New Orleans Jazzed about Engineering Education

Paper ID #16861

Agent-Based Modeling and Simulation of Depression and Its Impact on Student Success and Academic Retention

Sherif Elmeligy Abdelhamid, Virginia Tech

Sherif is a PhD candidate at the Department of Computer Science, Virginia Polytechnic Institute and State University and is a graduate research assistant at Network Dynamics and Simulations Science Laboratory. Sherif's research work lies at the intersection of computation, biology and education: in particular, he is interested in designing and building software systems to enable domain experts to easily access and effectively use high performance computing to perform and share the findings of simulations and large scale data analyses. Other aspects of his research focus on how to use these systems as learning tools for students and teachers.

Dr. Chris J. Kuhlman, Virginia Bioinformatics Institute

Chris is a Research Scientist at the Biocomplexity Institute at Virginia Tech. His research interests include discrete dynamical systems, agent-based modeling and simulation, distributed and high performance computing, algorithms, social sciences and modeling, and network science.

Prof. Madhav V. Marathe, Virginia Tech

Madhav Marathe is the director of the Network Dynamics and Simulation Science Laboratory and professor in the department of computer science, Virginia Tech. His research interests are in computational epidemiology, network science, design and analysis of algorithms, computational complexity, communication networks and high performance computing. Before coming to Virginia Tech, he was a Team Leader in the Computer and Computational Sciences division at the Los Alamos National Laboratory (LANL) where he led the basic research programs in foundations of computing and high performance simulation science for analyzing extremely large socio-technical and critical infrastructure systems. He is a Fellow of the IEEE, ACM and AAAS.

Prof. S. S. Ravi, University at Albany - SUNY

Ravi received his Ph.D. in Computer Science in 1984 and joined the Computer Science faculty at the University at Albany – State University of New York. His current title is DIstinguished Teaching Professor. His areas of interest include algorithms, discrete dynamical systems, data mining, network science and wireless networks.

Dr. Kenneth Reid, Virginia Tech

Kenneth Reid is the Assistant Department Head for Undergraduate Programs in Engineering Education at Virginia Tech. He is active in engineering within K-12, serving on the TSA Board of Directors. He and his coauthors were awarded the William Elgin Wickenden award for 2014, recognizing the best paper in the Journal of Engineering Education. He was awarded an IEEE-USA Professional Achievement Award in 2013 for designing the nation's first BS degree in Engineering Education. He was named NETI Faculty Fellow for 2013-2014, and the Herbert F. Alter Chair of Engineering (Ohio Northern University) in 2010. His research interests include success in first-year engineering, engineering in K-12, introducing entrepreneurship into engineering, and international service and engineering. He has written two texts in Digital Electronics, including the text used by Project Lead the Way.

Agent-Based Modeling and Simulation of Depression and Its Impact on Students' Success and Academic Retention

Abstract

In the U.S., major depressive disorder affects approximately 14.8 million American adults. Furthermore, depression can lead to a several other illnesses and disabilities. Economic burden of depression is estimated to be $53 billion annually in the U.S. alone. Depression can reach high levels that can lead to suicide, the third leading cause of death among the U.S. college-aged population.

Studies show a direct relation between mental health and academic success. In particular, depression is a significant predictor of lower GPA and increased drop out rate. A 15 point increase on the depression scale correlates with a 0.17 drop in GPA and corresponds to a 4.7 percent increase in probability of dropping out. High dropout rates also adversely impact both universities and society.

In this work, we construct and exercise an agent-based model (ABM) of the evolution of depression among a population of roughly 19,000 college students. This model includes within-agent interactions among depression symptoms and agent-to-agent interactions defined by a college student social network. We conduct simulation studies to identify (model) parameters and initial conditions that most influence population outcomes. Connectivity among within-agent symptoms is demonstrated to have a large effect on population levels of depression.

Introduction

Background

Dropout rates of engineering students are high in the United States[1,2]. Increased dropout rates have negative impacts on students, institutions and society. From a student's perspective, quitting school can lower self-confidence and self-esteem[3]. From an institutional perspective, student dropouts represent a loss of talented students and indicate an institution's lack of attention to the needs of students[4]. Dropout rates effect society. The cost to the nation of these dropouts amounted to $4.5 billion in lost earnings and taxes to state and federal governments in 2012[5]. Factors contributing to students' dropping out include preparation, ability, motivation,

engagement within the institution, college grade point average, financial aid, age, ethnicity and socioeconomic status[6,7,8,9].

Another important factor in withdrawing from school is side effects resulting from social interactions among students including friends, classmates, and roommates[10,11,12,13]. Depression is contagious[14,15,16,17,18,19,20,21], meaning that it can spread from person to person through different types of social contacts, such as face-to-face interactions and online social media. Studies show a direct relation between mental health and academic success. Students report depression and anxiety among top impediments to academic performance[22]. Sixty four (64) percent of young adults who are no longer in college cite a mental health-related reason for not attending[23]. In this work, we use agent-based modeling (ABM) to simulate the evolution of depression within a synthetic social contact network of undergraduate students at a large university.

Motivation for Agent-based Modeling of Depression

ABM of depression is important for several reasons. First, there has been much data gathered and analyzed since at least the 1970s regarding factors that affect academic retention and attainment. These data can be used to develop and inform social behavior models that can be used in simulations. Second, simulations resolve behaviors in time. This is often critical for understanding causality and how local agent behaviors give rise to population-level outcomes. This is fundamentally different from performing statistical analyses using final outcomes[24]. Third, simulations on appropriately labeled agents that compromise a population (e.g., in the form of social networks) can produce disaggregated results. Validated models can also be used in other settings (e.g. for different academic institutions), and can be exercised to explore counterfactuals. Finally, results from ABMs can inform experimental studies, including surveys and human-subjects testing. This is because ABMs can be exercised to identify which variables have the most impact on outcomes. These are the variables that are most important to characterize through experiments. Hence, there is a feedback loop between experiments and modeling.

Contributions

A summary of our major contributions follows.

1. A synthetic population of 19,000 college undergraduates. A social network of the state of Virginia was generated using the procedure in[25] and used as the starting point for our work. The social network contains synthetic individuals whose traits match in distribution the attributes of the actual population. From this network, we extracted college students (those agents in the age range 18-22 years that have college activities and are located in the vicinity of a major university). Edges in the original social network are retained if the incident vertices are both college students. This produces a social network of 18,866 students (agents) and 119,139 pairwise student daily interactions.

2. Agent-based model of depression. We extend a model of within-host depression evolution[26] to include the effects of social interactions. Consequently, our model accounts for internal, environmental, and social factors of the evolution of depression, consistent with many research

studies[17,18,19]. In particular, each agent has an internal network with 14 vertices that represent depressive symptoms. These symptoms can influence each other and we take these interactions from the literature. Each agent is connected to peer students in the social network with whom she comes in contact. Influence is transmitted through the edges. A student becomes depressed if a sufficient number of symptoms becomes activated through within-host and external interactions. A student may transition back and forth between depressive and healthy states.

3. Simulation of depression in a college population. We code the agent-based model in a simulation system and exercise it using various inputs. We show that the number of depressed students changes with the strength of influence in the symptom interactions and in the social network. We explore the effect of initial conditions, and illustrate the interesting result that within-agent symptom connectivity changes the magnitude of the steady state level of depression within the population.

Related Work

Depression is a primary factor for dropping out, and it can spread from person-to-person through social interactions, as stated in the Introduction. Several studies also identify peer influence as a main trigger for dropping out. Peer influence can exist in many forms. Mayer[27] found, for high school students, that the better a student's peers performed in school, the more likely the student is to drop out. Gaviria[28] also found peer-based effects for dropping out of high school. Crane[29] proposes a contagion (i.e., peer influence) approach to understanding social problems (including school dropping out), using an epidemic-like approach that is similar in spirit to the first epidemic-inspired social model[30]. The model is closely related to segregation models[31] and threshold models[32,33,34]. A relatively recent overview of contagion-like influence is given in[35]. Bank[36] identified different types of influence on students' persistence such as peer, faculty, and parental influences. Parents and peers were found to have stronger influences than were the faculty on the persistence of students.

Pyari[37] investigated the effects of anxiety among medical and engineering students. Results showed that medical students exhibited low anxiety in comparison to engineering students. Between 42% and 48% of PhD students in science and engineering at the University of California are depressed[38]. Vitasaria[39] studied the relationship between anxiety and academic performance. Results showed a significant correlation of high level anxiety and low academic performance among engineering students.

Students may find difficulty in keeping up with schoolwork; 36 percent of the students reported feeling frequently stressed[40]. No student reported a complete lack of worry about keeping up with schoolwork. Twenty-five percent of students reported frequent inability to pursue non-academic activities due to lack of time. Ten percent reported feeling most of time that they did not have a social life, while another 41 percent reported occasionally feeling this way. These life imbalances can eventually contribute to development of depression. Another study[41] revealed relationships between mental health and year of study, academic program, and gender. Rizwan[42] found factors that affect the stress level of female engineering students. Results indicate that teachers' discouraging attitudes have the strongest effect on the stress levels of female engineering

students. A study[43] by Cornell's College of Engineering involving 35.5% of the student population showed that the main sources of stress for engineering students include heavy workloads in engineering courses, large amount of time needed to finish assignments, not enough sleep, competition with classmates, and inflexibility of the Engineering curriculum.

Astin[44] in the theory of involvement, proposed that the greater the student's social involvement in college, the greater the student's learning and personal development and the less likely she is to leave. In the theory of attrition, Bean[45] identified three categories of reasons leading to student attrition: 1) background, 2) organizational and environmental factors, and 3) attitudinal and behavioral outcome. Spady[46] proposed a theory based on the use of Durkheim's[47] theory of suicides to explain freshman attrition. Spady's theory was the basis of Tinto's work[48]. It states that when a person shares values with a group, this person is less likely to commit suicide (or by analogy, drop out of school). Tinto identified in the interactionist theory three reasons for student departure: 1) academic difficulties, 2) inability of students to achieve their goals, and 3) failure to adapt to the institution social environment. Tinto's model focused more on academic and social integration.

Most of the mentioned studies follow an experimental or clinical approach, and may use quantitative, qualitative or mixed-method techniques to study student drop out phenomena and leading causal factors. These studies involve human subjects as well. Other studies take another approach through the use of agent-based modeling to study depression. Aziz[49] proposed a dynamic agent model of recurrences of depression for an individual. Borkulo[26] shows the effect of interactions among different depression symptoms, which is called the causal interactions network. Both[50] proposed another model that has been used to simulate different scenarios in which personal characteristics determine the effect of stress on the (long-term) mood of a person. These agent-based models simulate the evolution of depression within a single agent/person. Our model goes beyond this point to simulate the evolution of depression both within a single agent and across agents of a population.

These theories, studies, and experiments indicate that depression is a major contributor to dropping out of school. Thus as a first step, we model the evolution of depression within agents and its transmission across agents. The results from simulations will be useful in follow-on work to forecast the impact on retention.

Data and Methodology

College Social Network

We use a realistic college population over which we study depression dynamics and peer influence. We model the undergraduate student body of a large university. A modeling process[25] was used to construct this population, which creates anonymous students and endows them with traits such as age, gender, and sets of activities that result in daily face-to-face interactions with other students. The result of this process is a college social network, where nodes/agents represent students and undirected edges represent interactions between students.

To produce this network, we start with the social contact network of the state of Virginia. We then

extract from this network all people with age between 18 and 22 (inclusive) years that have at least one edge (i.e., interaction) with another college-age student, and who are geographically located in the vicinity of a particular public university. These agents and their interactions in a normative day form the social network, whose traits are given in Table 1. An agent has roughly 12 interactions a day with other students. There are over 200,000 "friend of friend" relationships (i.e., triangles) in the network. The diameter of the network is quite large compared to many social networks. Although there are 105 connected components, the network has a giant component that contains 98.7% of all nodes (agents).

Table 1: College network structural characteristics.

Network Property	**Description**	**Value**
Number of Nodes	Total number of students	18866
Number of Edges	Total number of peer-to-peer interactions	119139
Number of Triangles	Number of student groups of size three and form a cycle	202318
Average Degree	Average number of edges connected to a node	12.63
Diameter	Longest of all the calculated shortest paths in the network	16
Average Path Length	Sum of shortest paths between all pairs of nodes divided by the total number of pairs	4.785
Density	Ratio of the number of edges to the number of possible edges	0.001
Modularity	Fraction of edges that fall within a group, minus the expected number of edges within group	0.484
Number of Communities	Number of node groups in the network	151
Average Clustering Coefficient	Measure of the degree to which the nodes tend to cluster together	0.348
Number of Connected Components	Number of node groups that are mutually reachable by undirected edges	105
Size of Giant Component	The fraction of nodes in the largest connected component	0.987

Contagion (Behavioral) Model

We propose a model that quantifies the diffusion of activated depression symptoms among students. The model accounts for the major factors that affect the dynamics of depression. We study two types of dynamics: (i) internal dynamics within a single agent and (ii) external dynamics between agents. Internal dynamics describe how depression evolves within a student as a result of symptom interactions. Symptoms include depressed mood, loss of interest, weight loss, weight gain, decreased appetite, increased appetite, insomnia, hypersomnia, psychomotor agitation, psychomotor retardation, fatigue, worthlessness or guilt, concentration problems, and suicidal thoughts[26]. It is based on the hypothesis that symptoms of mental disorders have direct

causal relations with one another and is called the causal network perspective[51,52,53,26]. External dynamics focus on a student's peer interactions with her roommates, classmates or friends. The social contact network edges are the external interactions.

Our modeling approach follows graph dynamical systems (GDSs)[54]. A GDS is composed of four elements: (i) a network $G(V, E)$ with edge set E and vertex set V where $n = |V|$ and $m = |E|$; (ii) a set K of vertex states (a vertex is in one of these states at each time t in a simulation); (iii) a set F of vertex functions (one function f_i for each vertex/agent i); and (iv) a specification W of the order in which vertex functions are executed. Figure 1 shows an illustrative example of a six-agent network, with an internal view of two agents 4 and 6. Within-agent edges represent the causal relations between symptoms of that agent, while undirected edges across agents are the peer interactions. In the following discussion, we will refer to the internal symptom network as G^1, and to the across agent network as G^2.

In Figure 1, there is a single edge between agents 4 and 6. This single edge represents the multiple edges between symptoms of agents 4 and 6. We do not show all of these to reduce clutter. In general, each symptom of agent 4 can be connected to any number of symptoms of agent 6. In this work we confined ourselves to between-agent edges that connect the same symptoms, so there are 14 edges between the symptoms of agents 4 and 6.

In our model, the network is described as $G(V, E)$ where G is a composite graph of G^1 and G^2. $G^1(V^1, E^1)$ is the undirected graph of depressive symptoms within each student, and is fixed for all students. It represents the causal network of within-host symptom interactions. The set V^1 is the vertex set of symptoms, $n^1 = |V^1| = 14$. Let $v^1_{ki} \in V^1$ be the stress-generating symptom k for agent i. The edge set E^1 represents the direct causal relations between two symptoms where $e^1_{kl,i} \in E^1$ is the undirected edge between symptoms k and l for agent i and $m^1 = |E^1| = 17$ (see Figure 1). The state x^1_{ki} of v^1_{ki} is either 0 if symptom k is not activated or 1 if it is. The state set $K^1 = \{0, 1\}$.

Graph $G^2(V^2, E^2)$ is the network describing the student contacts in the population, where V^2 is the vertex set of (human) agents and E^2 is the edge set of their daily interactions, $n^2 = |V^2|$ and $m^2 = |E^2|$. For this problem, n^2 and m^2 are given in Table 1. Each element $v^2_i \in V^2$ is a student i that can be considered a supernode, that contains a graph G^1 describing internal dynamics. Each edge in E^2 represents a set of connections between pairs of symptoms in neighboring agents. Specifically, $e_{k_1 i, k_2 j}$ represents an undirected edge between symptom k_1 of agent i and symptom k_2 of agent j. Throughout this study, to simplify simulations and interpret results, we take $k_1 = k_2$. That is, only the same symptoms are connected between two neighboring agents in the social network. This choice conforms to research findings[55,56,57], which show that homophily in depression symptoms may stem from peer influence.

The set K^2 of agent vertex states is defined as $K^2 = \{0, 1\}$. Student i can be in one of two states, healthy $x^2_i = 0$ or depressed $x^2_i = 1$. A student is considered depressed if there are eight or more developed/activated symptoms. Thus, an agent can transition back and forth between depression and no depression depending on its number of activated symptoms. The probability for symptom v^1_{ki} to become activated is represented as p^1_{ki}.

Two types of vertex functions are presented in the model: The set F^1 of vertex functions f^1_{ki} determines the state of each stressing symptom v^1_{ki} for student v^2_i at time step t. The set F^2 of

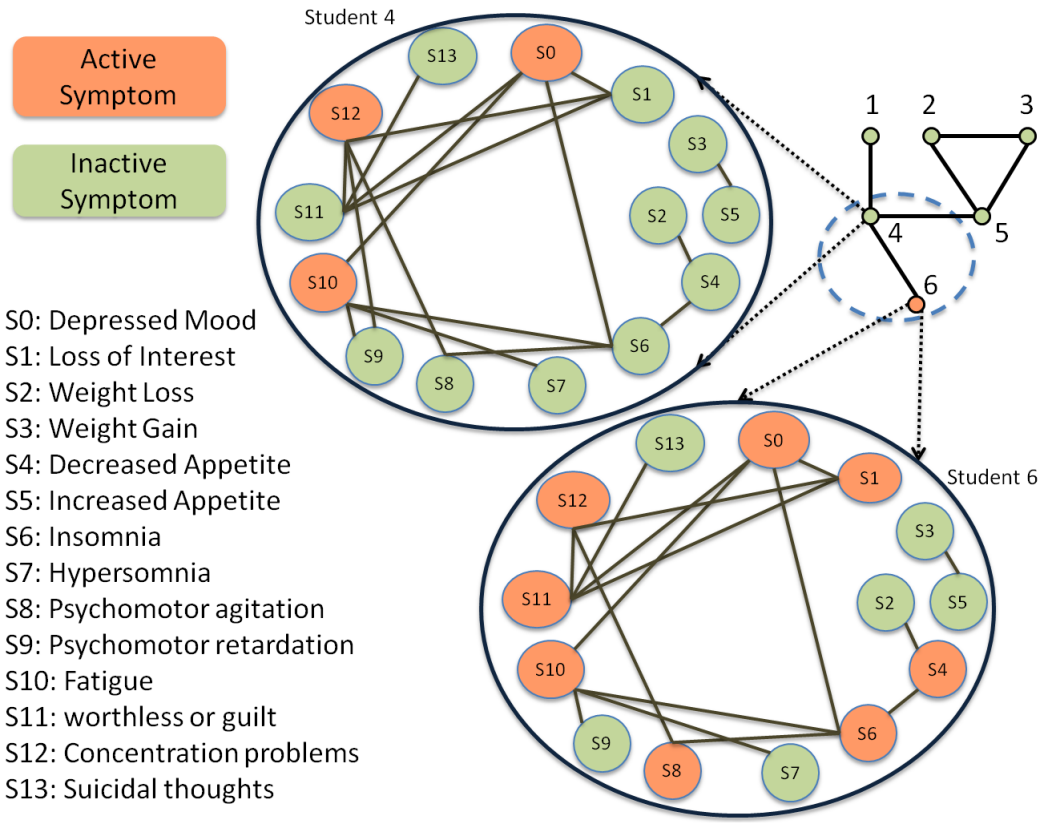


Figure 1: Illustrative example network, showing local interactions between two students labeled 4 and 6, who are part of a larger student population of six agents; this is network G^2. The color of symptoms in the two large circles determine the symptom state which can be orange (active) or green (inactive). Each within-agent symptom network is an instance of G^1. The total number of active symptoms determines whether the agent will be in a healthy or a depressed state. For example, agent 6 has 8 active symptoms, and as a result, the agent color turns to orange (depressed). The with-in agent network is taken from[26].

vertex functions f_i^2 determines the state of each student v_i^2 at each time step t. (See Table 2.)

The specification W of the order in which every f_{ki}^1 and f_i^2 is executed, is as follows. At each time t, each agent i computes the state x_{ki}^1 for symptom k, using f_{ki}^1. Then each agent i executes f_i^2 to determine the state x_i^2 at time t. An agent i is depressed if $x_i^2 = 1$ and not depressed if $x_i^2 = 0$. As we will see below, because of the forms of f_{ki}^1 and f_i^2, each agent can perform these computations in parallel. The critical number τ_{crit} of activated symptoms causing a person to change state is $\tau_{crit} = \lfloor n^1/2 \rfloor$. That is, at least one-half of a person's symptoms must be activated to cause a person to transition to, or remain in, state 1.

Definitions of each variable and constant are provided in Table 2. The values for a_{ki}^1 and b_{ki}^1 are based on the VATSPUD data used in[58,59]. The variable A_{ki}^1 is the total amount of stress on symptom v_{ki}^1. It consists of (i) the individual stress level z_{ki}^1, (ii) the amount of external activation c_i^2 (the amount of stress on the symptom network for student v_i^2), (iii) the influence of the activation of neighbors of symptom v_{ki}^1, which is denoted by q_{ki}^1, and (iv) the student's peer influence on symptom v_{ki}^1 denoted is by r_{ki}^2. The variable q_{ki}^1 depends on whether or not the

Table 2: List of model variables and user input parameters. *rnd* denotes a random number in $[0, 1]$.

Name	Definition	Type	Range	Equation
p^1_{ki}	Activation probability for symptom v^1_{ki} in G^1.	$\mathbb{R}$	$[0, 1]$	$\frac{1}{1+e^{a^1_{ki}(b^1_{ki}-A^1_{ki})}}$
A^1_{ki}	Total amount of stress on symptom v^1_{ki} in G^1.	$\mathbb{R}$	$[-\infty, +\infty]$	$z^1_{ki} + c^2_i + q^1_{ki} + r^2_i$
r^2_{ki}	Student's peer influence in G^2. $N^2(v^2_i)$ is the set of peers (distant-1 neighbors) for student v^2_i. w^2_{ij} is edge weight between student v^2_i and student v^2_j whose state is x^2_j.	$\mathbb{R}$	$[0, +\infty]$	$\sum_{v^2_j \in N^2(v^2_i)} (w^2_{ij} x^1_{kj})$
q^1_{ki}	Symptom's distant-1 neighbors influence in G^1. $N^1(v^1_{ki})$ is the set of neighbors for symptom v^1_{ki}. w^1_{kici} is edge weight between symptoms v^1_{ki} and v^1_{ci} whose state is x^1_{ci}.	$\mathbb{R}$	$[0, +\infty]$	$\sum_{v^1_{ci} \in N^1(v^1_{ki})} (w^1_{kici} x^1_{ci})$
τ^2_i	Number of activated symptoms for student v^2_i.	$\mathbb{N}$	$[0, 14]$	$\sum_{k=0}^{n^1-1} f^1_{ki}$
c^2_i	Amount of stress on symptoms network in student v^2_i.	$\mathbb{R}$	$[-8, 8]$	Model Parameter
z^1_{ki}	Individual stress level of symptom v^1_{ki}.	$\mathbb{N}$	$[-5, 5]$	Model Parameter
a^1_{ki}	Symptom-specific parameter controls steepness of p^1_{ki}.	$\mathbb{R}$	$[-\infty, +\infty]$	Model Parameter
b^1_{ki}	Symptom-specific parameter for the threshold of symptom v^1_{ki}.	$\mathbb{R}$	$[-\infty, +\infty]$	Model Parameter
w^1_{kici}	Edge weight between symptom v^1_{ki} and symptom v^1_{ci} within same agent i in G^1.	$\mathbb{R}$	$[0, 1]$	Model Parameter
w^2_{ij}	Edge weight between student v^2_i and student v^2_j in G^2 $(i \neq j)$. This is the weight between all corresponding symptoms of two agents.	$\mathbb{R}$	$[0, 1]$	Model Parameter
f^1_{ki}	Symptom k of agent i state transition function in G^1.	$\mathbb{N}$	$\{0, 1\}$	$f^1_{ki} = \begin{cases} 1 & \text{if } p^1_{ki} - rnd > 0 \\ 0 & \text{otherwise} \end{cases}$
f^2_i	Student's i state transition function in G^2.	$\mathbb{N}$	$\{0, 1\}$	$f^2_i = \begin{cases} 1 & \text{if } \tau^2_i > \tau_{crit} \\ 0 & \text{otherwise} \end{cases}$

symptom's neighbors are activated and on the strength of the connection between the activated neighbor and symptom v^1_{ki}. The variable r^2_{ki} quantifies peer influence in G^2 and depends on whether or not the symptom v_{ki} of student i's direct contacts (e.g., roommates, classmates or friends) are activated and on the strength of the relation between the depressed student and her/his peer v^2_j.

We relate particular equations in Table 2 to the GDS model and networks of Figure 1. First, the equation for q^1_{ki} captures the interactions among symptoms of the within-agent network G^1. For symptom k of agent i, $N^1(v^1_{ki})$ denotes the neighbors v^1_{ci} for symptoms c. The states x^1_{ci} of these distance-1 neighboring symptoms, along with the edge weights w^1_{kici} between symptoms k and c, contribute to the next state of symptom v^1_{ki} through q^1_{ki}.

Second, r^2_{ki} captures social influence on agent i in G^2. For each of the neighbors v_j of v_i in G^2, the symptom v_{kj} influences the symptom v_{ki} of v_i. Here, we take the strength w^2_{ij} of the interaction between corresponding symptoms of two agents i and j as the same for all symptoms k. We further assume that only corresponding symptoms interact. Neither of these assumptions is limiting and can be relaxed. In this case, then, the general interaction term r^2_{ki} would be $r^2_{ki} = \sum_{v^2_j \in N^2(v^2_i)} \left(\sum_{k'=0}^{n^1-1} (w^2_{k'jki} x^1_{k'j}) \right)$. In this case, $w^2_{k'jki}$ is the weight of an edge in G^2, the *social* network, between symptom k' in v^2_j and symptom k in v^2_i. If no such edge exists, the weight is zero. The outer sum in the last equation is over all neighbors of v^2_i.

Third, the state of depression of an agent i is computed in a two-step process. At each time t, f^1_{ki} is computed and symptom k of agent i is activated if $f^1_{ki} = 1$. Then, f^2_i is evaluated by determining whether τ^2_i—the number of activated symptoms—is more than one-half of symptoms for agent i. If so, then the agent i is depressed. Note that in these equations, an agent j may not be depressed, but if it has activated symptoms, then it can contribute to the depression of its neighbors. Note also that the depressive state x^2_i of an agent i is not directly used in the depression evaluation of its neighbors.

Model Behavior

The model has many parameters. Here we focus on the effects of stress level z^1_{ki}, edge weight w^2_{ij}, and agent's degree in the social network on the activation probability p^1_{ki} of symptoms, since activated symptoms govern depression of agents. We fix all other factors. Figures 2(a) shows the effect of varying the symptom individual stress z^1_{ki} from -5 up to 5 on symptom activation probability for all 14 symptoms. Other factors c^2_i, z^1_{ki} and w^2 are set to 0, 0 and 0.5 respectively. We use w^2 as the value of w^2_{ij} for all i and j, as seen in Table 2. Symptom 0 contributes the most to the increase in activation probability. Differences in results across symptoms are large for z^1_{ki} between 0 and 3.

In Figure 2(b), we investigate the effect on the activation probability of symptom 0 for each agent of increases in the numbers of depressed friends and changes in edge weights w^2. Other factors c^2_i and z^1_{ki} are set to 0. As the number of depressed friends increases, the internal symptom's activation probability increases. This effect is more pronounced as w^2 increases. This leads to an increase in the number of active symptoms and hence to an increase in the vulnerability to depression.

We repeated the experiment but now with all of the symptoms. The weights w^2 were fixed to 0.05 and factors c_i^2, z_{ki}^1 are set to 0. The results are shown in Figure 2(c). We find the same effect of depressed friends on all symptoms, but the magnitude in activation probability changes across symptoms. These results are expected as symptom 0 is highly connected in the symptom network of Figure 1.

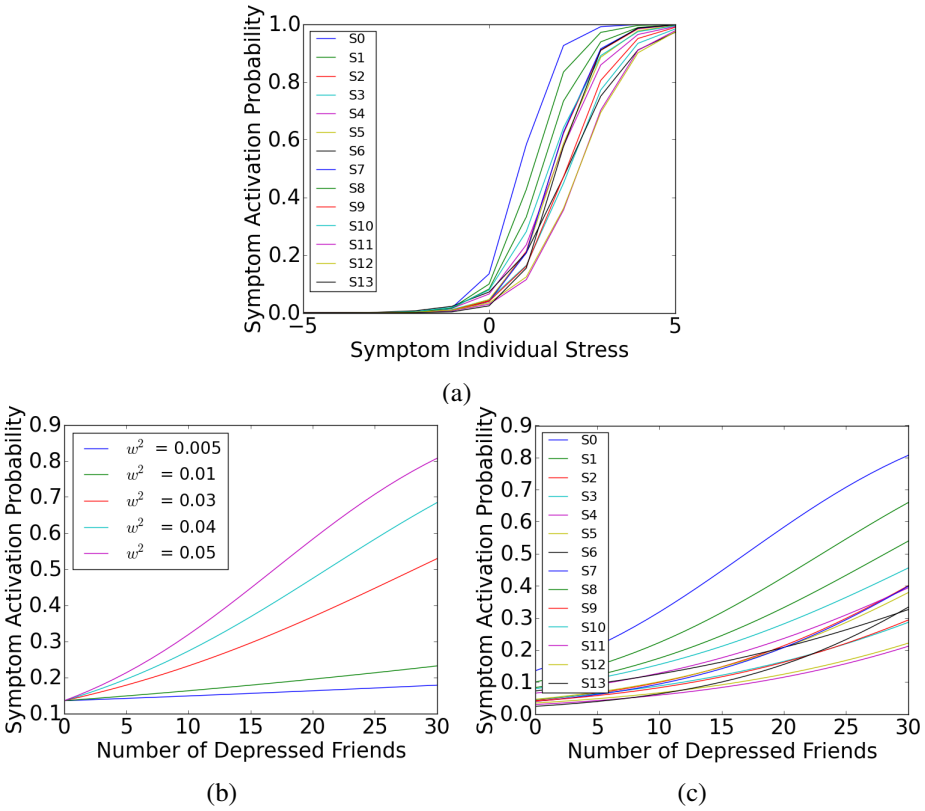


Figure 2: Results for the depression model. (a) The changes in symptom activation probabilities for all symptoms as a function of stress level for each symptom. (b) The changes in symptom activation probability for symptom 0 as a function of number of depressed neighbors and the strength of peer influence. (c) The changes in symptom activation probabilities for all symptoms as a function of number of depressed neighbors.

Simulation Description

Simulation starts at time $t = 0$. At each time step, a symptom v_{ki}^1 within student v_i^2 has a state x_{ki}^1, and its function f_{ki}^1 computes the symptom state change from state $x_{ki}^1(t)$ at time t to state $x_{ki}^1(t+1)$ at time $t+1$. Similarly at each time step, a student v_i^2 has a state x_i^2, and its function f_i^2 describes how the student changes her depressive state from state $x_i^2(t)$ at time t to state $x_i^2(t+1)$ at time $t+1$ based on the number of activated symptoms at time $t+1$. Healthy students are susceptible to depression at any time based on their current state and their peer influence. Students may also develop some symptoms based on external sources, such as experiencing or watching stressful events.

The main simulation steps are described in Algorithm 1, which is the entry point for the simulation. It describes how students' states change over time based on the current number of activated symptoms. The process of identifying and counting activate symptoms is described in Algorithms 2 and 3. We take $simDuration$ = 12 (months) and $depressionThreshold$ = 8.

Algorithm 1: depressionDevelopment

Input: Simulation parameters, including simDuration and depressionThreshold
Output: Students' depressionStatus
simTime = 0;
while *simTime* $<$ *simDuration* **do**
  foreach $v_i^2 \in V^2$ **do**
    $n \longleftarrow countActiveSymptoms(i, k, a_{ki}^1, b_{ki}^1, z_{ki}^1, c_i^2, r_{ki}^2, v_{ki}^1, N^1(v_{ki}^1), N^2(v_i^2), v_i^2)$;
    if *n >= depressionThreshold* **then**
      $depressionStatus \longleftarrow 1$;
    else
      $depressionStatus \longleftarrow 0$;
  $simTime \longleftarrow simTime + 1$;

Algorithm 2: countActiveSymptoms

Input: $i, k, a_{ki}^1, b_{ki}^1, z_{ki}^1, c_i^2, r_{ki}^2, v_{ki}^1, N^1(v_{ki}^1), N^2(v_i^2), v_i^2$
Output: count
count = 0;
foreach $v_{ki}^1 \in V^1$ **do**
  $p_{ki}^1 \longleftarrow activationProbability(i, k, a_{ki}^1, b_{ki}^1, z_{ki}^1, c_i^2, r_{ki}^2, v_{ki}^1, N^1(v_{ki}^1), N^2(v_i^2), v_i^2)$;
  $rnd \longleftarrow uniformRandomNumber()$;
  if $p_{ki}^1 - rnd > 0$ **then**
    $symptomState \longleftarrow 1$;
    $count \longleftarrow count + 1$;
  else
    $symptomState \longleftarrow 0$;
return count;

Simulation Results

Effect of strength of peer influence on the number of depressed students

We start by assuming that all agents are depression symptoms free. We evaluate the impact of peer influence on the final number of depressed students. We fix the model input parameters c_i^2, z_{ki}^1 and w_{kici}^1 to 0.5, 0 and 0.5 respectively. We systematically increase the edge weights w^2 between agents from 0.005 to 0.5. We use w^2 as the value of w_{ij}^2 for all i and j. Peer influence r_{ki}^2

Algorithm 3: activationProbability

Input: i, k, a^1_{ki}, b^1_{ki}, z^1_{ki}, c^2_i, r^2_{ki}, v^1_{ki}, $N^1(v^1_{ki})$, $N^2(v^2_i)$, v^2_i
Output: p^1_{ki}
$r^2_i \longleftarrow 0$;
$q^1_{ki} \longleftarrow 0$;
foreach $v^1_{ci} \in N^1(v^1_{ki})$ **do**
  $q^1_{ki} \longleftarrow q^1_{ki} + w^1_{kici} x^1_{ci}$;
foreach $v^2_j \in N^2(v^2_i)$ **do**
  $r^2_{ki} \longleftarrow r^2_{ki} + w^2_{ij} x^1_{kj}$;
$A^1_{ki} \longleftarrow c^2_i + q^1_{ki} + r^2_{ki} + z^1_{ki}$;
$p^1_{ki} \longleftarrow \frac{1}{1+e^{a^1_{ki}(b^1_{ki} - A^1_{ki})}}$;
return p^1_{ki};

increases as w^2 increases, as seen in Table 2. Each curve in the results of Figure 3(a) represents the average results from 50 simulation runs. Each curve also corresponds to a single value of w^2 for all edges in the social network. As w^2 increases, the peer influence increases, and the number of depressed students increases. When w^2 increases by 67%, from 0.03 to 0.05, the number of depressed students increases by a factor of six, from 400 to 2400. Hence, we see the importance of modeling between-agent interactions through the social contact network.

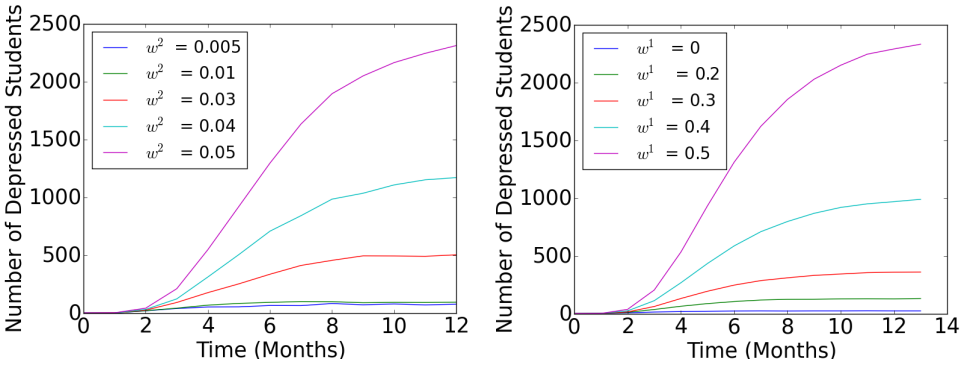
Effect of strength of symptom influence on number of depressed students

We study the symptom influence within the same agent. All agents are initially depression symptom free. We fix other factors, including the peer influence, to isolate this effect. The fixed values for model parameters c^1_i, z^1_i and w^2 are 0.5, 0 and 0.05 respectively. The results are shown in Figure 3(b). We take $w^1_{kici} = w^1$, for all symptoms k and c, and all agents i. As w^1 increases by $2.5\times$, from 0.2 to 0.5, the average number of depressed students increases by $16\times$, from 150 to 2400. The fraction of depressed students increases from 0.8% to 12.7%.

Students predisposed to depression

Students may start their college years with symptoms and signs of depression. In previous simulations we assumed that the entire population is free of depression symptoms. In this experiment, we study how pre-existing depression symptoms might affect the speed at which depression spreads among students. Model parameters c^1_{ki}, z^1_{ki}, w^1 and w^2 are set to values 0.5, 0, 0.5 and 0.05, respectively. We used p and n to represent the probability p of having n active symptoms at time = 0. Figure 4(a) shows that as n and p increase, the speed of depression contagion increases at early time. The final number of depressed students converges over time.

As mentioned earlier, the 14 symptoms are connected through an internal causal network shown in Figure 1 that has 2 connected components, including a giant component of 12 of the 14



(a) (b)

Figure 3: (a) The time evolution of number of depressed students, showing that as peer influence increases, the number of depressed agents increases. (b) The time evolution of number of depressed students, showing that as the strength of symptom-to-symptom influence increases, the number of depressed agents increases. These two plots illustrate the importance of capturing agent interactions in the social network (in (a)), and capturing within-agent interactions (in (b)).

symptoms. In Figure 4(b), we examine this symptom connectivity and two different (extreme) cases: all the symptoms are fully connected, and all are disconnected. The partially connected curve corresponds to the connectivity shown in Figure 1. Results illustrate the impact of having all symptoms fully connected in speeding the contagion process and having a larger number of depressed students at steady state. These data suggest the interesting result that internal agent connectivity dictates population level depression, at least for the conditions of these simulations. Also, the results for the connectivity of Figure 1 are much closer to those for disconnected symptoms than those for fully connected symptoms.

Another experiment on the symptom causal network aims to determine the effect of the most important symptoms (those most highly connected) being initially activated. Importance of symptoms is proportional to their degrees. Highest degree symptoms are (0, 1, 10 and 11) with degrees (4, 3, 4 and 4), respectively. Lowest degree symptoms are (3, 5, 7 and 13) with degrees (1, 1, 2 and 1), respectively. The results are shown in Figure 4(c). The results show that activating the highest degree symptoms accelerates the depression at early times.

Discussion

Clearly our model is focused mainly on modeling and simulation of depression among students. We are motivated by several studies, which are cited earlier in the Introduction and Related Work. These studies identify and emphasize the importance of depression on a student's academic performance and retention. A 15 point increase on the depression scale correlates with a 0.17 drop in GPA and corresponds to a 4.7 percent increase in probability of dropping out[60]. Therefore, we believe that the simulation results on depression can be used as the basis for evaluating the impact of depression on retention.

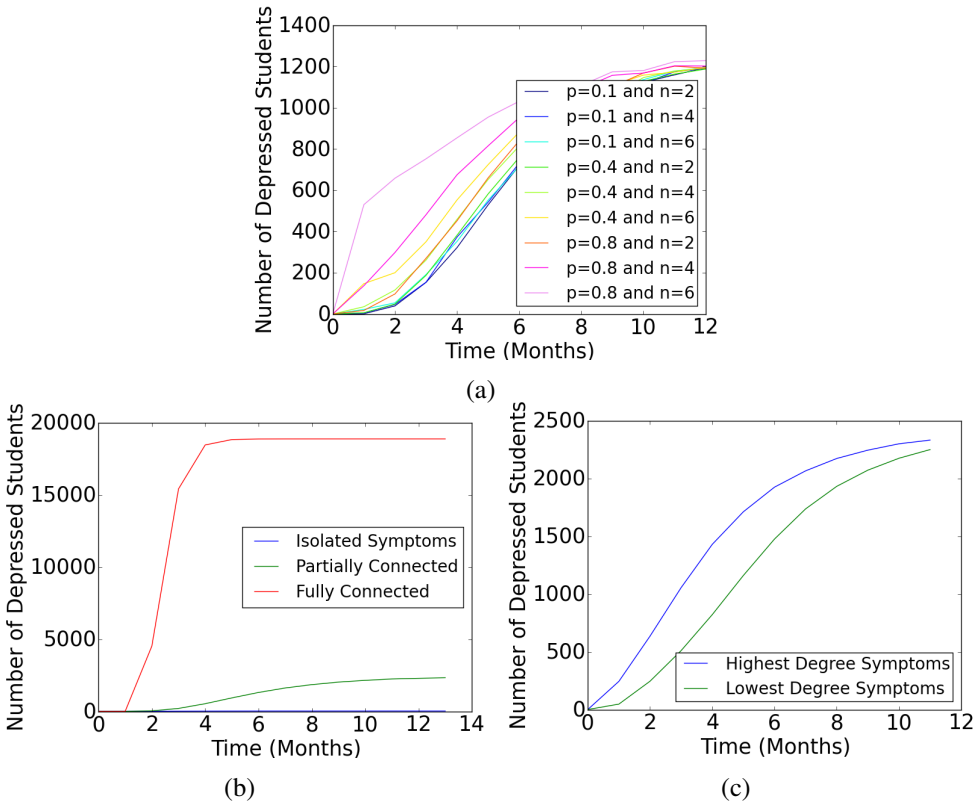


Figure 4: (a) The time evolution of number of depressed students, showing that as p and n increase, the contagious speed of depression increases. (b) The time evolution of number of depressed students, showing the impact of having all symptoms fully connected in speeding the contagion process and having a larger number of depressed students at steady state. (c) The time evolution of number of depressed students, showing the impacts of having the highest degree and lowest degree symptoms initially activated.

Implications

In order to decrease the rate of depressed students, both students and universities must fully understand the illness, the causes, and its effects. There is a need for developing programs that improve prevention, identification, and treatment of depression. Collected data reveals that around three percent of first-year students experiencing depressive symptoms seek help and counseling services[61]. Another study showed that living off campus was related to less knowledge of campus mental health services[62]. Existing programs focus on within-student depression symptoms. Considerations should also be given to the contagious aspect of depression through peer influence. Both students and universities need to be able to identify risk factors, external stressors, and signs and symptoms of the disease.

Some work[63] suggested that university mental health programs can also benefit from adapting ideas that have already shown promising results in settings outside of college campuses, such as interventions that have been effective in general populations of adolescents or young adults. Some of Tinto's recommendations[64] to promote retention can be utilized to control depression through providing academic, social and personal support, particularly in and before the first year, and

showing students that they are valued. Student led organizations may help lessen major depression stressors while also providing a secure social network for those who are at high risk of suffering from depression[65].

Conclusions and Future Work

Conclusions

In this work, we construct and exercise a causal agent-based model (ABM) of the evolution of depression among a population of college students. Like other models[66,49,50], we model the evolution of depression-related traits *within* each agent (this is where other studies stop, treating each person of a population as an isolated individual). Our within-agent model makes use of the model in[26]. Our approach goes beyond this point, and is novel in several respects, which form our contributions. First is the generation and use of a synthetic college population over which these dynamics occur. Second, we extend a model of within-host depression evolution to include the effects of social interactions. Finally, we code the agent-based model in a simulation system, and conduct simulations to understand population-level results that are produced from local agent interactions. We explore the effect of initial conditions, and among other findings, illustrate the interesting result that within agent connectivity changes the magnitude of steady state depression within the population.

Future Work

Our future work includes conducting simulation experiments on students populations of different universities (e.g. teaching vs. research and state vs private). As stated in the Related Work, several factors can affect the student's academic retention. Therefore depression data can not be used solely to inform a student's future behavior. We will investigate adding agent domain attributes to the model (e.g. income, marital status and medical history, gender or ethnicity), and examine how the model will behave.

Acknowledgements

We thank the anonymous reviewers for their useful comments and suggestions. We thank our external collaborators and members of the Network Dynamics and Simulation Science Laboratory (NDSSL). This work was partially supported by DTRA Grant HDTRA1-11-1-0016, DTRA CNIMS Contract HDTRA1-11-D-0016-0001, and NSF NetSE Grant CNS-1011769.

References

[1] W. K. LeBold, R. Delauretis, and K. D. Shell. The purdue interest questionnaire: An interest inventory to assist engineer students in planning their career. *annual Frontiers in Education Conference*, 1977.

[2] John P. Bean. Dropouts and turnover: The synthesis and test of a causal model of student attrition. *Research in higher education*, 12(2):155–187, 1980.

[3] James A. Beane and Richard P. Lipka. *Self-concept, self-esteem, and the curriculum*. Columbia University, Teachers College, 1986.

[4] Malene Rode Larsen, Hanna Bjornoy Sommersel, and Michael Søgaard Larsen. *Evidence on Dropout Phenomena at Universities*. Danish Clearinghouse for educational research, 2013.

[5] Mark Schneider and Lu Yin. The high cost of low graduation rates: How much does dropping out of college really cost? *American Institutes for Research*, 2011.

[6] B. Zimmerman and D. H. Schunk. Competence and control beliefs: Distinguishing the means and ends. *Handbook of educational psychology*, pages 349–367, 2006.

[7] Watson S. Swail. The art of student retention: A handbook for practitioners and administrators. In *Educational Policy Institute. Texas Higher Education Coordinating Board 20th Annual Recruitment and Retention Conference Austin, TX June*, volume 21, page 2004, 2004.

[8] Clifford Adelman. Answers in the tool box. academic intensity, attendance patterns, and bachelor's degree attainment, 1999.

[9] William K. LeBold. Research in engineering education: An overview. *Engineering Education*, 70(5):406, 1980.

[10] Jie Lu and Daniel Churchill. The effect of social interaction on learning engagement in a social networking environment. *Interactive learning environments*, 22(4):401–417, 2014.

[11] Cheryl L. Carmichael, Harry T. Reis, and Paul R. Duberstein. In your 20s it's quantity, in your 30s it's quality: The prognostic value of social activity across 30 years of adulthood. *Psychology and aging*, 30(1):95, 2015.

[12] S. S. Umar, I. O. Shaib, D. N. Aituisi, N. A. Yakubu, and O. Bada. The effect of social factors on students' academic performance in Nigerian tertiary institutions. *Library Philosophy and Practice (e-journal)*, page 334, 2010.

[13] Kadian M. Callahan. Academic-centered peer interactions and retention in undergraduate mathematics programs. *Journal of College Student Retention: Research, Theory & Practice*, 10(3):361–389, 2008.

[14] Matthew Boyko, Ruslan Kutz, Julia Grinshpun, Vladislav Zvenigorodsky, Shaun E Gruenbaum, Benjamin F Gruenbaum, Evgeni Brotfain, Yoram Shapira, and Alexander Zlotnik. Establishment of an animal model of depression contagion. *Behavioural brain research*, 281:358–363, 2015.

[15] Gerald J. Haeffel and Jennifer L. Hames. Cognitive vulnerability to depression can be contagious. *Clinical Psychological Science*, 2(1):75–85, 2014.

[16] Elizabeth A. Stevens and Mitchell J. Prinstein. Peer contagion of depressogenic attributional styles among adolescents: A longitudinal study. *Journal of Abnormal Child Psychology*, 33(1):25–37, 2005.

[17] Mitchell J. Prinstein. Moderators of peer contagion: A longitudinal examination of depression socialization between adolescents and their best friends. *Journal of Clinical Child and Adolescent Psychology*, 36(2): 159–170, 2007.

[18] Christopher K. Hsee, Elaine Hatfield, and Claude Chemtob. Assessments of the emotional states of others: Conscious judgments versus emotional contagion. *Journal of social and clinical psychology*, 11(2):119–128, 1992.

[19] Munmun De Choudhury, Scott Counts, and Eric Horvitz. Social media as a measurement tool of depression in populations. In *Proceedings of the 5th Annual ACM Web Science Conference*, pages 47–56. ACM, 2013.

[20] Adam D. I. Kramer, Jamie E. Guillory, and Jeffrey T. Hancock. Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences*, 111(24): 8788–8790, 2014.

[21] Bernadette von Dawans. *Neuropeptidergic modulation of social behavior in health and social phobia*. Cuvillier Verlag, 2008.

[22] American College Health Association et al. American college health association national college health assessment spring 2006 reference group data report (abridged). *Journal of American College Health*, 55(4):195, 2007.

[23] Darcy Gruttadaro and Dana Crudo. College students speak: A survey report on mental health. *Survey Report). Arlington, VA: National Alliance on Mental Illness*, 2012.

[24] Joshua Epstein. *Generative Social Science: Studies in Agent-Based Computational Modeling*. Princeton University Press, 2006.

[25] Christopher L. Barrett, Richard J. Beckman, Maleq Khan, V. S. Anil Kumar, Madhav V. Marathe, Paula E. Stretz, Tridib Dutta, and Bryan Lewis. Generation and analysis of large synthetic social contact networks. In *Winter Simulation Conference*, pages 1003–1014. Winter Simulation Conference, 2009.

[26] C. D. Van Borkulo, H. L. J. Van der Maas, D. Borsboomm, and A. O. J. Cramer. Netlogo vulnerability_to_depression, 2013. URL `http://ccl.northwestern.edu/netlogo/models/community/Vulnerability_to_Depression`.

[27] Susan E. Mayer. How much does a high school's racial and socioeconomic mix affect graduation and teenage fertility rates? In Christopher Jencks and Paul E. Peterson, editors, *The Urban Underclass*. The Brookings Institution, 1991.

[28] Alejandro Gaviria and Steven Raphael. School-based peer effects and juvenile behavior. *The Review of Economics and Statistics*, 83:257–268, 2001.

[29] Jonathan Crane. The epidemic theory of ghettos and neighborhood effects on dropping out and teenage childbearing. *American Journal of Sociology*, 96(5):1226–1295, 1991.

[30] F. M. Bass. A new product growth for model consumer durables. *Management Science*, 15:215–227, 1969.

[31] T. Schelling. *Micromotives and Macrobehavior*. W. W. Norton and Company, 1978.

[32] Mark Granovetter. Threshold Models of Collective Behavior. *American Journal of Sociology*, 83(6): 1420–1443, 1978.

[33] Mark Granovetter and R. Soong. Threshold models of diffusion and collective behavior. *J. Mathematical Sociology*, 9:169–179, 1983.

[34] D. Watts. A Simple Model of Global Cascades on Random Networks. *Proceedings of the Naitonal Academy of Sciences (PNAS)*, 99(9):5766–5771, 2002.

[35] Duncan Watts and Peter Dodds. Threshold models of social influence. In *The Oxford Handbook of Analytical Sociology*, pages 475–497, 2009.

[36] Barbara J. Bank, Ricky L. Slavings, and Bruce J. Biddle. Effects of peer, faculty, and parental influences on students' persistence. *Sociology of Education*, 63:208–225, 1990.

[37] Dayal Pyari. A comparitive study of depression and anxeity among medical and engineering students. *American research thoughts*, 1(9):12, 7 2015.

[38] Rachel Bernstein. Depression afflicts almost half of STEM graduate students at UC berkeley. *Science*, may 2015. doi: 10.1126/science.caredit.a1500125.

[39] Prima Vitasari, Muhammad Nubli Abdul Wahab, Ahmad Othman, Tutut Herawan, and Suriya Kumar Sinnadurai. The relationship between study anxiety and academic performance among engineering students. *Procedia-Social and Behavioral Sciences*, 8:490–497, 2010.

[40] H. G. Loshbaugh, Tawni Hoeglund, Ruth Streveler, and Kimberley Breaux. Engineering school, life balance, and the student experience. In *Proceedings of the American Society for Engineering Education Annual Conference, Chicago, IL*, 2006.

[41] Melissa Deziel, Dayo Olawo, Lisa Truchon, and Lukasz Golab. Analyzing the mental health of engineering students using classification and regression. In *EDM*, pages 228–231, 2013.

[42] Ali Rizwan, Saadia Farooq, M. S. I. Alvi, and Silvet Nawaz. Analysis of factors affecting the stress level of female engineering students. *Global Journal of Human-Social Science Research*, 12(10-A), 2012.

[43] Lisa Schneider. Perceived stress among engineering students. In *St, Lawrence Section Conference, Toronto, Canada*, 2007.

[44] Alexander W. Astin. Student involvement: A developmental theory for higher education. *Journal of college student personnel*, 25(4):297–308, 1984.

[45] John P. Bean. The synthesis of a theoretical model of student attrition, 1981.

[46] William G. Spady. Dropouts from higher education: An interdisciplinary review and synthesis. *Interchange*, 1 (1):64–85, 1970.

[47] Emile Durkheim. *Suicide: A Study In Sociology*. Free Press, 1997. ISBN 0684836327.

[48] Vincent Tinto. Research and practice of student retention: what next? *Journal of College Student Retention: Research, Theory & Practice*, 8(1):1–19, 2006.

[49] Azizi A. Aziz, Michel C. A. Klein, and Jan Treur. An agent model of temporal dynamics in relapse and recurrence in depression. *Next-Generation Applied Intelligence*, pages 36–45, 2009.

[50] Fiemke Both, Mark Hoogendoorn, Michel C. A. Klein, and Jan Treur. Modeling the dynamics of mood and depression, 2008.

[51] Denny Borsboom. Psychometric perspectives on diagnostic systems. *Journal of clinical psychology*, 64(9): 1089–1108, 2008.

[52] Angélique O. J. Cramer, Lourens J. Waldorp, Han L. J. van der Maas, and Denny Borsboom. Comorbidity: a network perspective. *Behavioral and Brain Sciences*, 33(2-3):137–150, 2010.

[53] Verena D. Schmittmann, Angélique O. J. Cramer, Lourens J. Waldorp, Sacha Epskamp, Rogier A. Kievit, and Denny Borsboom. Deconstructing the construct: A network perspective on psychological phenomena. *New Ideas in Psychology*, 31(1):43–53, 2013.

[54] Henning Mortveit and Christian Reidys. *An introduction to sequential dynamical systems*. Springer Science & Business Media, 2007.

[55] James C. Coyne. Depression and the response of others. *Journal of abnormal psychology*, 85(2):186, 1976.

[56] Aaron Hogue and Laurence Steinberg. Homophily of internalized distress in adolescent peer groups. *Developmental psychology*, 31(6):897, 1995.

[57] Noona Kiuru, William J. Burk, Brett Laursen, Jari-Erik Nurmi, and Katariina Salmela-Aro. Is depression contagious? a test of alternative peer socialization mechanisms of depressive symptoms in adolescent peer networks. *Journal of Adolescent Health*, 50(3):250–255, 2012.

[58] Kenneth S. Kendler and Carol A. Prescott. Genes, environment, and psychopathology. *New York: Guilford*, 2006.

[59] Carol A. Prescott, Steven H. Aggen, and Kenneth S. Kendler. Sex-specific genetic influences on the comorbidity of alcoholism and major depression in a population-based sample of us twins. *Archives of General Psychiatry*, 57(8):803–811, 2000.

[60] Daniel Eisenberg, Ezra Golberstein, and Justin B Hunt. Mental health and academic success in college. *The BE Journal of Economic Analysis & Policy*, 9(1), 2009.

[61] Adam Raunic and Sophia Xenos. University counselling service utilisation by local and international students and user characteristics: A review. *International Journal for the Advancement of Counselling*, 30(4):262–267, 2008.

[62] Jeremy B. Yorgason, Deanna Linville, and Bryan Zitzman. Mental health among college students: do those who need services know about and use them? *Journal of American College Health*, 57(2):173–182, 2008.

[63] Justin Hunt and Daniel Eisenberg. Mental health problems and help-seeking behavior among college students. *Journal of Adolescent Health*, 46(1):3–10, 2010.

[64] Vincent Tinto. Taking retention seriously: Rethinking the first year of college. *NACADA journal*, 19(2):5–9, 1999.

[65] Tiana Keith. Depression and its negative effect on college students. *Undergraduate Research Journal for the Human Sciences*, 9(1), 2010.

[66] Selver Demic and Sen Cheng. Modeling the dynamics of disease states in depression. *PLoS ONE*, 9(10): e110358, 10 2014. doi: 10.1371/journal.pone.0110358.