilon_j)}$$

where SD is the abbreviation for standard-deviation.

In this study, four different partial correlation networks were estimated. First two networks were computed for only the depression symptoms and the anxiety symptoms separately. Then *EBICglasso* function of the "qgraph" package [38] was used to estimate all four networks. In the first network, only the depression symptoms (BDI-II items) were included, therefore 5015 participant data were used to estimate it. Similarly, in the second network, only the anxiety symptoms (BAI items) were used with 5430 participant data. In the remaining two networks, selected PSWQ and MCQ-30 items were combined with depression symptoms and anxiety symptoms in separate networks, hence 743 participants who attended all measures were used. Later two networks were mainly used for bridge analysis only.

To quantify the importance of each node in the network, centrality indices are computed. Centrality indices reflect how connected a node is within the network and hence how potentially clinically relevant it may be. The betweenness centrality of a node equals the number of times that it lies on the shortest path between any pair of other nodes. Closeness centrality signifies the average distance of a node to all other nodes in the network, calculated as the inverse of the weighted sum of shortest path lengths of a given node to reach all other nodes network. Node strength is the sum of the absolute value of the edge weights connected to a node. While expected influence is the sum of all the value of the edge weights connected to a node. If there are red edges in the network, expected influence becomes important. Of the three centrality indices, node strength may be the most relevant index of importance for modelling symptom networks since it is proven to be the most stable of them all. Similarly, centrality metrics were plotted for the normalized (z-scored) values for each node by using "networktools" and qgraph packages. Robustness of our findings is also evaluated by using the R package "bootnet". First, stability of the centrality measures were checked with the case dropping boot-strap by sampling the data 10000 times. Then, the accuracy of the edge weights are estimated by employing a non-parametric boot-strap approach to calculate the 95% confidence intervals for the edges by sampling the data 10000 times (with replacement), thereby generating a distribution of edge weights.

Bridge Analysis

Psychometric researchers in psychopathology have relied on visual inspection of networks to identify bridge symptoms [39–42] Bridge analysis is a tool to detect comorbidity between different measures. After estimating a GLASSO network, nodes are divided into communities to estimate bridge analysis between them. Bridge centrality measures are derived as an extension of regular centrality measures that we use in network analysis and we call them bridge strength, bridge betweenness, bridge closeness, and bridge expected influence. The precise definitions of those centrality metrics in mathematical terms are given below for psychological networks [43]. Before passing through to the definitions, we need to introduce the following notations that apply to all of the statistics below. A network with a set of V nodes and E edges can generally be conceptualized as G(V,E). Let C be a community in this network, implying that C is a proper subset of V. Each of the bridge centrality statistics below are defined with respect to a selected node a within community C. The notation N(a) is used to represent the neighbour nodes of a while w_{ab} denote the weight on each edge ab belongs to E. Following the aforementioned notations, we can define four bridge centrality measure as below.

1. Bridge strength indicates a node's total connectivity with other community.

$$\text{bridge strength} = \sum_{b\in(N(a)-C)} |w_{ab}|.$$

2. Bridge betweenness assesses the number of times a node lies on the shortest path between any two nodes. Letting P_{ij} be a shortest path between $i, j \in V$, where the nodes i and j are not belong to same community.

$$x = \begin{cases} 0.5, & a \in P_{ij} \\ 0, & otherwise \end{cases}$$

$$\text{bridge betweenness} = \sum_{i\in V} x_i.$$

3. Bridge closeness reflects the average distance from a node to all nodes outside of its own disorder. if $a \in C \wedge b \notin C$, and P_{ab} be the shortest path among a and b, including edges $E(P_{ab} = e_i, \cdots, e_k, \cdots, e_n$ where each edge has weight w_k for $1 \leq k \leq n$.

$$\text{bridge closeness} = \frac{|V-C|}{\sum_{b\in(V-C)} \sum_{e_k\in E(P_{ab})} \frac{1}{w_k}}.$$

4. Bridge expected influence, much like bridge strength, indicates a node's sum connectivity with other disorders and can be expressed as:

$$\text{bridge expected influence} = \sum_{b\in(N(a)-C)} w_{ab}.$$

In the case of bridge expected influence, absolute value of edges is not taken before summing them. For this reason, this statistic is more useful in the presence of negative edges.

So, In addition to regular centrality measures, we calculated bridge centrality measures and their stabilities. Bridge centrality is calculated for each node connecting to every node in the other community. After detecting the greatest bridge centrality, correlation matrix between nodes is checked to find which node in the other community it is most strongly connected to. Thus, both sides of the bridge can be established. In this study, worry and meta-cognition nodes were grouped together as one community. Their relationship between depression symptoms and anxiety symptoms were examined as separate networks.

Directed Acyclic Graph network(DAG)

For ages, philosophers and scientists have struggled to figure out how to infer causal relationships. Over the past few centuries, little progress has been made in elucidating the circumstances that allow causal inference to be made. This began to change in the late 1980s, when researchers realized that studying multivariate systems allow for more robust conclusions. Indeed, provided the researcher is willing to invest the required assumptions, causal conclusions can be drawn from correlational data under certain circumstances. Judea Pearl's book *Causality* [44] is responsible for many of the theories discussed here so that I will try to explain the theoretical foundation of DAGs based on this excellent book.

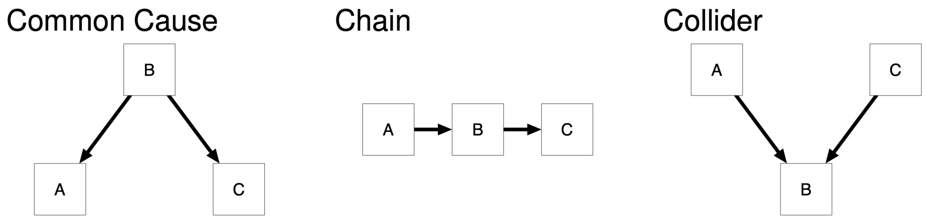


Figure 5: Building blocks of a Directed Acyclic Graph. Three important causal structures.

As it shown in Figure 5, there are three main blocks of a DAG. We can interpret the situation in abstract terms to exemplify each case. For example, in the common cause case, disease B simply causes two symptoms – A and C. Here, A and C conditionally independent given B. For the chain case, disease A causes B, which in turn causes C. Here again, A and C conditionally independent given B Finally, in the case of collider, A and C jointly cause a third variable B. Here, A and C conditionally dependent given B.

It is practically possible to narrow down the number of causal possibilities by looking at the conditional independence relations. However, as apparent, the systems in Figure 5 are only for demonstration. In this study, I generated DAGs with 21 and 34 items. Testing such causal models involves testing whether all conditional independence relations hold. Thus, there is a need for a method to analyze larger systems. Such method should test causal models against correlational data and also search the data for possible causal models that are consistent with them. Luckily, there are very simple rules by which we can look at a causal graph and derive which conditional independence relations should hold in the data if that graph were true. The method to do this is called d-separation. However, we need to explain the concept of blocking before explaining what d-separation is.

To identify if two variables are conditionally independent given a third variable or set of multiple variables, one needs to list all paths between the variables by ignoring direction of edges. After, the variable is checked if the following two condition is satisfied.

- The middle node in a chain or common cause structure
- Not the middle node in a collider structure or an effect of such a common effect

If the above conditions are satisfied then the path is called "blocked". Likewise, if all such paths are blocked, the two variables are d-separated and thus conditionally independent.

The d-separation is an extremely powerful concept, because one only has to look at the graph to determine that, for example, two variables A and B are d-separated by C; if one sees that they are, then it follows that the variables A and B are conditionally independent given C in all probability distributions that are consistent with the graph, regardless of how the variables are distributed or what the functional form of the relation between is.

To compute a Bayesian network [45], visualized as a DAG, the hill-climbing algorithm furnished by the R package bnlearn was run. The bootstrap function of bnlearn calculates the structural aspect of the network by adding edges, subtracting them, and reversing their direction to optimize the Bayesian Information Criterion (BIC) which is a target score. The inner mechanism of how this algorithm works can be described as follows: First step ascertains whether an edge between two symptoms exists; it does not determine its weight. Then, this procedure is randomly restarted with various candidate edges potentially linking different pairs of symptoms, perturbed the system, and so forth. As this iterative process unfolds, the algorithm discerns the network's structure.

The Bayesian analysis was done over 4 different datasets following the same framework of GLASSO networks. In the first directed network, only the depression symptoms (BDI-II items) were included, therefore 5015 participant data were used to estimate it. Similarly, in the second directed network, only the anxiety symptoms (BAI items) were used with 5430 participant data. In the remaining two networks, selected PSWQ and MCQ-30 items were combined with depression symptoms and anxiety symptoms in separate networks, hence 743 participants who attended all measures were used. To ensure the stability of the DAG, 10000 samples are bootstrapped, computing a network for each sample. Later, they are averaged to obtain the final, resultant network. There are two steps involved. First, determining how often an edge appeared in the 10000 bootstrapped networks. Only edges that appears more then a given threshold makes it to the final network. A fixed threshold can be used such as 85% as used in other papers.Scutari and Nagarajan's (2013) [46] statistically-driven method can be used to find an optimal cutpoint for retaining edges in the final, averaged network. Second, determining the direction of each edge in each of the 10000 bootstrapped networks. If an edge pointed from symptom X to symptom Y in at least 51% of the networks, then this direction was depicted in the final, averaged network.

Visualizations of the final, averaged network is two-fold. First, a DAG whose edges depicted the BIC value of an edge was computed. High absolute BIC values signify the importance of an edge to the model that best captures the structure of the data. Edge thickness depicts the magnitude of the BIC value. The thicker an edge, the more damaging it would be to model fit if the edge were removed from the network.

Second, a DAG whereby edge thickness signifies the probability that the edge points in the direction depicted was computed. Hence, if an edge pointed from symptom X to symptom Y in 9000 of 10000 bootstrapped networks, it would appear very thick. If it pointed from symptom X to symptom Y in only 5100 of 10000 bootstrapped networks, it would appear very thin.

IV Results

4.1 Regularized partial correlation networks

In this study, three different participant groups were used. For depression only analysis 5015, for anxiety only analysis 5430 and for worry and meta-worry included version of depression and anxiety analysis 743 students' data were used. Therefore, descriptive statistics for depression and anxiety were calculated for two groups, one being a subset group of students who attended all four measures (N=743).

Table 1: Descriptive statistics for two participant groups of the depression measure (BDI-II).

	N=5015 (2379 F, 2636 M)			N=743 (366 F, 377 M)		
Symptoms	mean	sd	r	mean	sd	r
Age	23.01	2.35	-	23.12	2.13	-
1.Sadness	0.45	0.55	0.61	0.50	0.58	0.62
2.Pessimism	0.49	0.65	0.61	0.50	0.67	0.61
3.Past Failure	0.44	0.64	0.60	0.48	0.67	0.60
4.Loss of Pleasure	0.49	0.66	0.63	0.52	0.70	0.60
5.Guilty Feelings	0.60	0.70	0.54	0.61	0.68	0.51
6.Punishment Feelings	0.26	0.62	0.51	0.28	0.65	0.50
7.Self-dislike	0.38	0.70	0.63	0.42	0.73	0.67
8.Self Criticalness	0.44	0.71	0.63	0.48	0.73	0.60
9.Suicidal Thoughts or Wishes	0.24	0.48	0.51	0.25	0.49	0.51
10.Crying	0.32	0.63	0.47	0.36	0.67	0.48
11.Agitation	0.33	0.53	0.56	0.37	0.54	0.56
12.Loss of Interest	0.52	0.71	0.63	0.56	0.72	0.60
13.Indecisiveness	0.58	0.68	0.55	0.62	0.70	0.56
14.Worthlessness	0.27	0.58	0.59	0.32	0.61	0.64
15.Loss of Energy	0.56	0.70	0.69	0.63	0.74	0.70
16.Changes in Sleeping Pattern	0.83	0.75	0.54	0.88	0.75	0.51
17.Irritability	0.41	0.61	0.55	0.44	0.63	0.56
18.Changes in Appetite	0.66	0.71	0.50	0.70	0.73	0.48
19.Concentration Difficulty	0.42	0.60	0.62	0.45	0.62	0.60
20.Tiredness or Fatigue	0.59	0.61	0.64	0.64	0.61	0.66
21.Loss of Interest in Sex	0.23	0.54	0.38	0.28	0.63	0.40
Total	9.49	8.09	1.00	10.26	8.47	1.00

First, single measure GLASSO networks were estimated starting from BDI-II measure, depicting the relationship between depression symptoms(see Figure 6). Centrality stability coefficients calculated by the case dropping boot were lowest for betweenness being 0.28 which is considered just above the acceptable stability for interpretation. Closeness, strength, and expected influence scored 0.75 which was the highest level tested and considered as excellent scores. However, in the centrality stability plot seen in Figure 7, it can be observed that stability of closeness is lower than strength and expected influence which are close to the ceiling limit. In the bottom graph of the same figure, it shows excellent edge weight stability that was plotted through the non-parametric bootstrap.

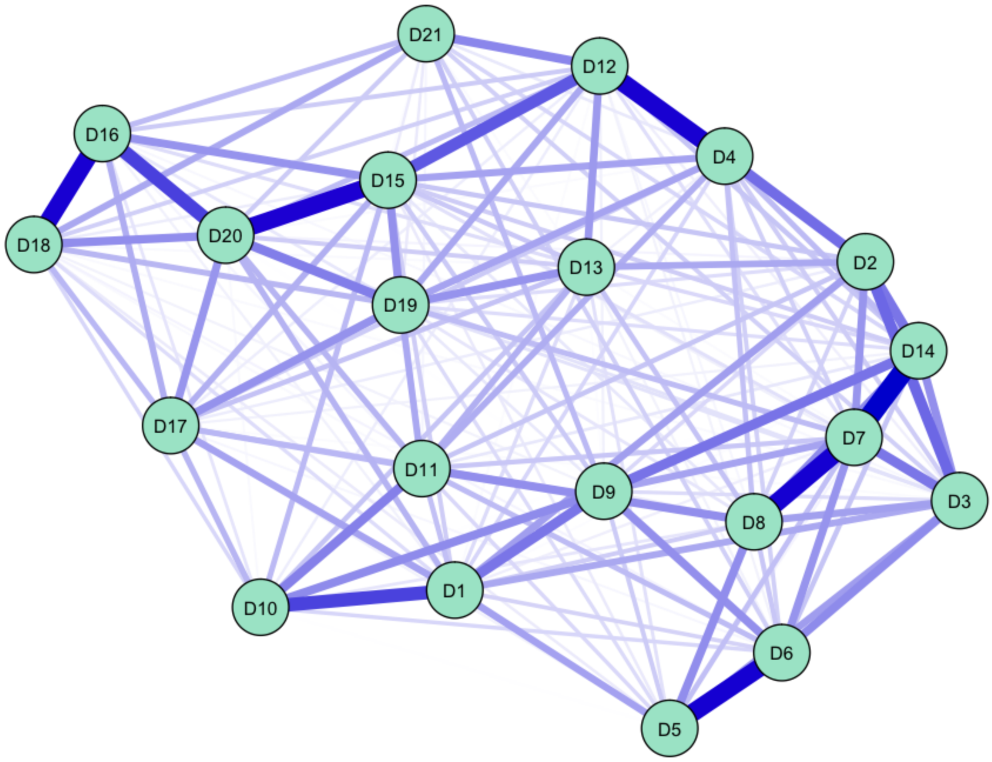
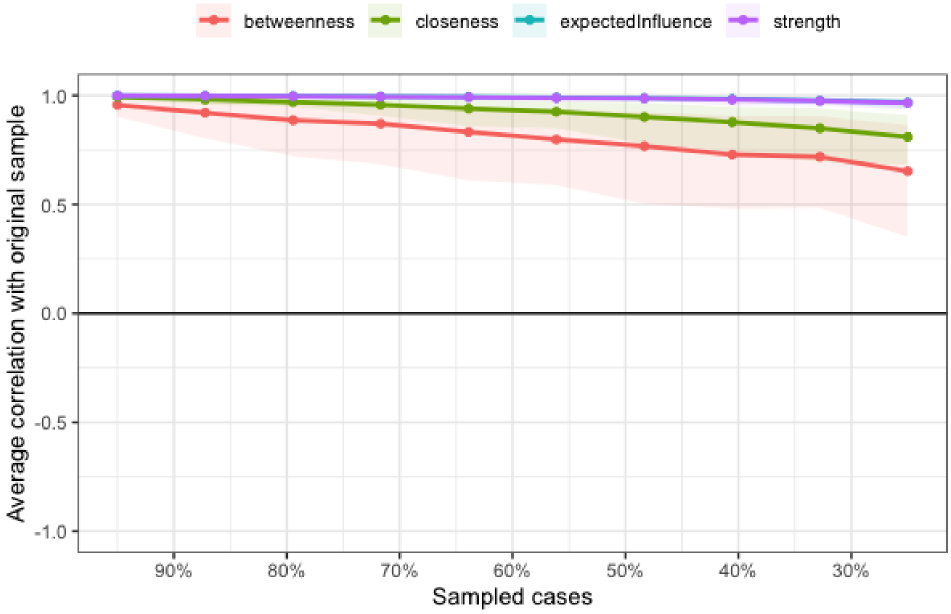


Figure 6: Regularized partial correlation network returned via the graphical LASSO depicting associations between pairs of depression symptoms.

In Figure 6, edges that connect the "(D7)Self-dislike" node to both "(D14)Worthlessness" and "(D8)Self-criticalness" nodes appear to be among the strongest edges. Apart from the above-mentioned nodes, other strongest edges were observed in between the following nodes in order: "(D5)Guilty feelings" and "(D6)Punishment feelings", "(D4)Loss of pleasure" and "(D12)"Loss of interest", "(D16)Changes in sleeping pattern" and "(D18)Changes in appetite", and lastly "(D15)Loss of energy" and "(D20)Tiredness or fatigue". (See Supplementary Figure 36 for more information)

In the figure 8, "(D7)Self-dislike" and "(D15)Loss of energy" symptoms emerged as having the greatest node strength followed by "(D14)Worthlessness" and "(D20)Tiredness or fatigue" symptoms.



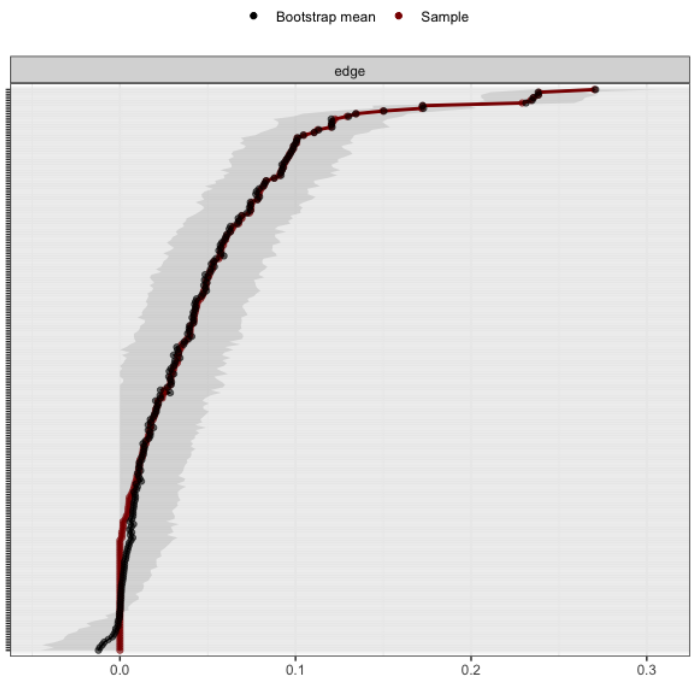


Figure 7: Above : Centrality stability graphs for strength, closeness, betweenness, expected-influence in the case-dropping stability analysis for depression network. Below : Edge stability graph in the non-parametric stability analysis for depression network

These nodes were revealed to be significantly different from the remaining nodes (See Figure 9 for more detail). The expected influence measure supports the same result as the strength measure since there were no apparent red edges in the network. Although the stability of the betweenness and closeness measures were acceptable and good respectively, they did not provide any centrally strong nodes. Because the nodes having the strongest betweenness and closeness values were not significantly different from the majority of the remaining nodes.

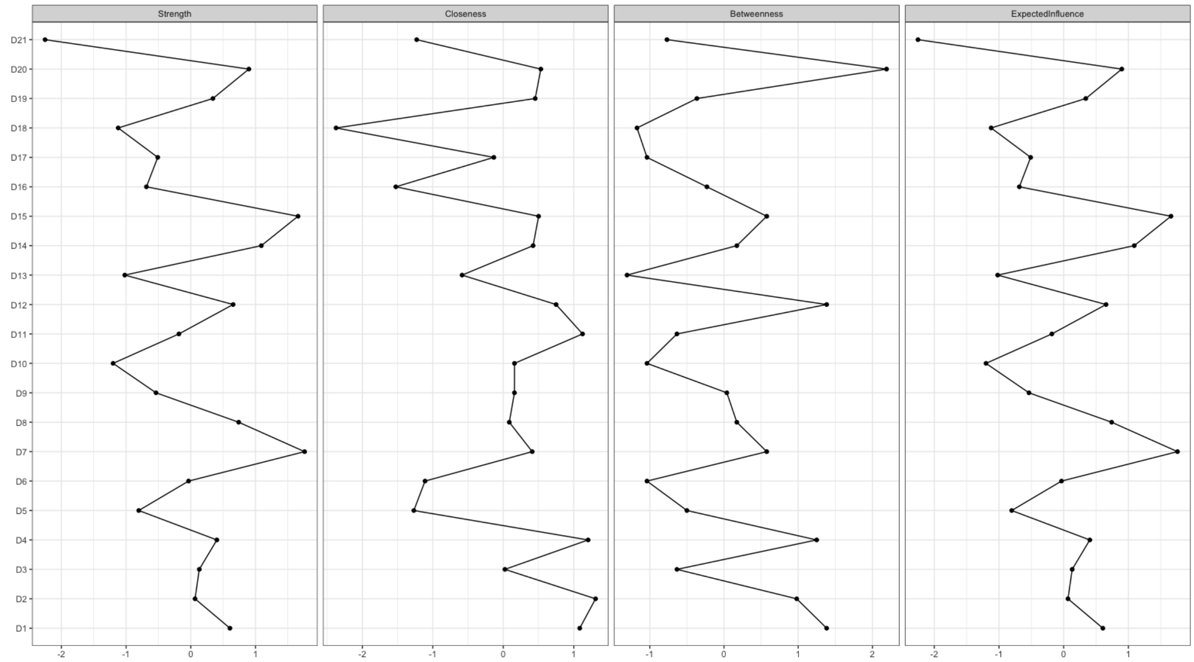


Figure 8: z-scored centrality metrics (betweenness, closeness, expected-influence, strength) for each depression symptom.

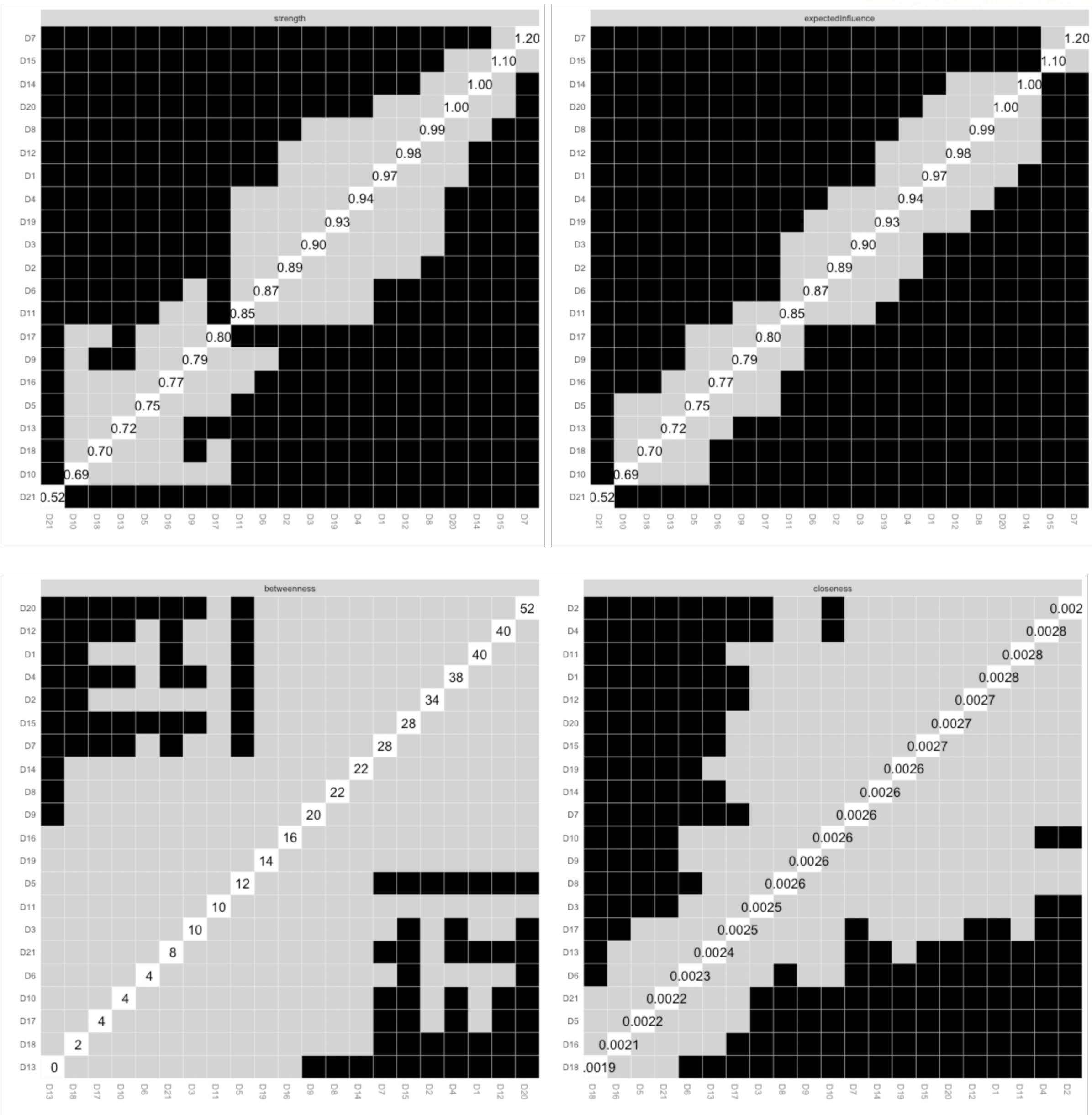


Figure 9: Centrality differences between depression nodes for strength, expected influence, betweenness and closeness measures. Black boxes indicate significant difference while grey boxes means significantly not different.

In the Figure 10A, there is no recognisable clustering between nodes meaning latent variable models may not be able to fully explain the complexity of the relationship between depression symptoms in given survey. In the previous studies which showed validity and reliability of the Korean version of the BDI-II survey, factor analysis was reported along with them. In the Figure 10B, each color represents different factors. Three factors structure reported to be the best fit in the Yu et al. for Korean university students. Although same colored nodes are relatively close, it does not fit well with current data [28]. In the Figure 10C, two factors structure reported to be the best fit in the Song et al. for Korean university students. This fitting seems to be a total miss in terms of representing the current data [27] Lastly in the Figure 10D, three factors structure reported by Park et al. seems to be the best fit for our data. However, unlike other two studies, factors reported here were generalized adult population not for student population [26].

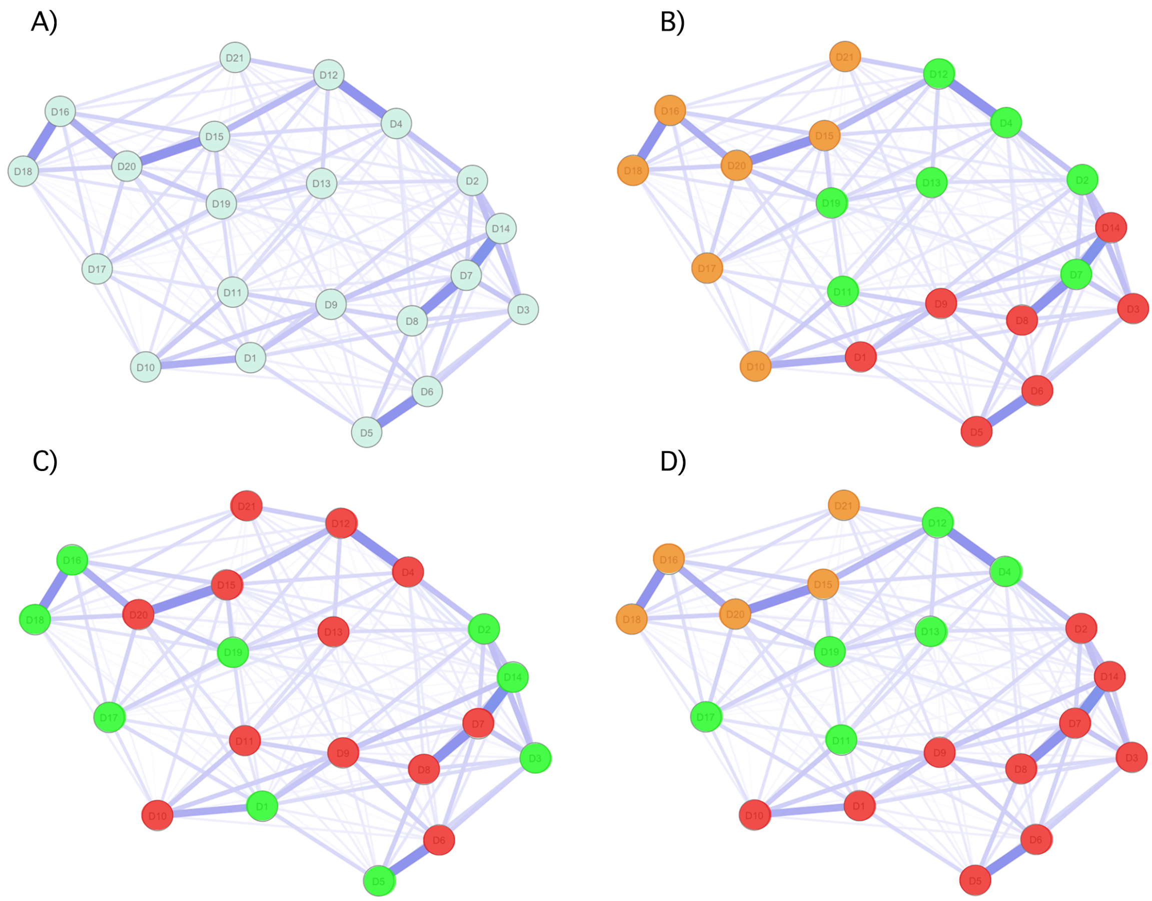


Figure 10: Representation and comparison of Factor analysis on BDI-II items done by previous studies in Korean subjects.

Table 2: Descriptive statistics for two participant groups of the anxiety measure (BAI).

	N=5430 (2694 F, 2736 M)			N=743 (366 F, 377 M)		
Symptoms	mean	sd	r	mean	sd	r
Age	22.86	2.46	-	23.12	2.13	-
1.Numbness or tingling	0.51	0.62	0.58	0.44	0.62	0.58
2.Feeling hot	0.72	0.66	0.53	0.61	0.63	0.55
3.Wobbliness in legs	0.55	0.67	0.54	0.47	0.64	0.50
4.Unable to relax	0.44	0.65	0.57	0.40	0.62	0.55
5.Fear of worst happening	0.49	0.68	0.63	0.48	0.69	0.65
6.Dizzy or lightheaded	0.72	0.74	0.59	0.63	0.74	0.62
7.Hearth pounding / racing	0.77	0.75	0.63	0.67	0.74	0.66
8.Unsteady	0.65	0.67	0.59	0.55	0.64	0.59
9.Terrified or afraid	0.62	0.77	0.63	0.55	0.74	0.64
10.Nervous	0.68	0.78	0.68	0.57	0.75	0.63
11.Feeling of choking	0.21	0.53	0.49	0.17	0.47	0.46
12.Hands trembling	0.33	0.61	0.47	0.28	0.57	0.45
13.Shaky / unsteady	0.36	0.59	0.62	0.32	0.54	0.59
14.Fear of losing control	0.19	0.49	0.52	0.15	0.43	0.51
15.Difficulty in breathing	0.23	0.53	0.46	0.18	0.48	0.46
16.Fear of dying	0.12	0.39	0.40	0.08	0.31	0.37
17.Scared	0.50	0.66	0.65	0.47	0.65	0.66
18.Indigestion	0.79	0.88	0.58	0.76	0.87	0.62
19.Faint / lightheaded	0.11	0.37	0.38	0.09	0.35	0.37
20.Face flushed	0.55	0.78	0.48	0.50	0.77	0.52
21.Hot / cold sweats	0.33	0.62	0.45	0.28	0.61	0.48
Total	9.85	7.88	1.00	8.67	7.55	1.00

Similar to the depression network, a GLASSO network depicting the relationship between anxiety symptoms were estimated(see Figure 11). Centrality stability coefficients calculated by the case dropping boot were lowest for betweenness being 0.57 which is still considered good. Closeness, strength, and expected influence scored 0.75 which was the highest level tested and considered as excellent scores. Similar to the depression network stability, the centrality stability plot in Figure 12 indicates that the stability of closeness is lower than the strength and the expected influence which are close to the ceiling limit. Bottom part of the same figure shows excellent edge weight stability that was plotted through the non-parametric bootstrap.

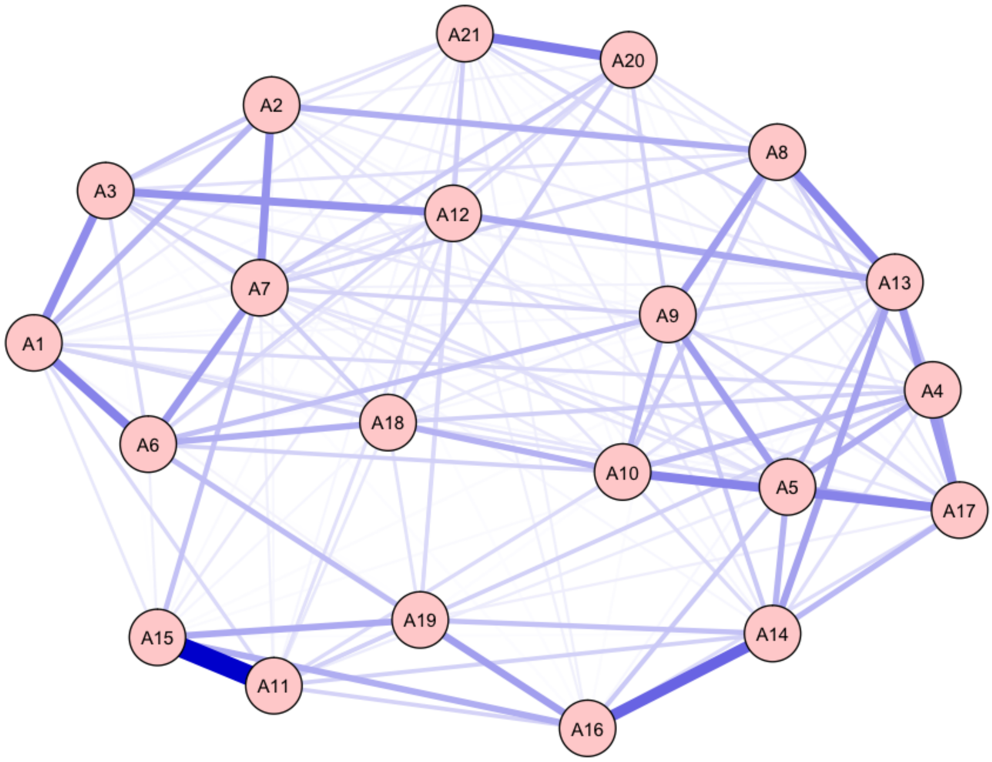


Figure 11: Regularized partial correlation network returned via the graphical LASSO depicting associations between pairs of anxiety symptoms.

The edge that connects the "(A11)Feeling of choking" node to "(A15)Difficulty in breathing" node appears to be the strongest compared to all the other edges in Figure 11. Edges between "(A14)Fear of losing control" and "(A16)Fear of dying" nodes as well as "(A20)Face flushed" and "(A21)Hot / cold sweats" nodes come right after. Apart from the above-mentioned nodes, edges observed in between the following nodes follows in terms of strength: "(A1)Numbness or tingling" with both "(A6)Dizzy or lightheaded" and "(A3)Wobbliness in legs", "(A10)Nervous" and "(A17)Scared", "(A8)Unsteady" and "(A13)Shaky / unsteady" (See Figure 37)

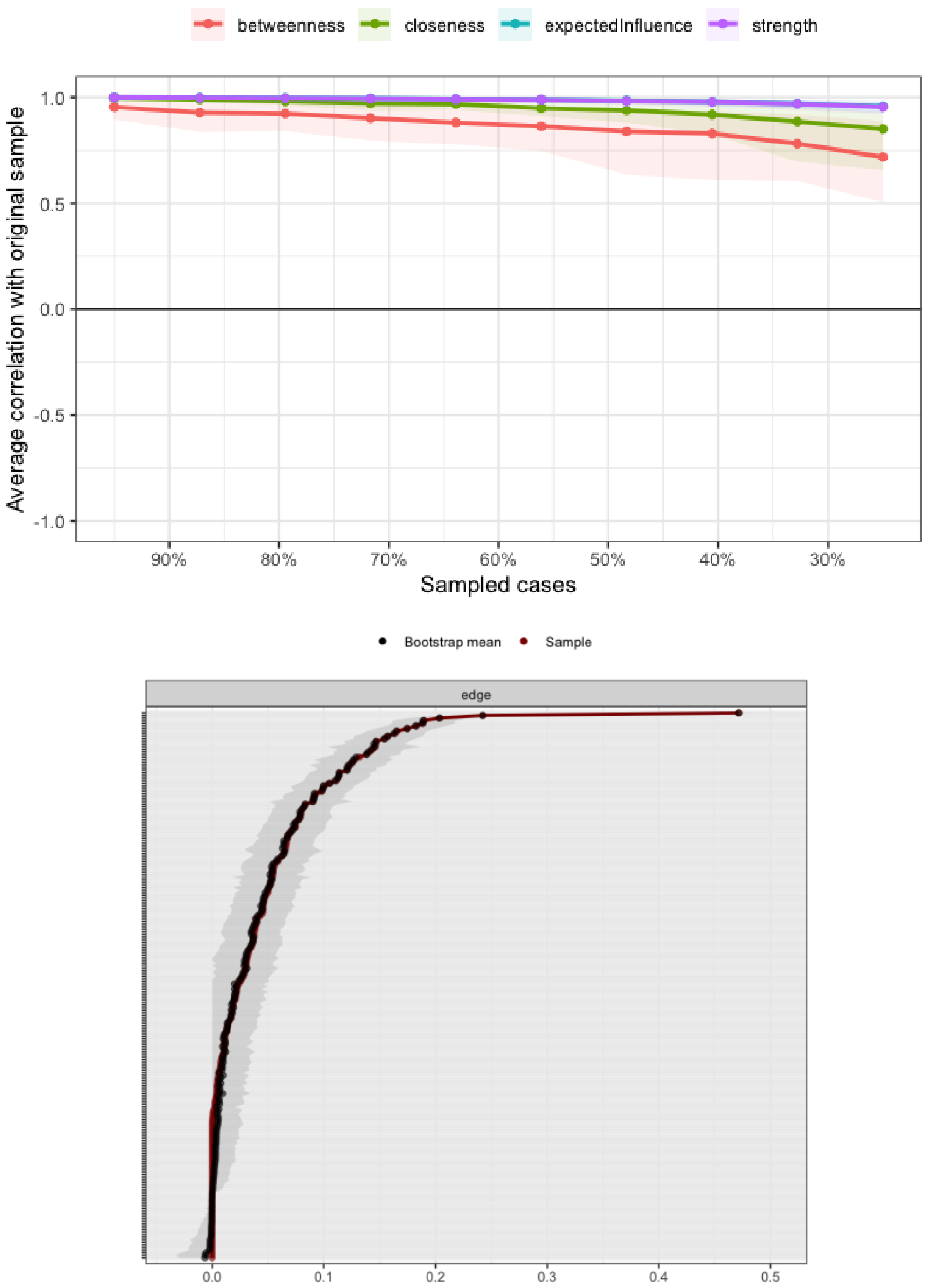


Figure 12: Above : Centrality stability coefficients for strength, closeness, betweenness, expected-influence in the case-dropping stability analysis for anxiety network. Below : Edge stability graph in the non-parametric stability analysis for anxiety network.

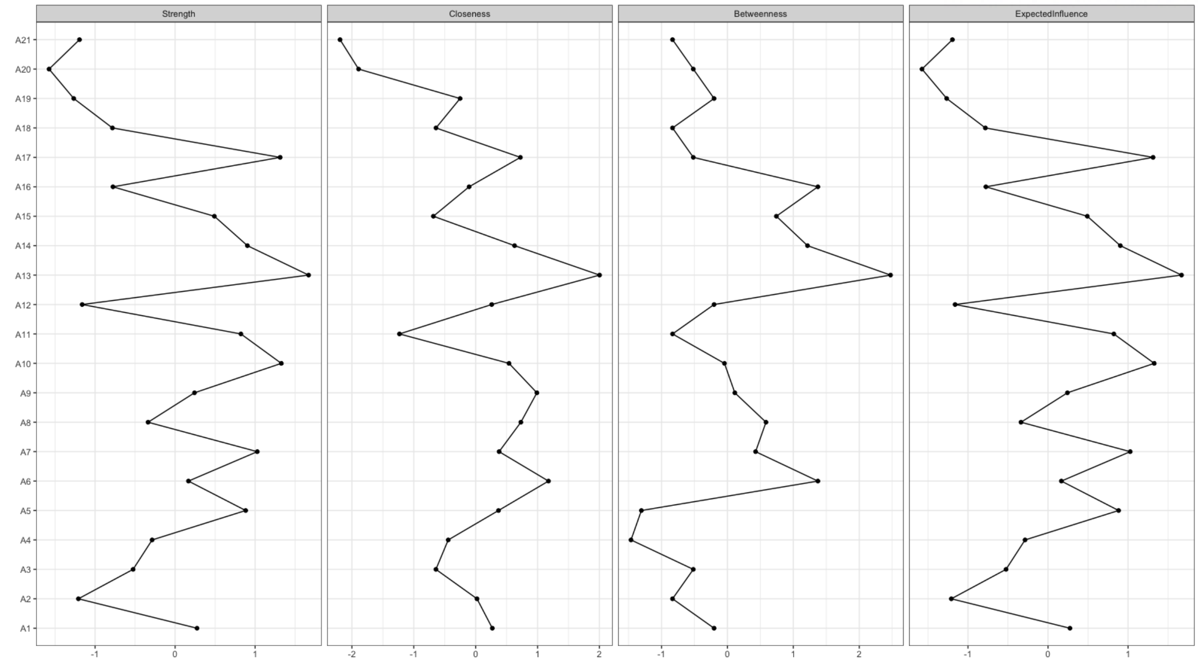


Figure 13: z-scored centrality metrics (betweenness, closeness, expected-influence, strength) for each anxiety symptom.

In the figure 13 "(A13)Shaky / unsteady" symptom emerged as the greatest for all centrality measures. "(A7)Hearth pounding / racing" symptom was greatest for both strength and betweenness. "(A10)Nervous" and "(A17)Scared" nodes were only greatest in the strength centrality while "(A14)Fear of losing control" were only greatest in the betweenness centrality. Meanwhile, "(A6)Dizzy or lightheaded" node was greatest at both betweenness and closeness measures. These nodes were revealed to be significantly different from the remaining nodes (See Figure 14 for more detail). The expected influence measure supports the same result as the strength measure since there were no apparent red edges in the network.

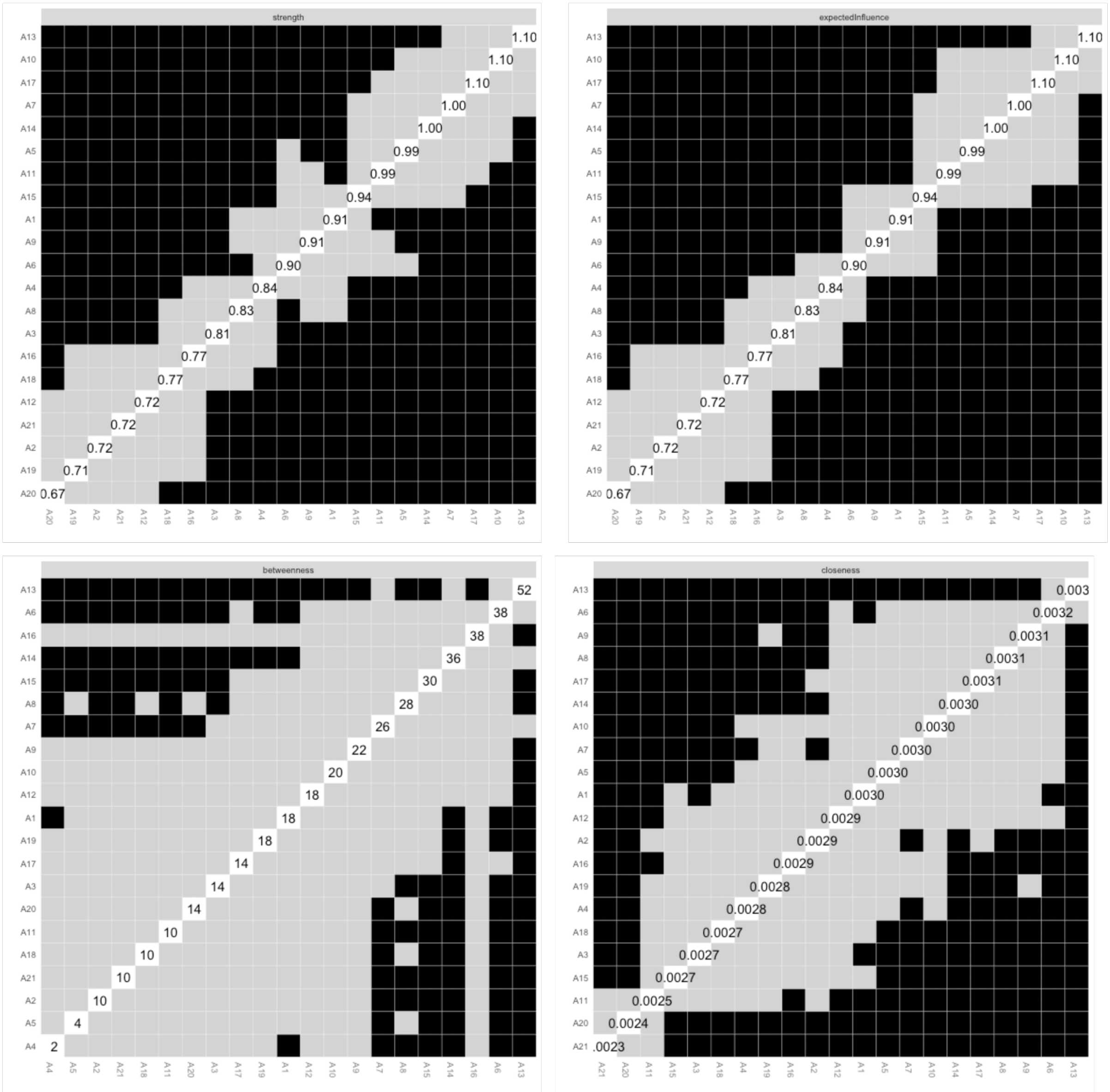


Figure 14: Centrality differences between anxiety nodes for strength, expected influence, betweenness and closeness measures. Black boxes indicate significant difference while grey boxes means significantly not different.

In Figure 15A, contrarily to the depression network, there seems to be three observable clustering among the nodes. First cluster consisting of A11, A14, A15, A16, A19 nodes, second cluster including A4, A5, A8, A9, A10, A13, A14, A17 nodes and the third cluster forming with the remaining nodes. This could hint into 3 latent variable behind the clusters. Further analysis needs to be done to reveal the truth. In Figure 14B, different colors belong to four different factor structure which was reported to be the best fit for Korean adult population by Lee et al. [30]. Clearly it is far from explaining our data set.

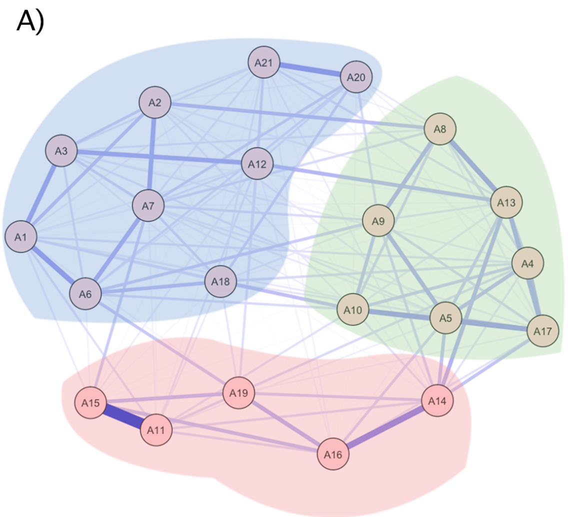


Figure 15: Representation and comparison of Factor analysis on BAI items done by previous studies in Korean subjects.

4.2 Bridge Analysis

Table 3: Descriptive statistics for Worry(PSWQ) items and Meta-worry(MCQ-30) subtotals (N=743).

Symptoms	mean	sd
PSWQ		
2.My worries overwhelm me	4.48	3.40
4.Many situations make me worry	6.45	3.41
5.I know I should not worry about things, but I just cannot help it	7.42	3.37
6.When I am under pressure I worry a lot	4.76	3.54
7.I am always worrying about something	3.64	3.02
9.As soon as I finish one task, I start to worry about everything else I have to do	1.93	0.98
12.I have been a worrier all my life	2.52	1.02
13.I notice that I have been worrying about things	2.34	1.17
MCQ-30		
1.(Lack of) cognitive confidence	2.84	1.11
2.Positive Beliefs about worry	2.05	1.08
3.Cognitive self-consciousness	1.82	1.00
4.Negative beliefs about uncontrollability & danger	2.33	1.20
5.Need to control thoughts	2.62	1.26

For bridge analysis, two different GLASSO networks were estimated by using the same data set of 743 participants. In the first network, depression, worry, and meta-cognition nodes/items were included(Figure 16). Similar to the single measure GLASSO networks above, the stability of the network needs to be good to be able to go forward and interpret the results. Centrality stability coefficients calculated by the case dropping boot for betweenness being 0.20 and closeness being 0.36 were within the acceptable range to establish the network as stable. Strength and expected influence scored 0.67 and considered good. Bridge centrality stability coefficients were the same as before-mentioned peers(Figure 17). Excellent edge weight stability that was plotted through the non-parametric bootstrap can be seen in the same figure as well.

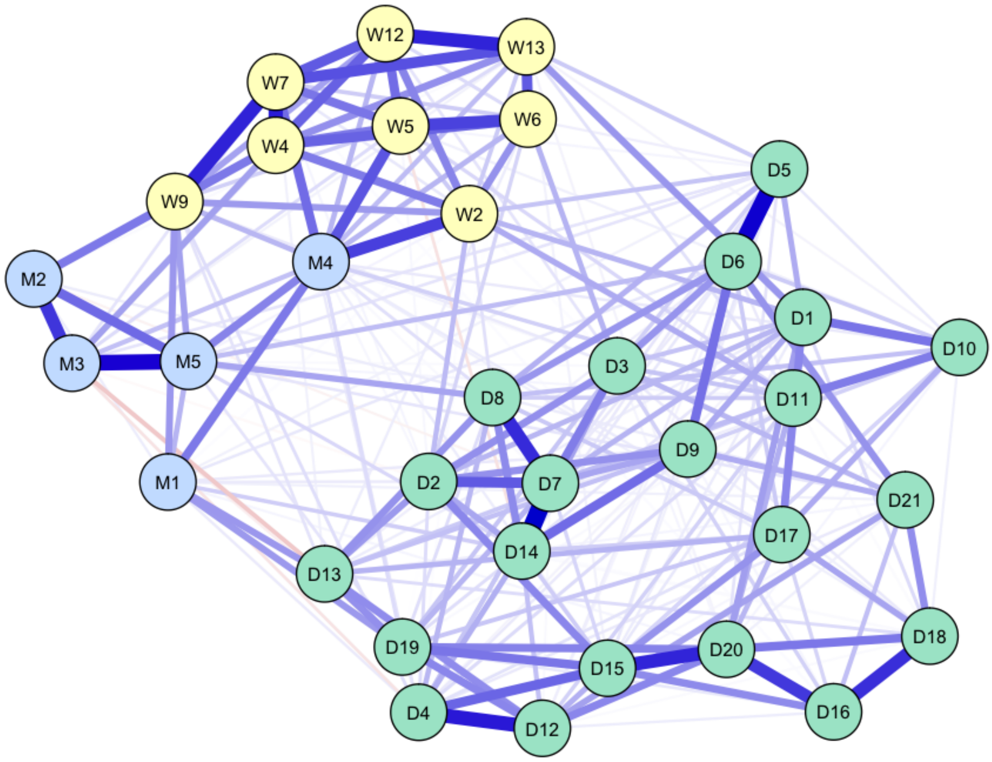
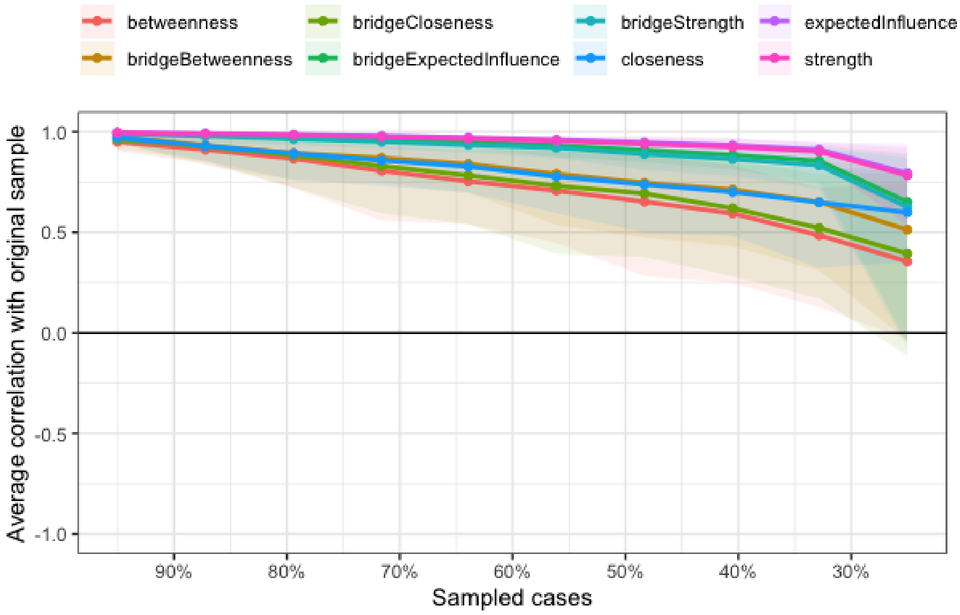


Figure 16: Regularized partial correlation network depicting associations between depression symptoms, meta-cognition subsections and worry items.



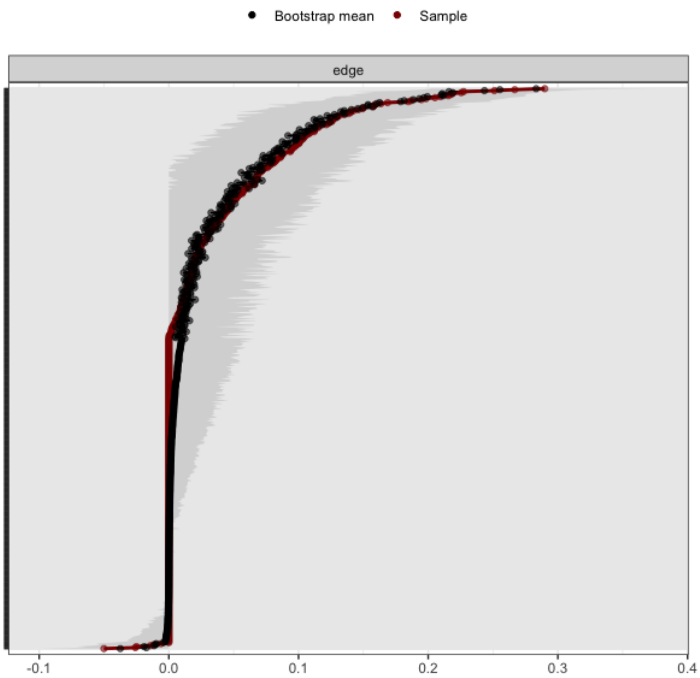


Figure 17: Stability of centrality measures and edge weights for depression, meta-cognition and worry network

In the interpretation of the results, unlike the single measure GlASSO networks, the most central nodes or strongest edges were not the interest but the bridge nodes which connect 2 communities were. Therefore, bridge centrality measures were considered in the analysis. Due to the lower stability of the bridge-betweenness and the bridge-closeness along with the lack of red edge number and sizes which results in bridge-expected influence being almost identical to bridge strength, only the bridge-strength measure was examined (Figure 18). "(M1)Lack of cognitive confidence", (M4)Negative beliefs about uncontrollability and danger", "(M5)Need to control thoughts", and "(W2)My worries overwhelm me" nodes showed the highest bridge strength among the 34 nodes, meaning that they are highly connected to the depression community.

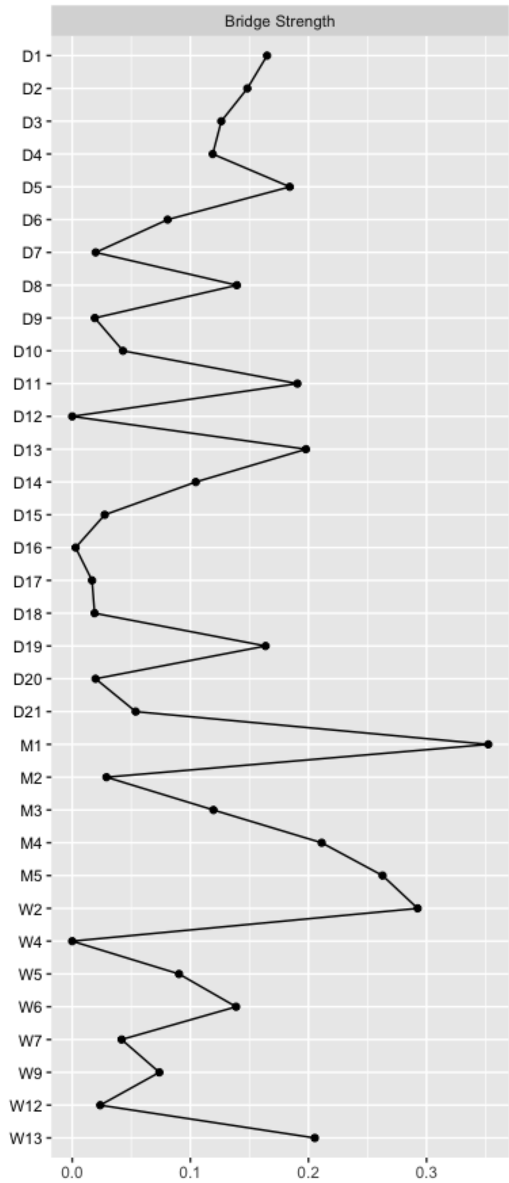


Figure 18: z-scored Bridge strength centrality metrics for Depression, Worry & Meta-worry.z-scored centrality metrics (betweenness, closeness, expected-influence, strength) for each anxiety symptom

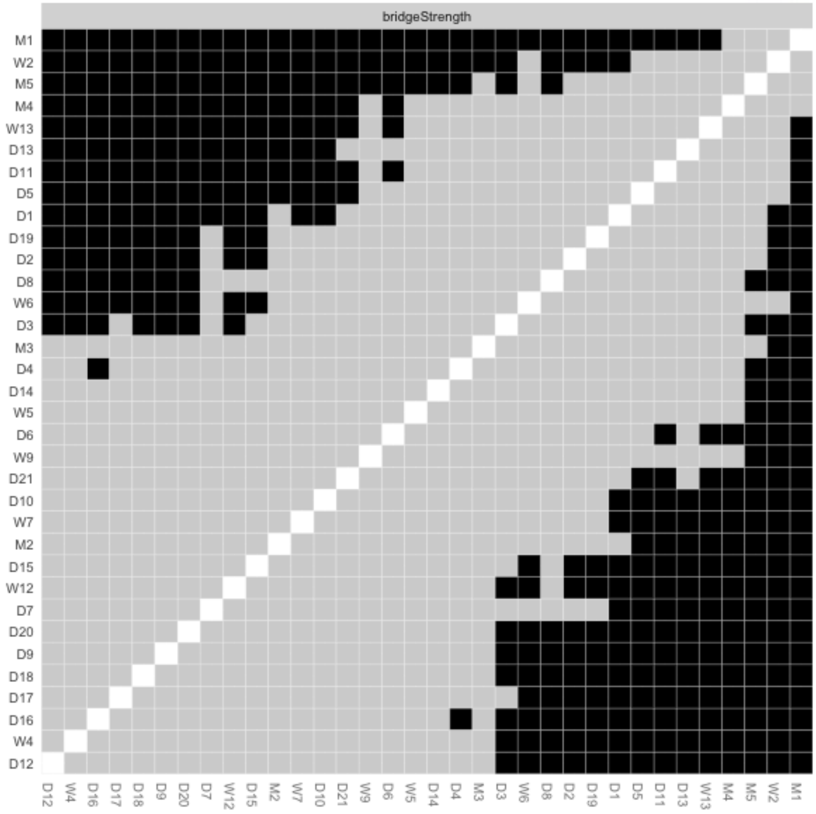


Figure 19: Difference between nodes for depression, meta-cognition and worry network

Correlation matrix between all nodes was checked to detect which node in the depression community they are most connected to (Figure 20). "(M1)Lack of cognitive confidence" node is connected to "(D13)Indecisiveness" while "(M5)Need to control thoughts" node is connected to "(D8)Self-criticalness" symptom in the depression community. "(M4)Negative beliefs about uncontrollability and danger" and "(W2)My worries overwhelm me" nodes both connect to "(D11)Agitation" symptom.

	M1	M4	M5	W2
D1	0	0	0	0.06779894
D2	0.01849938	0.00696849	0	0.06242686
D3	0	0.00655822	0.00515741	0.00170568
D4	0.0187209	0.02539228	0	0.00222748
D5	0.01209628	0	0.02357604	0.05350525
D6	0.02025417	0	0.06052976	0
D7	0.00277185	0.00543758	0	0
D8	0	0.0287339	0.08348412	0.02704947
D9	0.01910629	0	0	0
D10	0	0.01882985	0	0
D11	0	0.04253101	0.00895319	0.07776066
D12	0	0	0	0
D13	0.09498449	0	0	0
D14	0.04400267	0	0.03475022	0
D15	0	0.01584114	0	0
D16	0	0	0	0
D17	0	0.01676745	0	0
D18	0.00496862	0.01398683	0	0
D19	0.09407817	0.03018371	0.02372389	0
D20	0.00924307	0	0	0
D21	0.01355891	0	0.02260194	0

Figure 20: Correlation matrix between nodes for depression, meta-cognition and worry network

In the second estimated network, anxiety, worry, and meta-cognition nodes/items were included(Figure 21). Centrality stability coefficients calculated by the case dropping boot for betweenness being 0.13 was very low making this centrality measure not interpretable. Strength and expected influence scored 0.67 and closeness scored 0.52 that considered in the good range for stability analysis. Contrarily, bridge-closeness scored lowest level tested being 0.05 and making it impossible to use for interpretation. Others fell into the stably good category; bridge-strength and bridge-expected influence with scores of 0.67 and bridge-betweenness with a score of 0.52.(Figure 22). It also shows excellent edge weight stability that was plotted through the non-parametric bootstrap in the same figure.

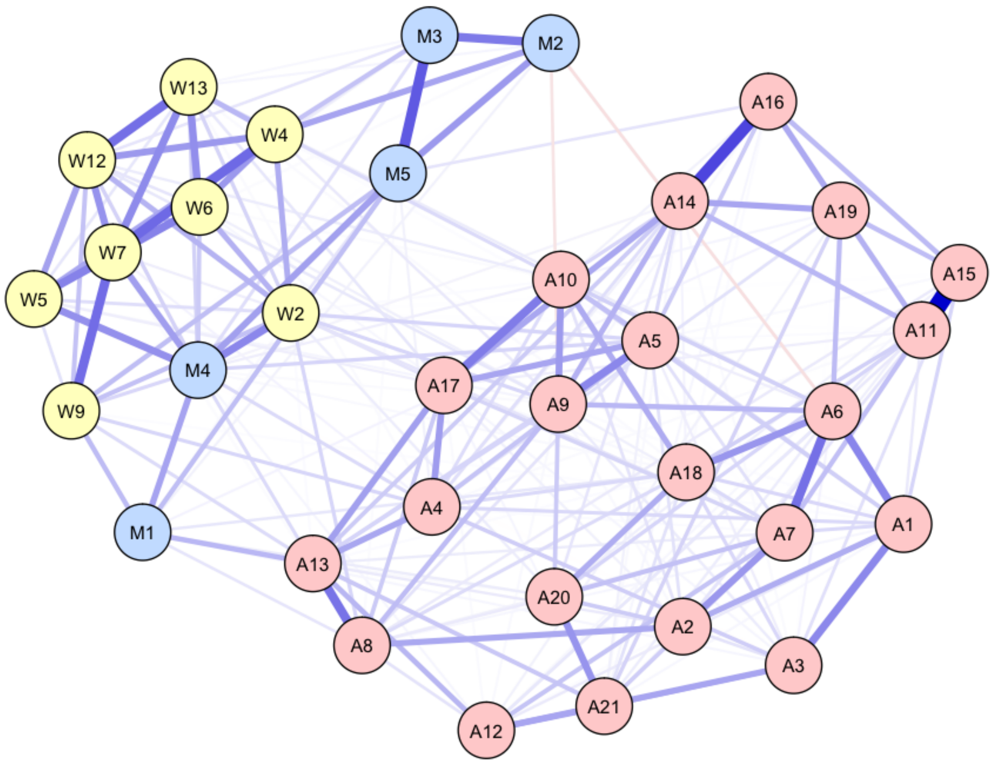


Figure 21: Regularized partial correlation network depicting associations between anxiety symptoms, meta-cognition subsections and worry items.

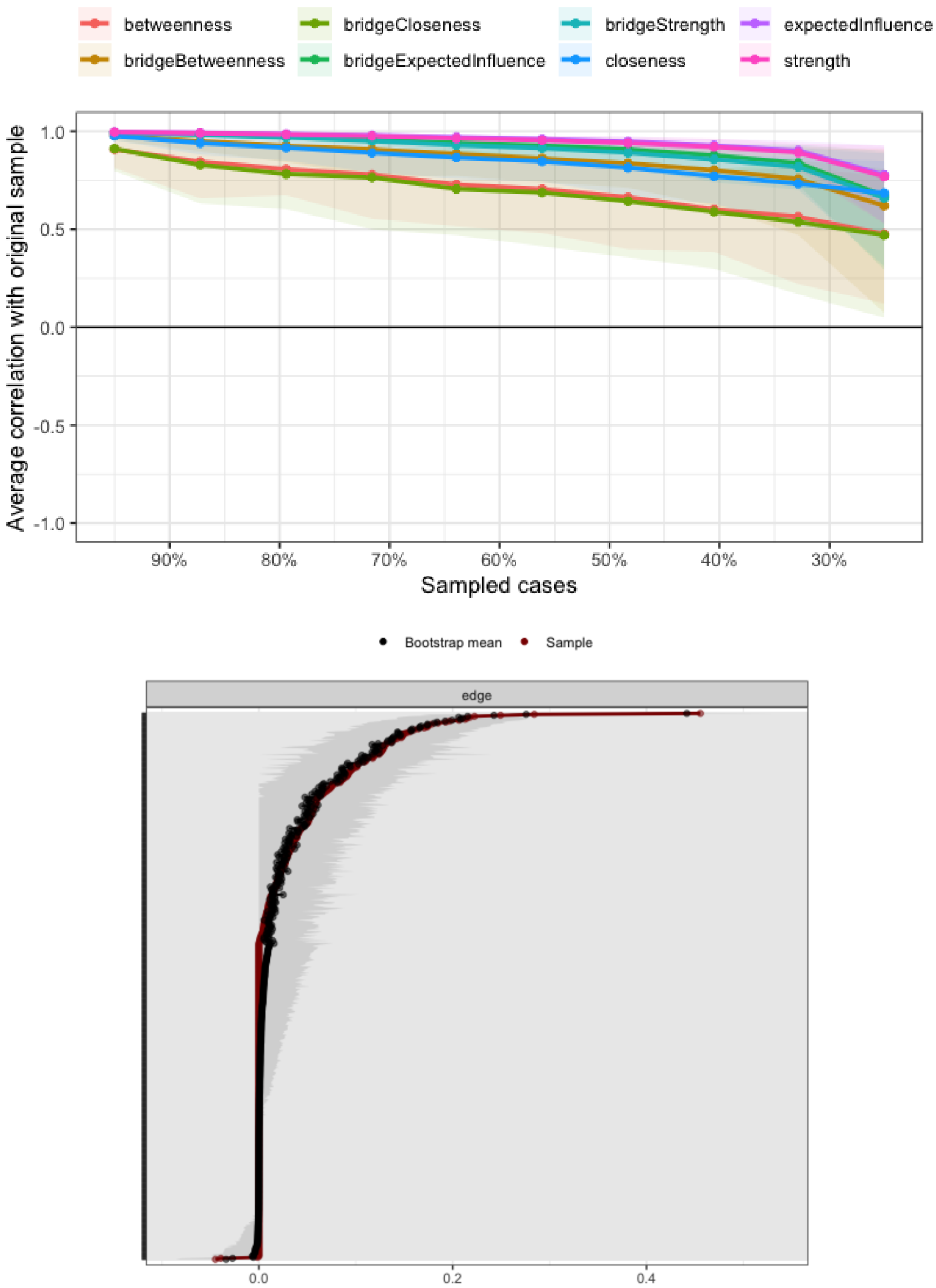


Figure 22: Stability of centrality measures and edge weights for anxiety, meta-cognition and worry network

Similarly, only the bridge-strength measure was examined (Figure 23). "(M1)Lack of cognitive confidence", (M4)Negative beliefs about uncontrollability and danger", and "(W2)My worries overwhelm me" nodes showed the highest bridge strength from the worry&meta-worry community, while "(A9)Terrified or afraid" and "(A13)Shaky / unsteady" nodes from the anxiety community came up as having highest bridge strength.

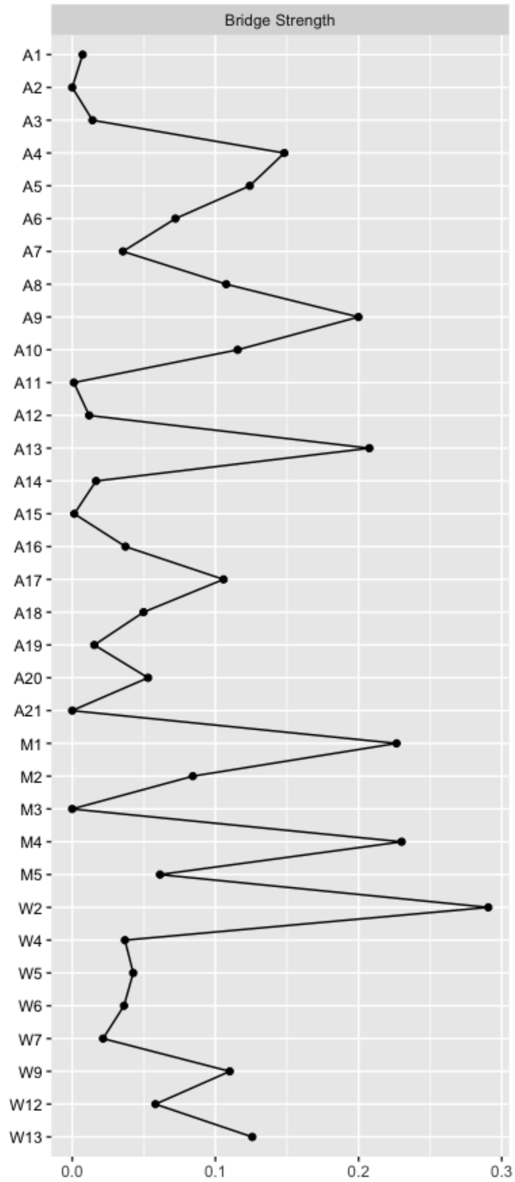


Figure 23: Bridge strength and its difference between nodes for anxiety, meta-cognition and worry network

The correlation matrix between all nodes was checked to detect M1, M4, and W2 nodes' strongest connections to the anxiety community nodes(Figure 25). "(M1)Lack of cognitive confidence" node and "(A13)Shaky / unsteady" node are connected each other while "(M4)Negative beliefs about uncontrollability and danger" node's connections spread out to "(A4)Unable to relax", "(A5)Fear of worst

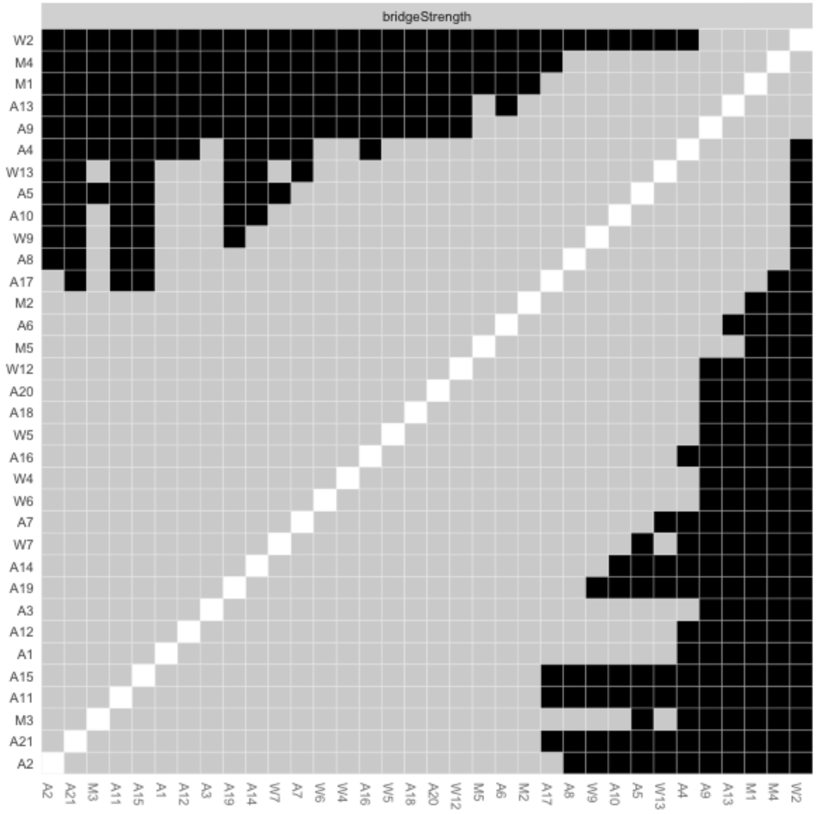


Figure 24: Difference between nodes for anxiety, meta-cognition and worry network

happening" and "(A13)Shaky / unsteady" while " "(W2)My worries overwhelm me" node's connections spread out to "(A5)Fear of worst happening", "(A8)Unsteady", and "(A9)Terrified or afraid".

To sum up, "(M1)Lack of cognitive confidence" and "(M4)Negative beliefs about uncontrollability and danger" subcategories of the meta-cognition questionnaire and "(W2)My worries overwhelm me" item of the Penn-state worry questionnaire are co-morbid with both depression and anxiety symptoms while "(M5)Need to control thoughts" subsection of the meta-cognition questionnaire is co-morbid with depression symptoms. However, any bridge nodes from the depression side of the network were not highly central nodes. Contrarily M4 and M1 were connected to A13 which has high centrality among anxiety symptoms.

	M1	M4	W2
A1	0	0	0.00427624
A2	0	0	0
A3	0.01265568	0	0
A4	0	**0.05164914**	0
A5	0	**0.05849079**	**0.05908392**
A6	0	0	0
A7	0	0	0.03541389
A8	0.03575485	0.00221761	**0.05765003**
A9	0	0.00696343	**0.05717519**
A10	0	0.02256742	0.0051101
A11	0	0	0.0011498
A12	0	0.0118172	0
A13	**0.101229**	**0.0512262**	0.00457919
A14	0	0.01663632	0
A15	0	0	0.00143804
A16	0	0	0
A17	0	0	0.0393284
A18	0.03386273	0.0085001	0
A19	0.01543523	0	0
A20	0.02758901	0	0.02536515
A21	0	0	0

Figure 25: Correlation matrix between nodes for anxiety, meta-cognition and worry network

4.3 Directed Acyclic Graph (DAG)

In the DAG which only covers depression symptoms, 0.85 was used as a treshold for the cutpoint of edge strength, and 0.5 cutpoint was used as directionality strength (Figure 26). Appearing on top, "(D7)Self-dislike" has the possibility of being the initiator of the depression symptoms. It was followed by "(D3)Past failure" and "(D14)Worthlessness" and "(D8)Self-criticalness". In the GLASSO network of the depression symptoms, D7 and D14 had high centrality as well, making them important targets for intervention. When checked for directional strength between nodes, edges between top nodes did not have strong directionality (Figure 27).So, it is better to assess them together as intervention targets. The DAG ended with "(D10)Crying", "(D13)Indecisiveness" nodes, and "(D16)Changes in sleeping pattern", "(D18)Changes in appetite", "(D21)Loss of interest in sex" chain nodes.

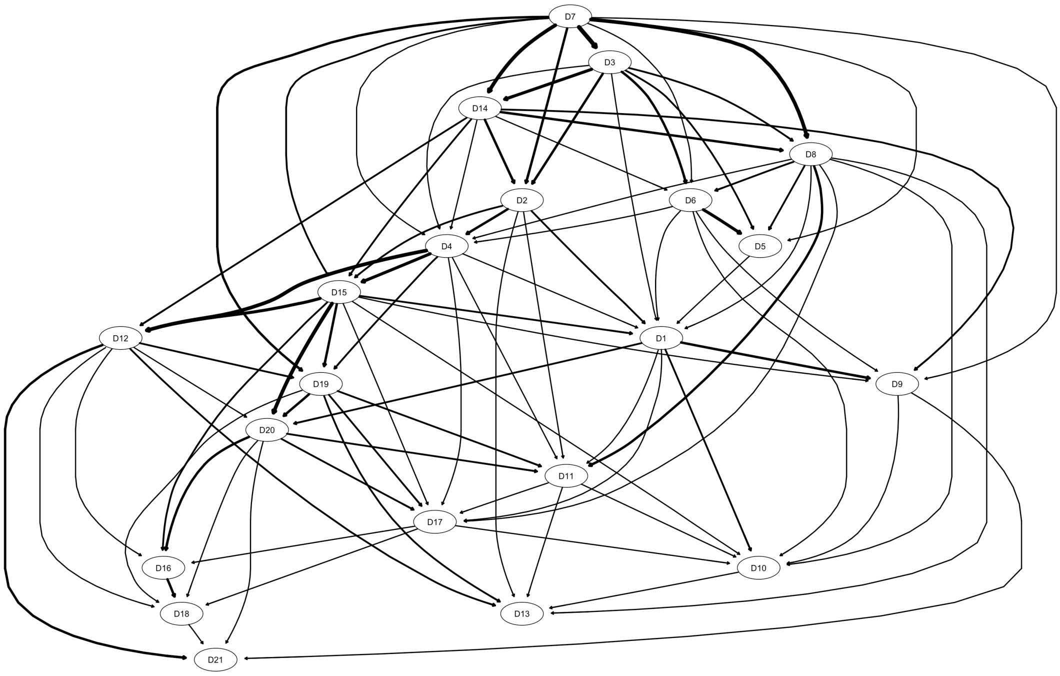


Figure 26: Directed acyclic graph (DAG) by using the %85 cutpoint. Edges signify the importance (BIC value) of the edge to model fit (depression).

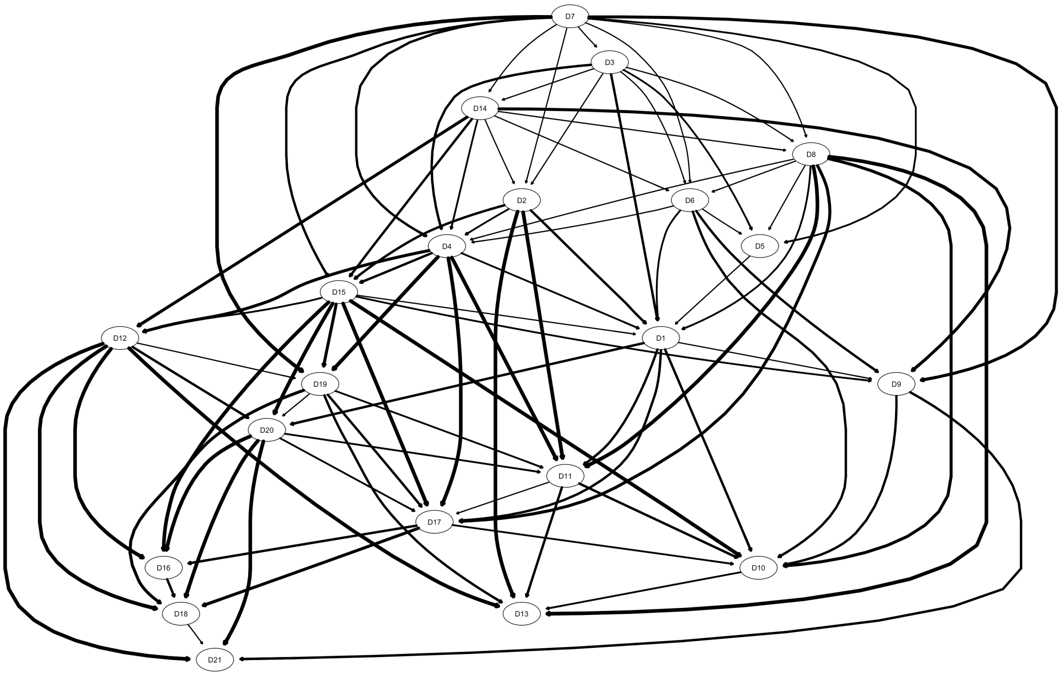


Figure 27: Direction probabilities for edge width. Thick arrows indicate high directional probabilities, thin arrows low directional probabilities(depression).

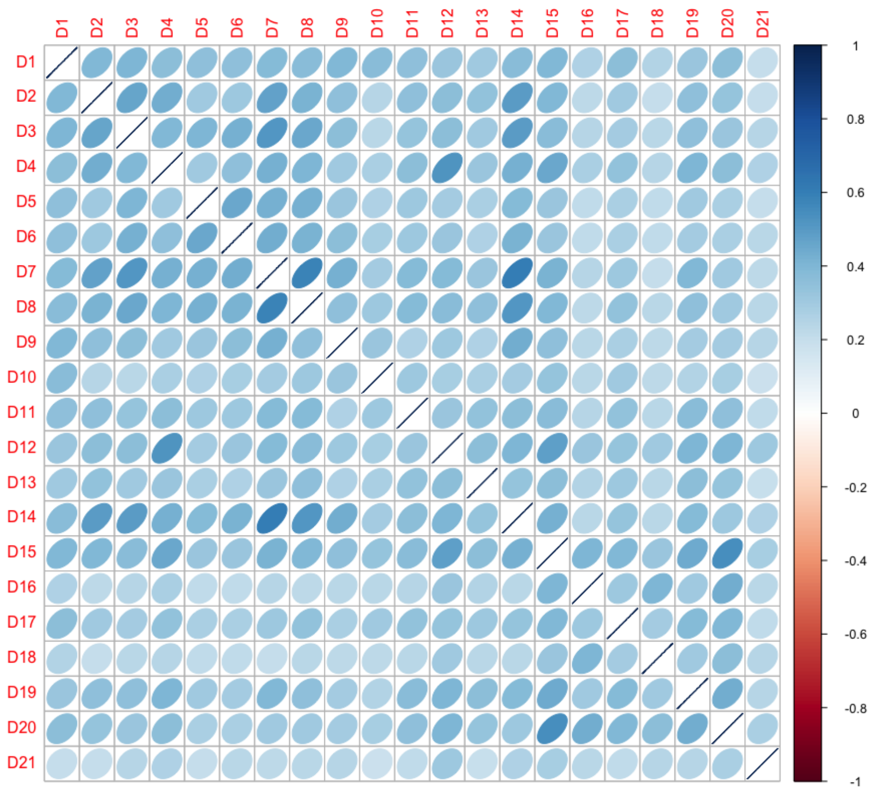


Figure 28: Correlation matrix for depression symptoms.

Similarly, in the DAG which only covers anxiety symptoms, 0.85 was used as the cutpoint for edge strength, and 0.5 cutpoint was used as directionality strength (Figure 29). "(A17)Scared" appeared on top followed by "(A10)Nervous" and "(A5)Fear of worst happening" making them a focal point for intervention targets. In the GLASSO network, A17 and A10 had the greatest strength between all the anxiety symptoms as well. Like DAG with depression symptoms, the directional strength between nodes that appeared at the top did not have strong directionality (Figure 30). So, it is better to assess them together as intervention targets. The anxiety DAG ended with "(A2)Feeling hot" on one side and, "(A15)Difficulty in breathing", "(A19)Faint / lightheaded", "(A21)Hot / cold sweats chain nodes on the other side.

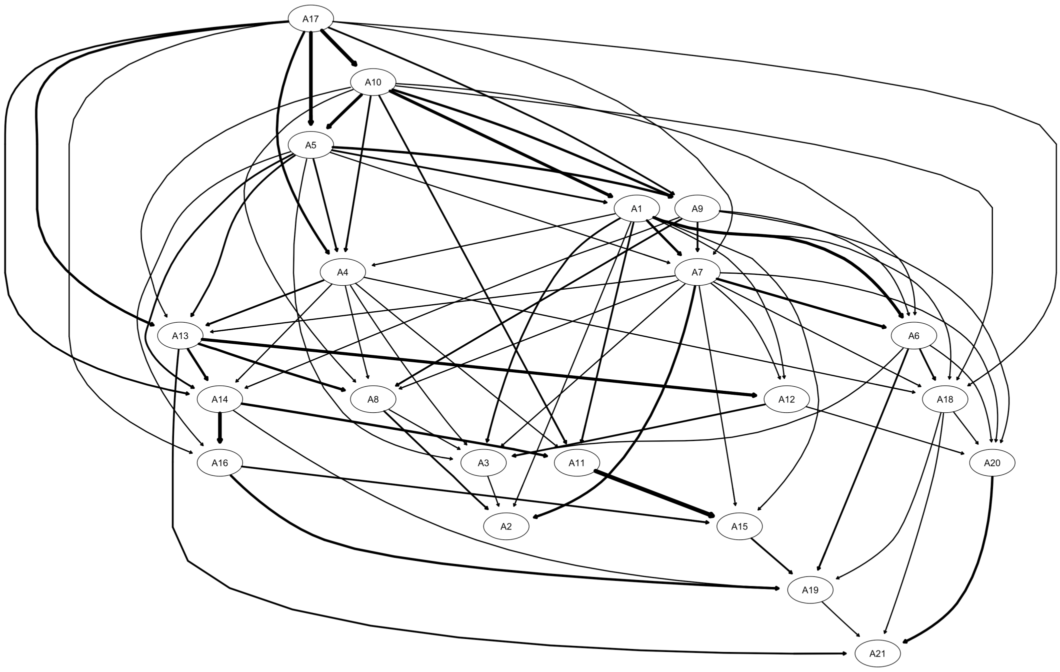


Figure 29: Directed acyclic graph (DAG) by using the 0.85 cutpoint. Edges signify the importance (BIC value) of the edge to model fit (anxiety).

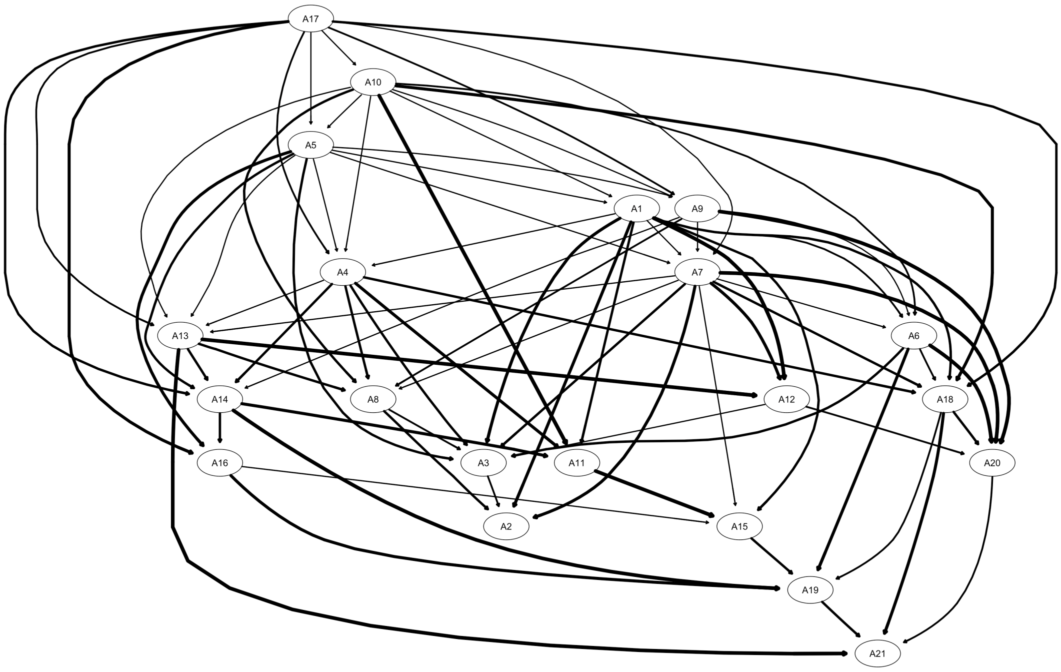


Figure 30: Direction probabilities for edge width. Thick arrows indicate high directional probabilities, thin arrows low directional probabilities(anxiety).

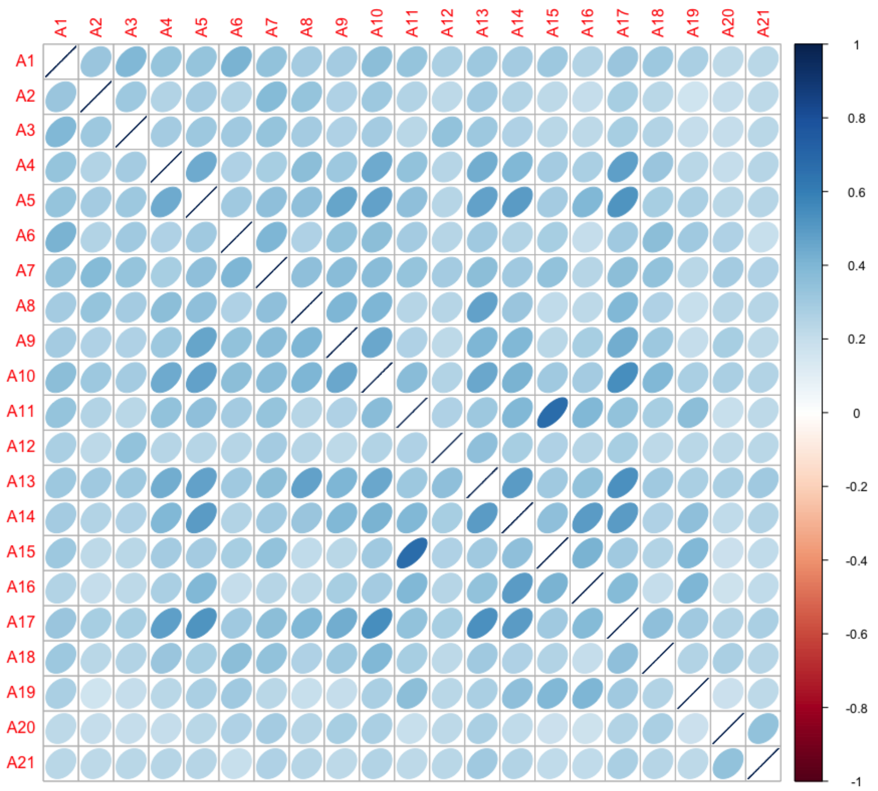


Figure 31: Correlation matrix for anxiety symptoms.

When looking for the causality between worry & meta-cognition community and depression or anxiety symptoms, firstly, the optimal threshold was calculated by using the algorithm provided by (put ref) for edge strength cutpoint, and 0.5 cutpoint was used as directionality strength.

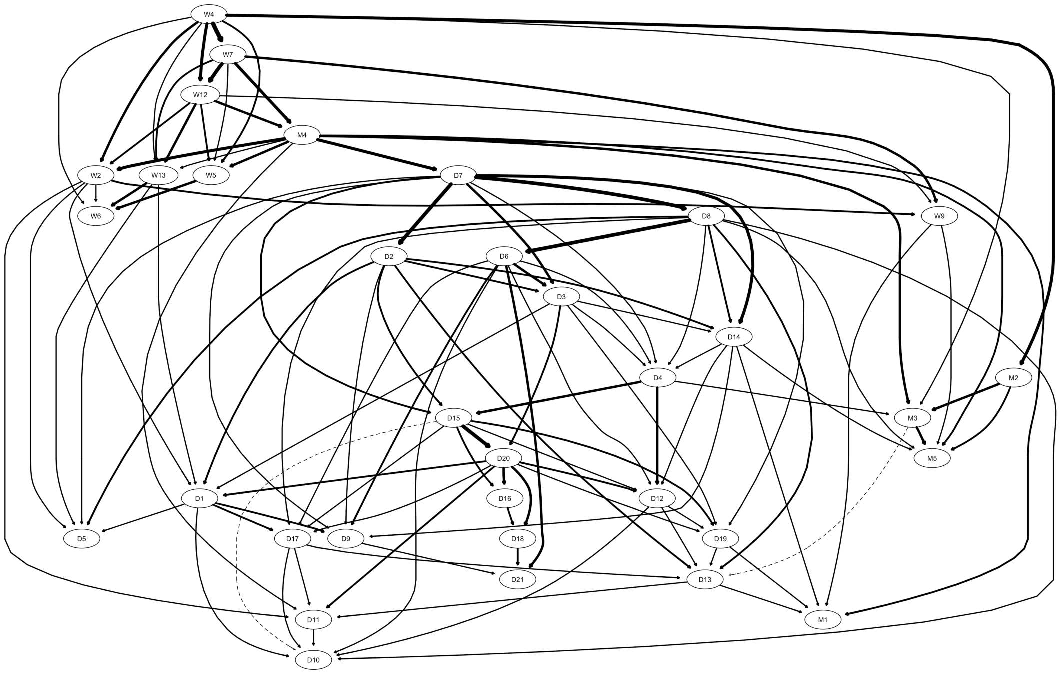


Figure 32: Directed acyclic graph (DAG) by using the optimal cutpoint. Edges signify the importance (BIC value) of the edge to model fit (depression, worry, meta-cognition).

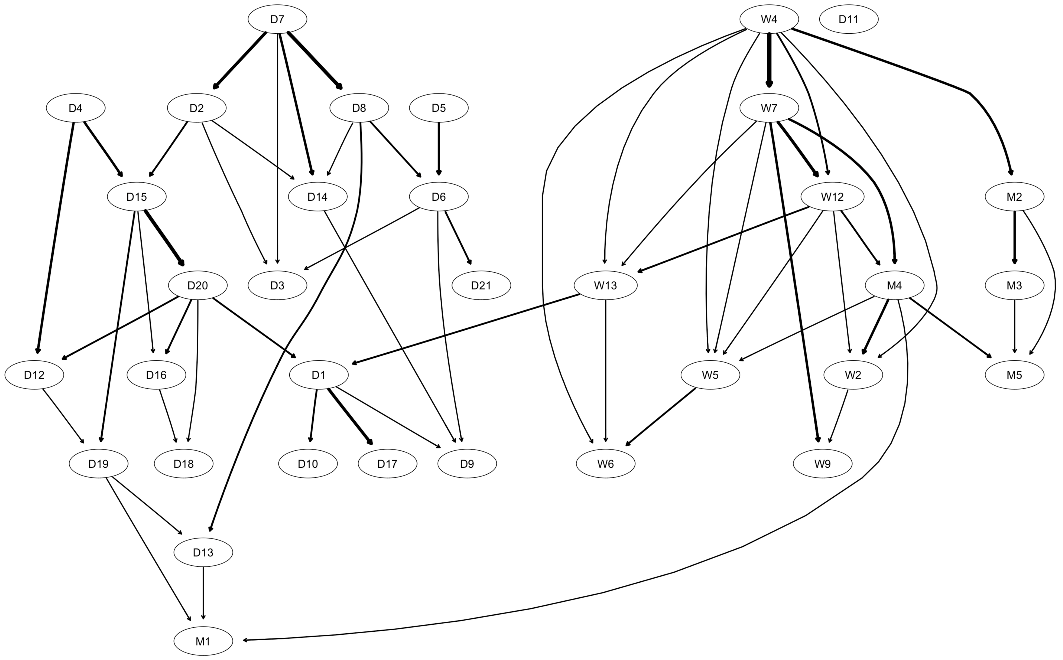


Figure 33: Directed acyclic graph (DAG) by using the 0.85 cutpoint. Edges signify the importance (BIC value) of the edge to model fit (depression, worry, meta-cognition).

In the figure 32, "(M4)Negative beliefs about uncontrollability and danger" bridge symptom for the depression appeared on top of all depression symptoms, meaning it can be a contributor for depression symptoms. "(M1)(Lack of) cognitive confidence" and "(M5)Need to control thoughts" bridge symptoms appeared at the bottom meaning they are most likely a result. Meanwhile, "(W2)My worries overwhelm me" only contributed to the lower branches of the depression symptoms. However, when the cutpoint was raised to 0.85 for the edges, the interaction between the two communities mostly disappeared leaving only "(M1)(Lack of) cognitive confidence" as an end result of the depression symptoms (Figure 33).

In the figure 34, "(M4)Negative beliefs about uncontrollability and danger" and "(W2)My worries overwhelm me" bridge symptoms for anxiety appeared to be contributors to the top anxiety symptoms. Meanwhile, "(M5)Need to control thoughts" only contributed to the lower branches of the anxiety symptoms. However, when the cutpoint was raised to 0.85 for the edges, again, the interaction between the two communities completely disappeared making two different islands of causal relationships (Figure 35). Directional probability graphs of last 4 figures can be found at the supplementary material at the end.

In summary, "(M4)Negative beliefs about uncontrollability and danger" for both depression and anxiety, and "(W2)My worries overwhelm me" for only anxiety could give us a red flag for students who scored high on those items as indicators of future mental health problems.

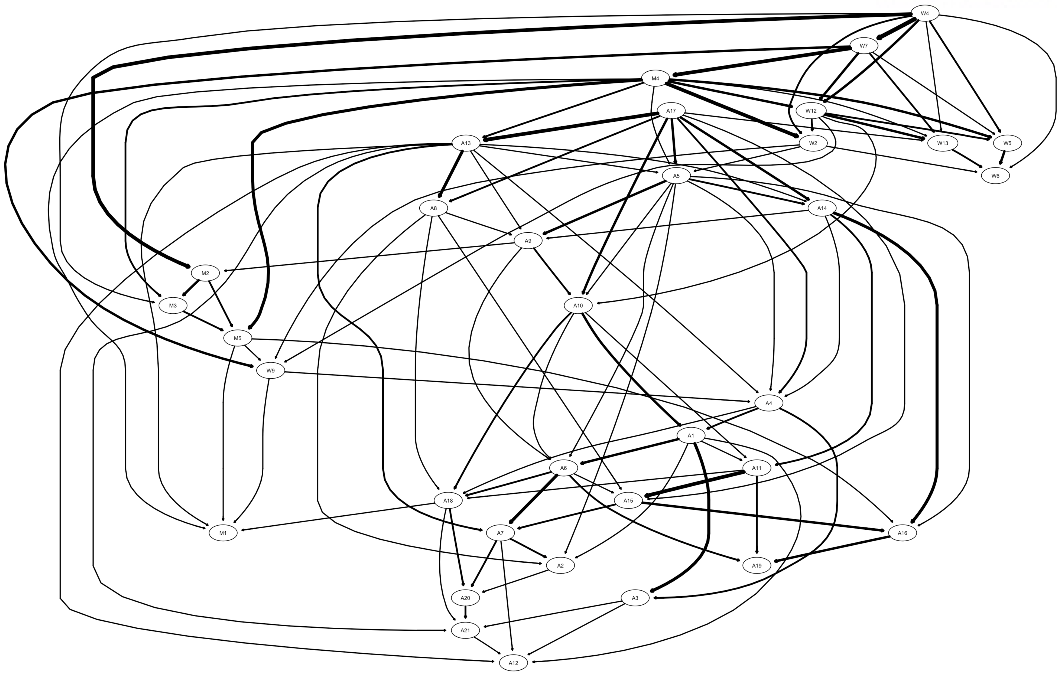


Figure 34: Directed acyclic graph (DAG) by using the optimal cutpoint. Edges signify the importance (BIC value) of the edge to model fit (anxiety, worry, meta-cognition).

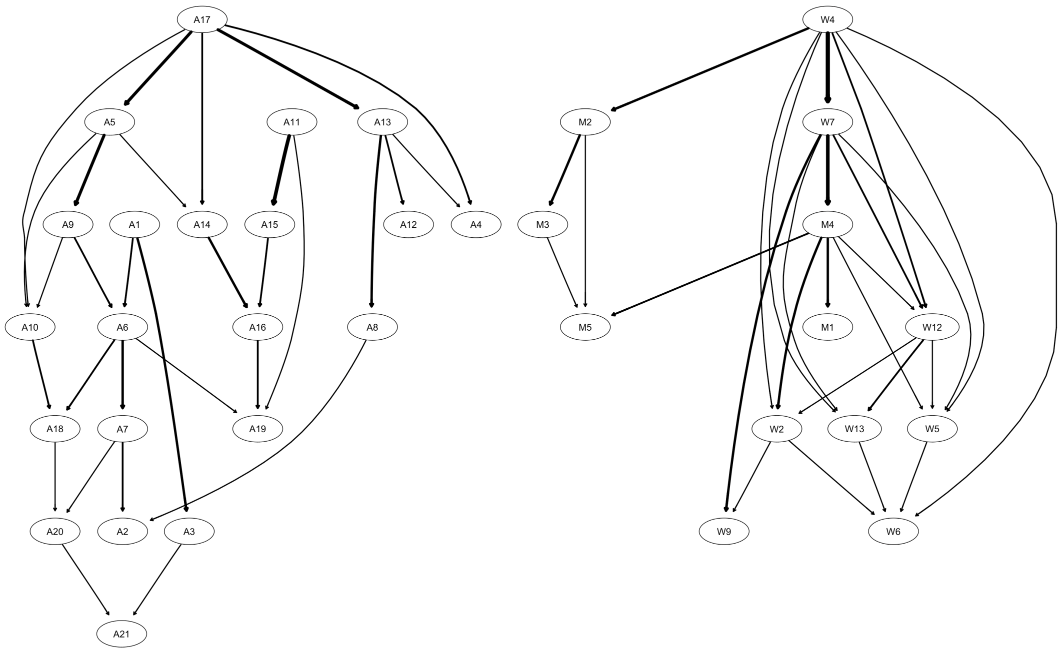


Figure 35: Directed acyclic graph (DAG) by using the 0.85 cutpoint. Edges signify the importance (BIC value) of the edge to model fit (anxiety, worry, meta-cognition).

V Conclusion

In this study, descriptive statistic regarding the depression, anxiety, worry and meta-worry items were revealed. Factor analysis done by previous studies were compered with the network analysis results only based on appearance . Although it is insufficient to draw conclusions, the goal was to suggest and encourage further research. The main focus of this thesis was to explore four aims explained in the objective of the research at top.

The first two aims of the study were achieved successfully by a detailed analysis of relationships between depression and anxiety symptoms. For depression, "Self-dislike" and "Loss of energy" symptoms came up as strongest nodes followed by "Worthlessness" and "Tiredness or fatigue" symptoms emerged as most central symptoms. Strength was the only centrality measure to give significant results. Also, "Self-dislike", "Worthlessness" along with "Past failure" was suggested by DAG as an initiator of other depression symptoms. For anxiety, "Shaky / unsteady", and "Hearth pounding / racing" symptoms had the greatest node strength betweenness. "Nervous" and "Scared" nodes were only greatest in the strength centrality while "Fear of losing control" were only greatest in the betweenness centrality. Meanwhile, the "Dizzy or lightheaded" node was greatest at both betweenness and closeness measures. "Nervous" and "Scared" together with "Fear of worst happening" comes up as initiator symptoms of anxiety in the DAG. Central symptoms are substantial targets for intervention since they are highly connected to the symptoms of those around them. Additionally, initiator symptoms can be considered to come up with preventive methods for mental health disorders. Another strategy for intervention can be developed by considering highly connected nodes together.

The last two goals of the study were achieved as well. From 8 PSWQ(worry) items, the "My worries overwhelm me" item showed comorbidity with both depression and anxiety. When its strongest connection to the symptoms was checked, they did not appear to be one of the central symptoms. When causality was investigated with optimal cutpoint DAG, it had an impact on the initiator symptoms of anxiety, while only contributing to the lower branches of the depression symptoms. However, when the cutpoint was raised to 85%, no ties remained. From 5 subscales of MCQ(meta-worry), "Lack of cognitive confidence" and "Negative beliefs about uncontrollability and danger" showed comorbidity with both depression and anxiety. Their strongest connections to anxiety symptoms included one of the central symptoms of anxiety, namely "Shaky / Unsteady". "Need to control thoughts" subscale only showed comorbidity with depression symptoms. However, non of the bridge symptoms connected to the central symptoms of depression strongly. When causality was investigated with optimal cutpoint DAG, "Negative beliefs about uncontrollability and danger" appeared on top for both depression and anxiety, meaning, higher scores on this subscale can be initiating factors for both mental health disorders. Contrarily, "Lack of cognitive confidence" appeared as an end result in both directed networks, meaning, high scores in this subscale could be an indicator of already developed mental health problems. However, when the cutpoint was raised to 85%, only the "Lack of cognitive confidence" subscale's relation with depression remained.

Whether or not they made it to the final network model, all 4 items from worry and meta-worry concepts showing comorbidity with depression and anxiety have importance by giving a warning sign of existing mental health disorders or potential hazard for the future. Them, not making it into the final model means they may not be as important to include them when designing a prevention or intervention method for depression and anxiety. All findings of this study give suggestions for intervention and prevention methods of depression and anxiety disorders by detecting the most important symptoms of each. However, these findings need to be tested and proven to be a helpful and effective way of intervention or prevention technique by conducting further clinical research.

This study's biggest strength was high number of participants which is rare in network analysis. But it was achieved by collecting over long period of time which in return caused some weaknesses in the study too. The period of time the data were collected, historical, economical, or sociological events that may potentially affect the mental health of the given population, and the effect of recently occurred pandemic. Unfortunately, timestamps were not included in the data collection process. Therefore, observing the mental health change during the time-period was not possible in this study.

VI Codes – R scripts

My conclusion will be here

```
# Install required packages..
library(bootnet)
library(networktools)
library(qgraph)
library("e1071")
library("igraph")
library("dplyr")
## Bayesian network package
library("bnlearn")
## used for visualizing Bayesian networks
library("Rgraphviz")

# Area proportional euler graph.
hope <- c("Depression" = 5015, "Anxiety" = 5430,
          "Meta-Worry" = 3589, "Worry" = 1761,
          "Depression&Anxiety" = 2157, "Depression&Meta-Worry" = 3584,
          "Depression&Worry" = 1761, "Anxiety&Meta-Worry" = 1729,
          "Anxiety&Worry" = 1761, "Meta-Worry&Worry" = 743,
          "Depression&Anxiety&Meta-Worry" = 1729,
          "Depression&Anxiety&Worry" = 1761,
          "Depression&Meta-Worry&Worry" = 743,
          "Anxiety&Meta-Worry&Worry" = 743,
          "Depression&Anxiety&Meta-Worry&Worry" = 743)

# fitting..
fita <- euler(hope, input = "union", shape = "ellipse")

#####
#This code was run 4 times by importing different datasets.
#####



# Import data
mydata <- read.table("~/DAWmetaW.csv", header=TRUE,
                     sep=",", na = "NA")
```

```
# Check the data
summary(mydata)
head(mydata)
dim(mydata)


# Running GoldBricker to find out if there are nodes sharing similar fate!
GBmydata <- goldbricker(mydata, p = 0.05, method = "pearson1898",
            threshold = 0.25,corMin = 0.5, progressbar = TRUE)

# If found remove them from the dataset..
Red_mydata <- net_reduce(data=mydata, badpairs=GBmydata)


#color palets, used different colors for each measure
#and kept it consistent in different networks.
color = c('#A8E6CF', '#ffd2d2','#cbe2ff', '#fffec8')


mynetwork <- estimateNetwork(mydata, default="EBICglasso",
            corMethod = "cor", corArgs = list(method = "spearman",
            use = "pairwise.complete.obs"))


#Change groups and colors!

png("mynetwork_anx.png", width = 1024, height = 768, units='px',
            res = 1600, pointsize = 10)
myplot <- plot(mynetwork, layout="spring", vsize=6,
            groups=list("depp"=c(1:21),"meta-worry"=c(22:26),
            "worry"=c(26:33)), color=c('#A8E6CF', '#DCEDC1', '#FDC705'),
            border.color="black", legend=FALSE)
dev.off()

#Check communities!

#Estimate bridge values for each node
```

```
#Name our bridge object
mybridge <- bridge(myplot,
           communities=c( '1','1','1','1','1','1','1',
                          '1','1','1','1','1','1','1',
                          '1','1','1','1','1','1','1',
                          '2','2','2','2','2','2','2',
                          '2','2','2','2','2','2'),
           useCommunities = "all", directed = NULL, nodes = NULL)

#Estimate bridge values for each node
mybridge

write(mybridge, "Centrality_bridge.txt")

#lets plot
plot(mybridge, include = c("Bridge Strength",
           "Bridge Expected Influence (1-step)",
           "Bridge Closeness", "Bridge Betweenness"))

#Save centrality values
CentralityTable <- centralityTable(mynetwork)
write.csv(CentralityTable, "CentralityTable_depp.csv")


#Centrality plot
centralityPlot(myplot,
           include = c("Strength", "Closeness",
           "Betweenness", "ExpectedInfluence"))


#Constructing a partial correlation matrix
N1edges <- getWmat(mynetwork)
write.csv(N1edges, "NetworkEdges_depp.csv")

#Estimating Network Stability & Accuracy
#Include communities only when bridge analysis is needed.


caseDroppingBoot <- bootnet(mynetwork, boots=10000, type="case",
```

```
            statistics=c("strength","closeness", "betweenness",
            "expectedInfluence", "edge","bridgeStrength",
            "bridgeExpectedInfluence", "bridgeCloseness",
            "bridgeBetweenness"),
            communities=c( '1','1','1','1','1','1','1',
                           '1','1','1','1','1','1','1',
                           '1','1','1','1','1','1','1',
                           '2','2','2','2','2','2','2',
                           '2','2','2','2','2','2'),
            useCommunities = "all")


nonParametricBoot <- bootnet(mynetwork, boots=10000,
            type="nonparametric",
            statistics=c("strength","closeness", "betweenness",
            "expectedInfluence", "edge","bridgeStrengt",
            "bridgeExpectedInfluence", "bridgeCloseness",
            "bridgeBetweenness"),
            communities=c( '1','1','1','1','1','1','1',
                           '1','1','1','1','1','1','1',
                           '1','1','1','1','1','1','1',
                           '2','2','2','2','2','2','2',
                           '2','2','2','2','2','2'),
            useCommunities = "all")


#get stability coefficients
mycorstability <- corStability(caseDroppingBoot)
write(mycorstability, "corstability_dmw.txt")

#Plot centrality stability
plot(caseDroppingBoot, statistics=c("strength", "closeness",
            "betweenness", "expectedInfluence",
            ))
plot(caseDroppingBoot, statistics=c("bridgeStrength",
            "bridgeExpectedInfluence"))

#Plot centrality difference
plot(nonParametricBoot, statistics="bridgeStrength",
```

```
          plot="difference", order = "sample")
plot(nonParametricBoot, statistics="strength", plot="difference",
          order = "sample")

# Above command can be used for other centrality measures as well.

#Save edge stability graph
plot(nonParametricBoot, labels = FALSE, order = "sample")

#Edge weights diff test
plot(nonParametricBoot, "edge", plot = "difference",
          onlyNonZero = TRUE, order = "sample")

# AIM 2: Estimate Bayesian network
## Fit a first Bayesian network,
#based on 50 random re-starts and
#100 perturbations for each re-start.

## convert to numerics
netdata <- as.data.frame(apply(mydata, 2, as.numeric))

# Quick correlation structure check
cormat <- cor(netdata)
corrplot(cormat, method = "ellipse")

set.seed(123)

## hc gives directed graph
fitBN1 <- hc(netdata, restart = 50, perturb = 100)
fitBN1

## global network score
bnlearn::score(fitBN1, data = netdata)

## connection strength
astr <- arc.strength(fitBN1, netdata, "bic-g")

## sorted edge strength from strongest to weakest
astr[order(astr[,3]), ]
```

```r
strength.plot(fitBN1, astr, shape = "ellipse")

## Now we stabilize the network across multiple
#samples through bootstrapping:
## Learn 10000 network structures
#(takes ~5 min, we keep the number of restarts
#and perturbations considerably low)

set.seed(123)
bootnet <- boot.strength(netdata, R = 10000,
             algorithm = "hc",
             algorithm.args = list(restart = 5, perturb = 10),
             debug = TRUE)

head(bootnet)

## filter the ones with a strength larger than 0.85
#and a direction probability > 0.5
bootnet[bootnet$strength > 0.85 & bootnet$direction > 0.5, ]

## Net1: build the average network using a 0.85 threshold
avgnet1 <- averaged.network(bootnet, threshold = 0.85)
avgnet1
bnlearn::score(avgnet1, data = netdata)

## compute edge strengths
astr1 <- arc.strength(avgnet1, netdata, "bic-g")
strength.plot(avgnet1, astr1, shape = "ellipse")

# Net2: use the optimal cutpoint,
#according to Scurati & Nagarajan (2013)
avgnet2 <- averaged.network(bootnet)
avgnet2
bnlearn::score(avgnet2, data = netdata)
thresh <- avgnet2$learning$args[[1]]
thresh

## optimal significance threshold
```

```
## compute edge strengths
astr2 <- arc.strength(avgnet2, netdata, "bic-g")
astr2

suppressWarnings(strength.plot(avgnet2, astr2,
          shape = "ellipse", threshold = 0.5))

## Net3: use net2 threshold, edge strenghts
#are determined by direction probability
## edges in net2
boottab <- bootnet[bootnet$strength > thresh &
          bootnet$direction > 0.5, ]
boottab

## table with direction probabilities
astr3 <- boottab

## use the direction probabilities for edge width
astr3$strength <- astr3$direction

strength.plot(avgnet2, astr3, shape = "ellipse")

## Net4: use net1 threshold, edge strenghts
#are determined by direction probability
## edges in net2
boottab <- bootnet[bootnet$strength > 0.85 &
bootnet$direction > 0.5, ]
boottab

## table with direction probabilities
astr4 <- boottab

## use the direction probabilities for edge width
astr4$strength <- astr4$direction

strength.plot(avgnet1, astr4, shape = "ellipse")
```

VII Appendix

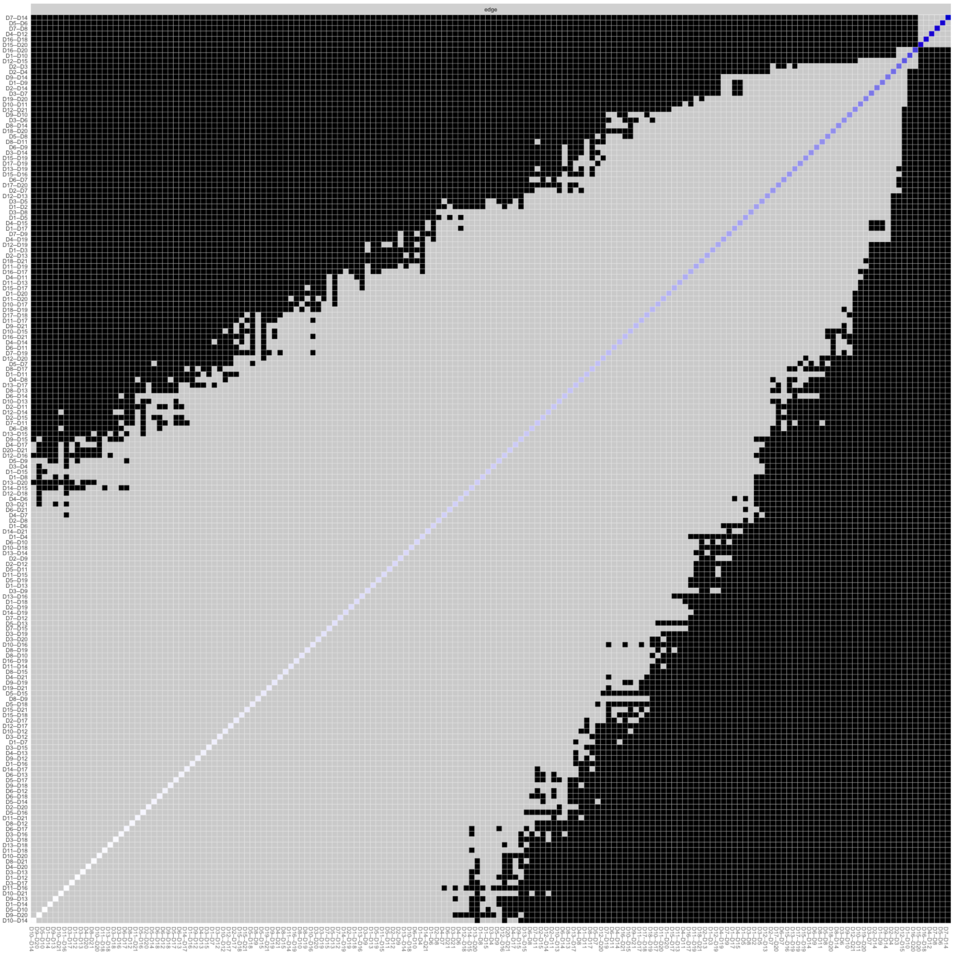


Figure 36: Edge weight differences between depression nodes. Black boxes indicate significant difference while grey boxes means significantly not different.

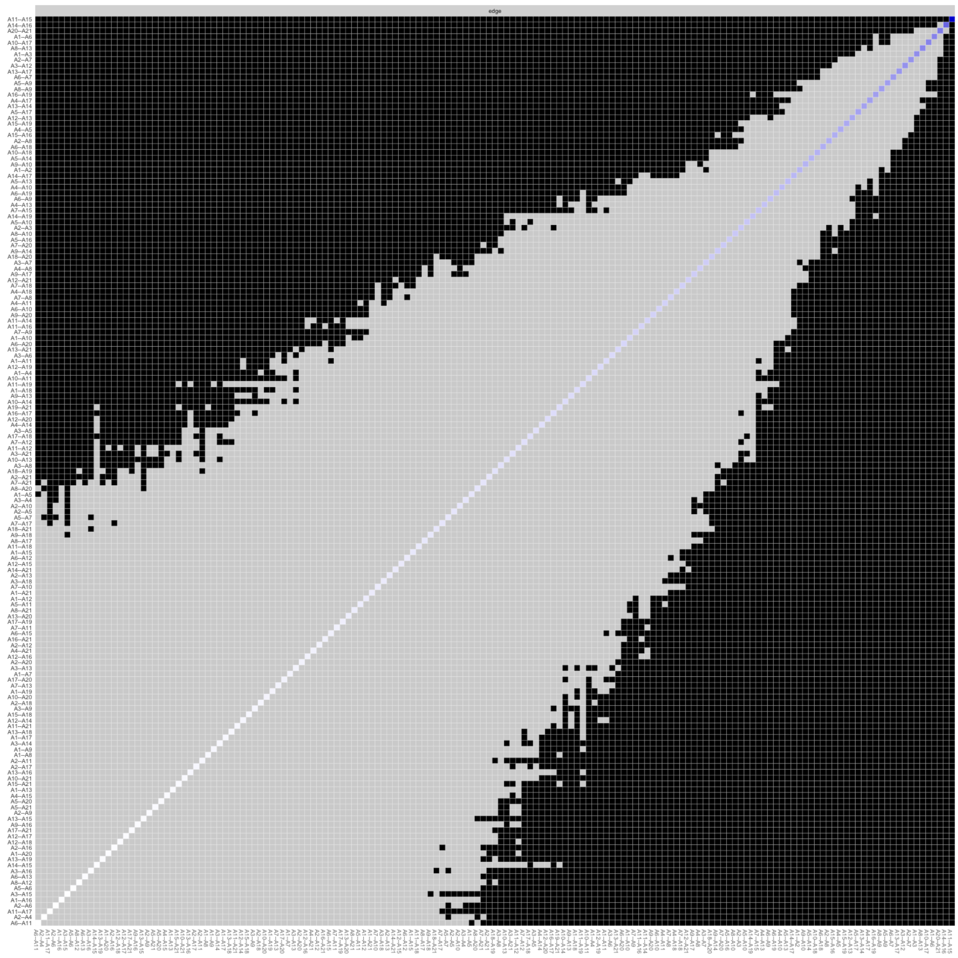


Figure 37: Edge weight differences between depression nodes. Black boxes indicate significant difference while grey boxes means significantly not different.

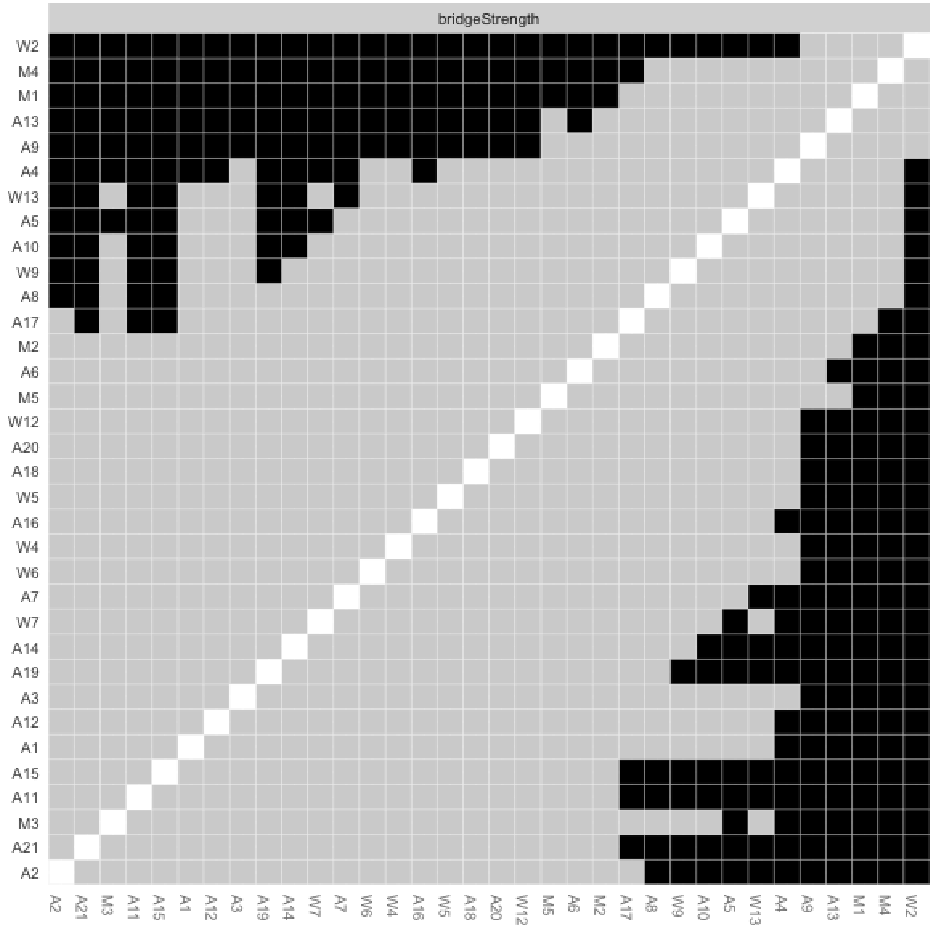


Figure 38: Bridge strength and its difference between nodes for anxiety, meta-cognition and worry network

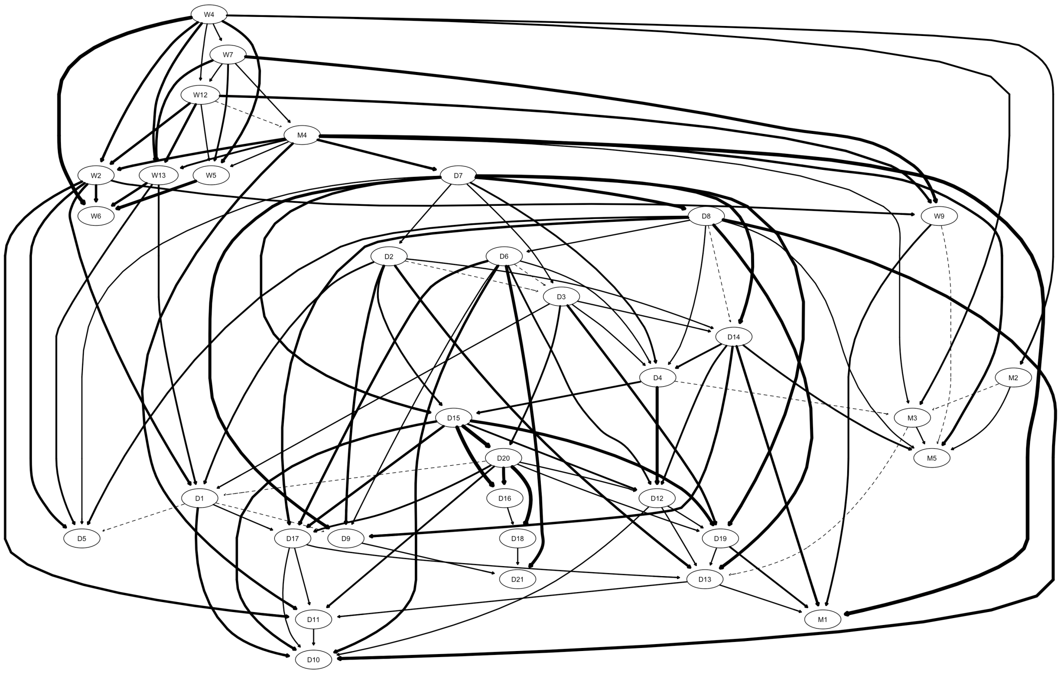


Figure 39: Direction probabilities for edge width. Thick arrows indicate high directional probabilities, thin arrows low directional probabilities(depression, worry, meta-cognition).

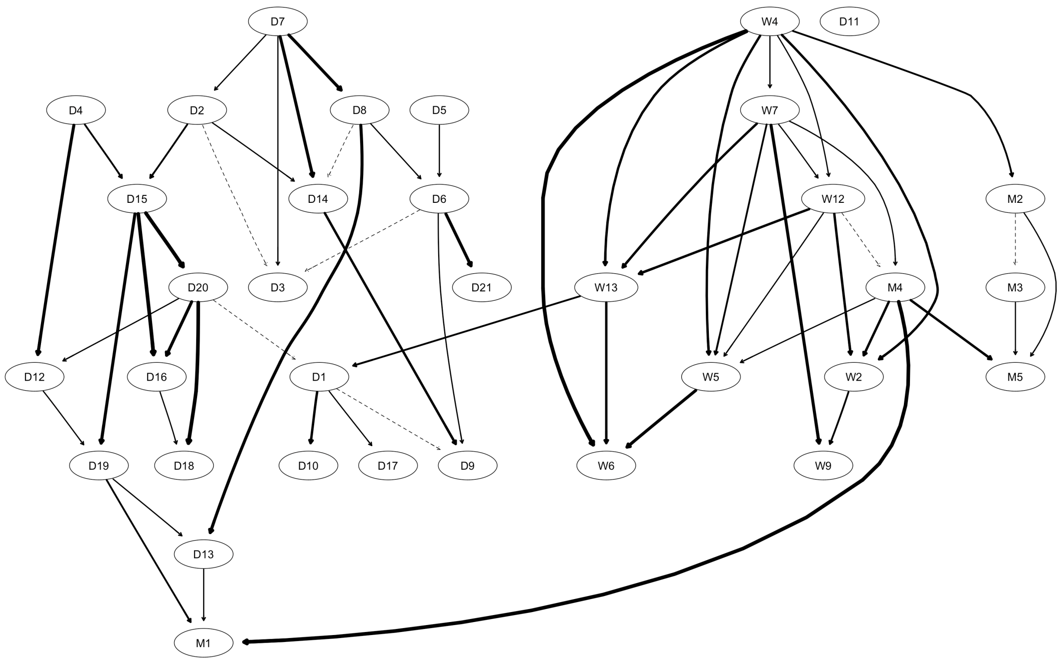


Figure 40: Direction probabilities for edge width. Thick arrows indicate high directional probabilities, thin arrows low directional probabilities(depression, worry, meta-cognition).

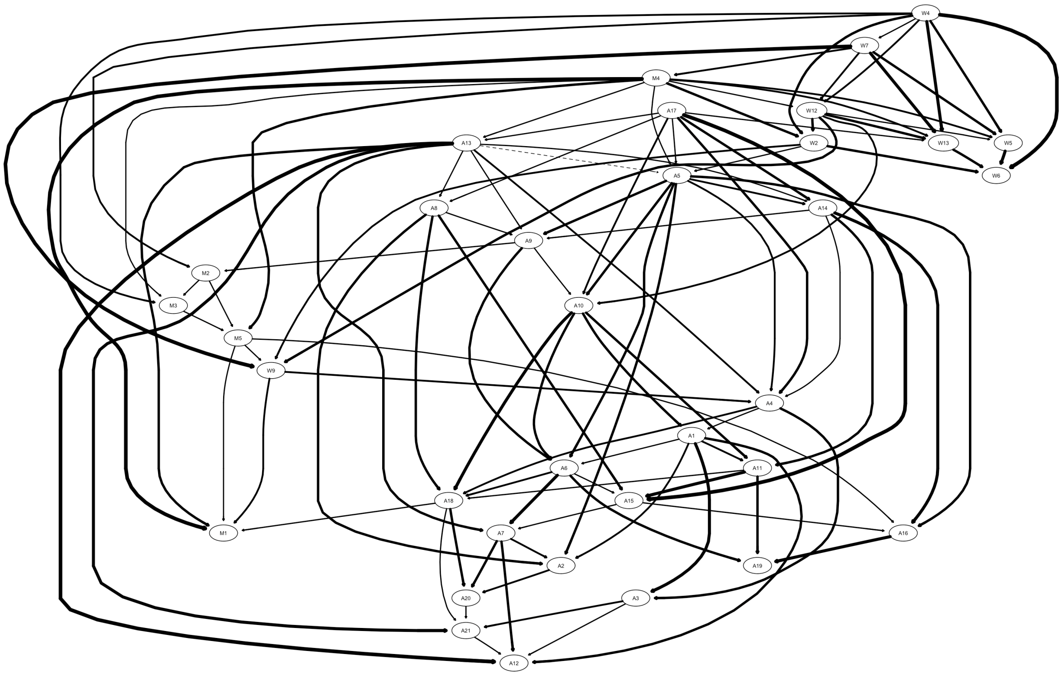


Figure 41: Direction probabilities for edge width. Thick arrows indicate high directional probabilities, thin arrows low directional probabilities(depression, worry, meta-cognition).

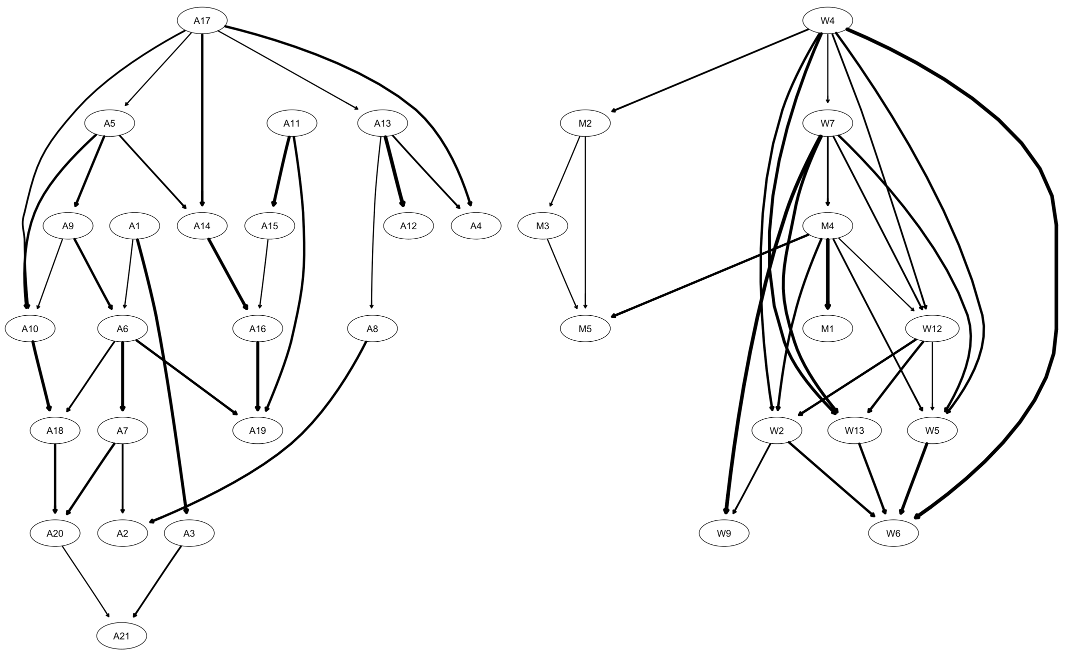


Figure 42: Direction probabilities for edge width. Thick arrows indicate high directional probabilities, thin arrows low directional probabilities(depression, worry, meta-cognition).

References

[1] D. Borsboom and A. O. Cramer, "Network analysis: An integrative approach to the structure of psychopathology," *Annual Review of Clinical Psychology*, vol. 9, pp. 91–121, 2013.

[2] V. D. Schmittmann, A. O. Cramer, L. J. Waldorp, S. Epskamp, R. A. Kievit, and D. Borsboom, "Deconstructing the construct: A network perspective on psychological phenomena," *New Ideas in Psychology*, vol. 31, no. 1, pp. 43–53, 2013. [Online]. Available: http://dx.doi.org/10.1016/j.newideapsych.2011.02.007

[3] S. Epskamp, D. Borsboom, and E. I. Fried, "Estimating psychological networks and their accuracy: A tutorial paper," *Behavior Research Methods*, vol. 50, no. 1, pp. 195–212, 2018.

[4] S. Wasserman, "Social network analysis : methods and applications," *Structural analysis in the social sciences ; 8*, 1994.

[5] N. Meinshausen and P. Bühlmann, "High-dimensional graphs and variable selection with the Lasso," *Annals of Statistics*, vol. 34, no. 3, pp. 1436–1462, 2006.

[6] J. Friedman, T. Hastie, and R. Tibshirani, "Sparse inverse covariance estimation with the graphical lasso," *Biostatistics*, vol. 9, no. 3, pp. 432–441, 2008.

[7] T. Hastie, "Statistical learning with sparsity : the lasso and generalizations," Boca Raton, 2015.

[8] R. Foygel and M. Drton, "Extended Bayesian information criteria for Gaussian graphical models," *Advances in Neural Information Processing Systems 23: 24th Annual Conference on Neural Information Processing Systems 2010, NIPS 2010*, pp. 1–14, 2010.

[9] A. O. Cramer, L. J. Waldorp, H. L. Van Der Maas, and D. Borsboom, "Comorbidity: A network perspective," *Behavioral and Brain Sciences*, vol. 33, no. 2-3, pp. 137–150, 2010.

[10] L. Boschloo, C. D. Van Borkulo, M. Rhemtulla, K. M. Keyes, D. Borsboom, and R. A. Schoevers, "The network structure of symptoms of the diagnostic and statistical manual of mental disorders," *PLoS ONE*, vol. 10, no. 9, pp. 1–12, 2015.

[11] E. I. Fried, C. Bockting, R. Arjadi, D. Borsboom, M. Amshoff, A. O. Cramer, S. Epskamp, F. Tuerlinckx, D. Carr, and M. Stroebe, "From loss to loneliness: The relationship between bereavement and depressive symptoms," *Journal of Abnormal Psychology*, vol. 124, no. 2, pp. 256–265, 2015.

[12] R. J. McNally, D. J. Robinaugh, G. W. Wu, L. Wang, M. K. Deserno, and D. Borsboom, "Mental disorders as causal systems: A network approach to posttraumatic stress disorder," *Clinical Psychological Science*, vol. 3, no. 6, pp. 836–849, 2015.

[13] K. T. Forbush, C. S. Siew, and M. S. Vitevitch, "Application of network analysis to identify interactive systems of eating disorder psychopathology," *Psychological Medicine*, vol. 46, no. 12, pp. 2667–2677, 2016.

[14] A. M. Isvoranu, D. Borsboom, J. Van Os, and S. Guloksuz, "A network approach to environmental impact in psychotic disorder: Brief theoretical framework," *Schizophrenia Bulletin*, vol. 42, no. 4, pp. 870–873, 2016.

[15] A. M. Isvoranu, C. D. Van Borkulo, L. L. Boyette, J. T. Wigman, C. H. Vinkers, D. Borsboom, R. Kahn, L. De Haan, J. Van Os, D. Wiersma, R. Bruggeman, W. Cahn, C. Meijer, and I. Myin-Germeys, "A network approach to psychosis: Pathways between childhood trauma and psychotic symptoms," *Schizophrenia Bulletin*, vol. 43, no. 1, pp. 187–196, 2017.

[16] C. Van Borkulo, L. Boschloo, D. Borsboom, B. W. Penninx, J. W. Lourens, and R. A. Schoevers, "Association of symptom network structure with the course of longitudinal depression," *JAMA Psychiatry*, vol. 72, no. 12, pp. 1219–1226, 2015.

[17] G. Costantini, S. Epskamp, D. Borsboom, M. Perugini, R. Mõttus, L. J. Waldorp, and A. O. Cramer, "State of the aRt personality research: A tutorial on network analysis of personality data in R," *Journal of Research in Personality*, vol. 54, pp. 13–29, 2015. [Online]. Available: http://dx.doi.org/10.1016/j.jrp.2014.07.003

[18] J. Dalege, D. Borsboom, F. Van Harreveld, H. Van den Berg, M. Conner, and H. L. Van der Maas, "Toward a formalized account of attitudes: The Causal Attitude Network (CAN) Model," *Psychological Review*, vol. 123, no. 1, pp. 2–22, 2016.

[19] J. J. Kossakowski, S. Epskamp, J. M. Kieffer, C. D. van Borkulo, M. Rhemtulla, and D. Borsboom, "The application of a network approach to Health-Related Quality of Life (HRQoL): introducing a new method for assessing HRQoL in healthy adults and cancer patients," *Quality of Life Research*, vol. 25, no. 4, pp. 781–792, 2016.

[20] K. J. Kan, H. De Jonge, H. L. Van Der Maas, S. Z. Levine, and S. Epskamp, "How to compare psychometric factor and network models," *Journal of Intelligence*, vol. 8, no. 4, pp. 1–10, 2020.

[21] R. van Bork, M. Rhemtulla, L. J. Waldorp, J. Kruis, S. Rezvanifar, and D. Borsboom, "Latent Variable Models and Networks: Statistical Equivalence and Testability," *Multivariate Behavioral Research*, vol. 56, no. 2, pp. 175–198, 2021. [Online]. Available: https://doi.org/10.1080/00273171.2019.1672515

[22] H. Guyon, B. Falissard, and J. L. Kop, "Modeling psychological attributes in psychology - An epistemological discussion: Network analysis vs. latent variables," *Frontiers in Psychology*, vol. 8, no. MAY, pp. 1–10, 2017.

[23] S. Gruijters, "Using principal component analysis to validate psychological scales: Bad statistical habits we should have broken yesterday II," *European health Psychologist*, vol. 20, pp. 544–549, 2020.

[24] L. Ren, Z. Yang, Y. Wang, L. B. Cui, Y. Jin, Z. Ma, Q. Zhang, Z. Wu, H. N. Wang, and Q. Yang, "The relations among worry, meta-worry, intolerance of uncertainty and attentional bias for threat in men at high risk for generalized anxiety disorder: A network analysis," *BMC Psychiatry*, vol. 20, no. 1, pp. 1–11, 2020.

[25] A. T. Beck, R. A. Steer, and G. Brown, "Beck depression inventory–ii," *Psychological Assessment*, 1996.

[26] K. Park, E. Jaekal, S. Yoon, S. H. Lee, and K. H. Choi, "Diagnostic Utility and Psychometric Properties of the Beck Depression Inventory-II Among Korean Adults," *Frontiers in Psychology*, vol. 10, no. January, pp. 1–10, 2020.

[27] Y.-m. Song, H.-k. Lee, and J. W. Kim, "Validation and factor structure of korean version of the beck depression inventory second edition (bdi-ii) : In a university student sample," *Korean journal of biological psychiatry*, pp. 402–408, 2012.

[28] Y.-M. Song, H.-k. Lee, and K. Lee, "Reliability and validity of the korean version of beck depression inventory-ii via the internet : Results from a university student sample," *Korean journal of biological psychiatry*, vol. 18, no. 3, pp. 126–133, 2011.

[29] A. T. Beck and R. Steer, "Beck anxiety inventory (bai)," *Überblick über Reliabilitäts-und Validitätsbefunde von klinischen und außerklinischen Selbst-und Fremdbeurteilungsverfahren*, vol. 7, 1988.

[30] Han-Kyeong Lee, Jihae Kim, SanghwangHong, Eun-ho Lee, and Hwang Soon Taeg, "Psychometric Properties of the Beck Anxiety Inventory in the Community-dwelling Sample of Korean Adults," *Korean Journal of Clinical Psychology*, vol. 35, no. 4, pp. 822–830, 2016.

[31] Y. Kim and Y. Jun, "The Role of Positive Meta-cognition and Negative Meta-cognition in Employment Stress among College Students," *International Conference on Studies in Humanities and Social Sciences*, pp. 91–95, 2015.

[32] T. J. Meyer, M. L. Miller, R. L. Metzger, and T. D. Borkovec, "Development and validation of the penn state worry questionnaire," *Behaviour research and therapy*, vol. 28, no. 6, pp. 487–495, 1990.

[33] Y. J. Lim, Y. H. Kim, E. H. Lee, and S. M. Kwon, "The Penn State worry questionnaire: Psychometric properties of the Korean version," *Depression and Anxiety*, vol. 25, no. 10, pp. 97–103, 2008.

[34] J. Cohen, P. Cohen, S. West, and L. Aiken, "Applied multiple regression/correlation analysis for the behavioral sciences," 2003.

[35] S. L. Lauritzen, *Graphical models*. Clarendon Press, 1996, vol. 17.

[36] N. Meinshausen and P. Bühlmann, "High-dimensional graphs and variable selection with the lasso," *The annals of statistics*, vol. 34, no. 3, pp. 1436–1462, 2006.

[37] M. Pourahmadi, "Covariance estimation: The glm and regularization perspectives," *Statistical Science*, vol. 26, no. 3, pp. 369–387, 2011.

[38] S. Epskamp, A. O. J. Cramer, L. J. Waldorp, V. D. Schmittmann, and D. Borsboom, "Qgraph: Network visualizations of relationships in psychometric data," *Journal of statistical software*, vol. 48, no. 4, pp. 1–18, 2012.

[39] C. Beard, A. J. Millner, M. J. Forgeard, E. I. Fried, K. J. Hsu, M. T. Treadway, C. V. Leonard, S. J. Kertz, and T. Björgvinsson, "Network analysis of depression and anxiety symptom relationships in a psychiatric sample," *Psychological Medicine*, vol. 46, no. 16, pp. 3359–3369, 2016.

[40] P. J. Jones, A. Heeren, and R. J. McNally, "Commentary: A network theory of mental disorders," *Frontiers in Psychology*, vol. 8, p. 1305, 2017. [Online]. Available: https://www.frontiersin.org/article/10.3389/fpsyg.2017.01305

[41] C. A. Levinson, S. Zerwas, B. Calebs, K. Forbush, H. Kordy, H. Watson, S. Hofmeier, M. Levine, R. D. Crosby, C. Peat, C. D. Runfola, B. Zimmer, M. Moesner, M. D. Marcus, and C. M. Bulik, "The Core Symptoms of Bulimia Nervosa, Anxiety, and Depression: A Network Analysis," *Journal of abnormal psychology (1965)*, vol. 126, no. 3, pp. 340–354, 2017.

[42] R. J. McNally, P. Mair, B. L. Mugno, and B. C. Riemann, "Co-morbid obsessive–compulsive disorder and depression: a bayesian network approach," *Psychological Medicine*, vol. 47, no. 7, p. 1204–1214, 2017.

[43] P. J. Jones, R. Ma, and R. J. McNally, "Bridge Centrality: A Network Approach to Understanding Comorbidity," *Multivariate Behavioral Research*, vol. 0, no. 0, pp. 1–15, 2019. [Online]. Available: https://doi.org/10.1080/00273171.2019.1614898

[44] J. Pearl, *Causality*. Cambridge University Press, 2000.

[45] M. Vignes, M. Sammarro, A. Schrefheere, A. White, and S. Marsland, "Bayesian Networks: With Examples in R. By M. Scutari and J.-B. Denis Boca Raton, Florida CRC Press. 2014. 241 pages. UK £69.99 (hardback). ISBN 978-1-48222-558-7," *Australian & New*

Zealand Journal of Statistics, vol. 59, no. 4, pp. 545–546, 2017. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1111/anzs.12220

[46] R. Nagarajan, M. Scutari, and S. Lèbre, *Introduction*. New York, NY: Springer New York, 2013, pp. 1–12. [Online]. Available: https://doi.org/10.1007/978-1-4614-6446-4_1

Acknowledgements

I would like to express my sincere gratitude to my advisor Prof. Dooyoung Jung for his patience and continuous support. I believe his advices helped me to find a research topic best fits to my background. Besides my advisor, I would like to thank the rest of my thesis committee: Prof. Oh-Sang Kwon, Prof. Dongil Chung, and Dr. Sangil Lee, for their encouragement, insightful comments, and support. I am indeed grateful to my lovely husband, his never-ending support and encouragement gave me the strength i need. I dedicate this thesis to my baby whom we are looking forward to welcome soon. Thanks for cheering me up with your kicks during the stressful time.

ULSAN NATIONAL INSTITUTE OF
SCIENCE AND TECHNOLOGY