Growth and impact of a statewide network focused on STEM success

Marjorie Darrah
West Virginia University

Christopher Wheatley
West Virginia University

John Stewart
West Virginia University

ABSTRACT

The First2 Network is a coalition of individuals from multiple universities, K-12 schools, industry, and government organizations from a rural eastern U.S. state who collaborate to ensure that rural, first-generation undergraduate students are prepared and motivated to persist in their science, technology, engineering, and mathematics (STEM) major. Since its inception in 2018, this National Science Foundation-funded project has utilized student summer immersive experiences for incoming freshmen and Networked Improvement Communities to produce replicable best practices, campus student clubs, student ambassador programs, institutional teams, statewide conferences, and many other methods, all for the purpose of promoting student STEM persistence across the state. This study employs social network analysis to explore the structure, growth, and impact of the connections across this Network over the five years of its existence. Social network analysis metrics indicate that the Network grew both in size and connectivity until 2022 when policy changes led to more institutional localization for the purpose of sustainability. Students have formed robust connections with other Network members throughout the course of the project, leading to a higher STEM persistence rate among students in the Network than average at their university. Faculty from different universities across the state have made connections, which has increased productivity as a result of network membership. The available data suggests that the Network has had a positive impact on both student retention and faculty collaboration, which should be sustained and have a positive impact on STEM persistence throughout the state in years to come.

KEYWORDS

Social Network Analysis, First Generation Students, STEM Retention, Institutional Change

INTRODUCTION

For many years it has been known that the United States (U. S.) and many other countries do not produce enough Science, Technology, Engineering, and Mathematics (STEM) workers to fill future positions. (Peri, Shih, and Sparber, 2015; van den Hurk, Meelissen and van Langen, 2019; Boggs, Dukes, and Hawthorne, 2022) According to a White House report, only 20% of U. S. high school graduates are ready for the rigors of STEM majors and, in fact, over the past 15 years, that report also found that the U. S. has produced only 10% of the world's science and engineering graduates. (Herman, 2019) On the other hand, according to the U. S. Bureau of Labor Statistics (2023) the number of STEM jobs is expected to see a 10.8% increase from 2022-2032 as compared to only 2.8% growth for non-STEM careers.

STEM student retention is a problem faced by institutions throughout the U. S. According to a study, the retention rate for STEM majors is 48% nationally. (Snyder and Cudny, 2017) However, the retention rate varies depending on the institution and sector and the rates seem to be lower in rural regions. Saw (2021) found that STEM enrollment at higher education institutions three years after high school graduation is higher in suburban (17%) relative to rural (13%) students.

For college students, as with faculty members, researchers, and industry employees, success often depends on forming the right connections. Most people today have come to appreciate the importance of networking. Forming meaningful connections while in college can make a difference to retention and success. These connections are especially important for rural students since they may struggle with feelings of isolation and homesickness. Students who must leave their communities and travel far from home may be less likely to persist (Cunningham et al., 2008; Howley et al., 2014). Additionally, rural students may place a high value on community and family relationships and focus more on family commitments than individual achievement (Wright, 2012; Howley et al., 2015). To capitalize on these cultural norms for rural STEM majors, it is important to help them build a sense of community at their institution and within their field of study.

Networking is also crucial for the career development of faculty members. Faculty often need to provide evidence of continued academic activity for promotion and tenure at their institutions. Building connections through networks can facilitate the development of social capital and professional achievement (Coleman, 1988, Burke, 1992; Regans & McEvily, 2003; Lin et al., 2017). Often at smaller, rurally located institutions, faculty members have fewer opportunities to engage with colleagues who are in the same field or have the same research interests. These faculty may benefit from engaging with networks outside their institutions where they can connect with other faculty members and industry professionals.

The First2 Network

The First2 (now referred to as the Network) is seeking to address the problem of STEM retention by creating a network within the state to have an impact on factors related to student retention. The Network brings together key stakeholders in 16 organizations representing higher education, K-12 education, federally funded research labs, and industry associations with the common goal of improving the graduation rate of undergraduate STEM students. The Network has the goal of doubling the number of STEM graduates in the rural state of West Virginia in 10 years, while building a collaborative culture to sustain improvement research around techniques that work to support students along the STEM pathway into the workforce.

The Network was conceived around the idea of STEM student success. It was observed that throughout the state, diverse groups of people were executing varied programs that address the STEM persistence problem, however, they were, for the most part, working in isolation from one another. The Network leadership decided that there would be more impact if a collaborative culture could be built where colleagues from various educational institutions learn from one another, where STEM-based industry employees would provide input, and where STEM students could contribute to the conversations at all levels. In response, the Network has built a collaborative working environment through Networked Improvement Communities (Russell et al., 2019) and an infrastructure for learning from one another to replicate promising practices and programs. This allowed for the development and implementation of key ideas and best practices that would build community, student STEM identity, self-efficacy, sense of belonging, and continued interest in pursuing STEM, which, in turn, should lead to retention. These best practices played out in activities including, but not limited to:

- Summer Immersive Experiences for incoming freshman STEM majors that provided authentic research experiences.
- Continued student research funding that provided STEM students money to continue research with a mentor during their undergraduate career.
- Network biannual gatherings that provided students, faculty and industry members a dedicated time to meet and exchange ideas and build relationships.
- Campus student STEM clubs that provided a place for students to lead and network with others in STEM.
- A website to join that included information and a social networking aspect.
- Working groups that focused on the goals of the project and included both faculty and students.
- Hometown Ambassadors program that allowed students to return to their former high schools to encourage students to study STEM.
- Institutional teams made up of faculty and students that helped implement best practices at the academic institutions.
- Funding for faculty and students to implement new ideas and capture results through a Plan-Do-Study-Act cycle.

Purpose of Research

The purpose of this research is to determine if the efforts of the project have resulted in a network of individuals from various backgrounds who can assist in retaining rural, first-generation (FG) students in STEM. It is important to determine if the network is growing and maintaining members over the five years of data collection. It is also important to explore student retention in STEM and faculty productivity and their connection to the network. In this study, social network analysis is utilized to determine how the Network grew over a five year timeframe and to explore the impact on the college and university students and faculty members who have been involved. The changes in the overall structure of the network, how collaborative groups are forming, and what impact these connections are having on student retention and faculty productivity are explored.

This research utilizes the techniques of Social Network Analysis (SNA) to answer the research questions below:

1. How is the overall structure of the network changing?
2. What collaborative groups are forming? What are their compositions?
3. What impact do the Network connections have on student retention and faculty productivity?

BACKGROUND

The background for this work includes theoretical frameworks that guide this research, a review of literature that connects community groups to retaining students, a review of literature that connects networks, collaboration and faculty productivity, and review of a variety of examples of the use of social network analysis in education.

Theoretical Frameworks

This section discusses frameworks that can help explain how community groups and student involvement can mitigate student departure and assist with student retention. It also provides a framework that can help explain how faculty involvement in a network can impact faculty productivity.

Framework for Student Involvement

The Network was formed with the intention of increasing student retention and persistence in STEM by involving rural and FG students in close-knit community groups that help develop their science identity and improve their self-efficacy. These methods are supported theoretically by both Tinto's Model of Student Departure (Tinto & Cullen, 1973; Tinto, 1975) and by Astin's Theory of Student Involvement (Astin, 1984).

Tinto states that the three primary reasons that students withdraw from an institution are difficulties becoming integrated into communities at the institution, difficulties with academics, and difficulties with career choice. The model emphasizes the importance of both formal and informal methods of student involvement in both academic systems and social systems for student retention. Tinto describes formal academic systems as those directly related to academic performance, informal academic systems as interactions between students and faculty/staff, formal social systems as any well-defined extracurricular activities, and informal social systems as interactions between students and their peer groups. According to Tinto, students who have both social and academic connections are more likely to persist than those who do not.

In 1984, Astin performed a longitudinal study to determine the factors that were most related to student persistence (Astin, 1984). He found that students' level of involvement in their institution was not only related to persistence, but it was also linked to performance. Astin's Student Involvement theory was developed from this study. His theory states that the quantity and quality of energy that students invest into their college experience is directly proportional to the amount of personal development and learning a student experiences in college. Involvement is defined loosely as faculty/staff interactions and participation in academic work and extracurricular activities. Astin posits that student involvement should be the primary metric by which effective educational practices or policies are measured.

Both Tinto's Model of Student Departure and Astin's Theory of Student Involvement independently come to the conclusion that students who engage with extracurricular groups and on-campus activities are more likely to persist at their institutions than students who are more isolated. These models lay the groundwork for the current study by theoretically connecting student persistence with student connectedness. This work examines student-student groups forming as well as students' participation in the Network as a whole. These models not only present an underlying argument for the existence of the Network itself, but they also provide context for the connections studied with network analysis.

Framework for Faculty Involvement

While student STEM retention is the main focus of the Network, faculty collaboration is a key component in the Network's success and sustainability. In order to study the Network's impact on faculty who supply sustainability and support in the Network, the lens of Social Capital Theory (Lin et al., 2017) is used. Which is defined as, "investment in social relations with expected returns in the marketplace. (pg. 19)" Lin goes on to discuss why resources that are embedded in a social network can enhance the outcomes of actions.

Social capital theory can provide valuable insights into the dynamics of networking. Coleman's (1988) early work emphasized the role of social capital in human capital formation, discussing the importance of social networks in facilitating educational and professional achievement. The literature also suggests that faculty members should identify and fill structural holes (Burt, 1992) within their networks. So as to bridge gaps within one's personal network in order to facilitate the flow of information, ideas, and resources. The faculty member should find ways to build connections between otherwise disconnected individuals or groups, for example, as in this study, connections with faculty outside their home institution, connections with faculty in other STEM fields, or connections with individuals in industry. Reagans and McEvily (2003) studied how the network structure relates to knowledge transfer and organizational performance. They emphasize how network cohesion and network reach can facilitate information exchange and innovation.

To create social capital, it is important to invest time in building and maintaining meaningful relationships. Faculty members who prioritize these networking activities to nurture relationships with colleagues, collaborators, and students can yield long-term benefits in career advancement, research collaborations that lead to productive work, and professional support. Social capital theory is used as a lens to help explain what meaning connections may have for faculty members.

Close-knit Community Groups and Retention

Some researchers have found that students who become part of a community during their college experience have increased persistence. Many studies reported that students who have more interactions with faculty and other students, live on campus, and engage in more co-curricular activities are more satisfied, get better grades, and have a higher rate of persistence than other students who are not as engaged (Pascarella & Terenzini, 1991; Astin, 1997).

Burke (2019) conducted a comprehensive literature review to examine student retention. He focused on the role of social interaction during the students' higher education experience. His review of the literature suggests that the type of social interaction a student engages in is of little consequence, but that participating in some type of social activity is what matters. Students engaged in co-curricular (e.g., residence hall living and learning community, recreation programming) or curricular (e.g., honors program, intrusive advising) had higher retention rates than their peers who were not as engaged. Burke states, “The literature is clear that student engagement during the higher education experience leads to higher student retention rates and increased institutional commitment. (pg. 12)” Burke’s claim that the type of social interaction does not matter is not in contradiction to Astin’s claim that both the quantity and quality of energy that students invest into their college experience affects learning outcomes because Burke’s claims are focused externally on the activities while Astin’s claims are focused on the internal motivations of students.

There are other ways students can get connected during their undergraduate career. Baker and Pomerantz (2000) studied how participation in a Learning Community (LC) program influenced student performance and satisfaction. They found that students enrolled in these LCs had higher GPA, completed more credit hours, were more satisfied overall, and academic probation was less of problem in the group relative to the non-LC students. They reported that the LCs fostered group bonding and increased interaction with instructors. Additionally, honors programming in higher education has been historically viewed as beneficial for students (Digby, 2005; Bowman and Culver, 2018). One of the reasons for this is the smaller classes that provide more interaction with faculty members and other students. These programs can provide a peer environment in which students can be accepted for their abilities and build relationships while developing their intellects.

The consensus from the literature is that community and social interactions for college students lead to persistence and success. Burke’s review points out that any form of social engagement during the undergraduate experience may contribute to higher retention. While Burke suggests that it is the social interaction and not the type of social interaction that matters, this does align with Astin’s theory of student involvement, which emphasizes both quantity and quality of the interaction. Both learning communities and honors programs offer benefits by allowing students to engage in smaller groups, have increased interaction with faculty and, develop peer acceptance based on abilities. The Network studied in this work attempts to provide this sense of community to rural, first-generation STEM students through continued interactions with STEM faculty members and their peers. These students may not have other means of social interaction or community building. However, through Network summer immersive research experiences, STEM clubs, bi-annual conferences, undergraduate research opportunities and many other project-sponsored activities, these students build a community and a network of other like-minded individuals with similar backgrounds and similar struggles.

Faculty Network, Collaboration, and Research Productivity

Faculty collaboration and productivity are essential for success in academia. Being able to publish is vital for faculty members’ careers and for the advancement of their disciplines. Faculty members need interaction with peers in their discipline and peers who are interested in the same research areas. Hampton and Parker (2011) found that face-to-face interaction was vital to success in synthesis groups and, in turn, resulted in increased peer-reviewed publications in the group. Developing an organization that helps to foster networking is an effective strategy for social integration and inclusion of faculty, particularly for underrepresented minorities and women (Schweer et al., 2011). By recognizing the value of these social relationships and providing opportunities for faculty to gather, the Network has been able to harness social capital and enhance scholarly collaboration.

Warner et al. (2016) studied faculty promotion and retention by analyzing coauthor networks at Harvard Medical School. Their findings show that the coauthor network metrics provided insight into faculty advancement and retention in academic medicine. Similarly, Lybarger, et al. (2014) used SNAto evaluate research collaborations and productivity. Their work provides a structured methodology and a model for others to follow when evaluating the connections formed by faculty members and determining if these connections lead to increased productivity.

Overall, providing collaborative environments and providing networking opportunities are pivotal strategies to enhance faculty collaboration and, ultimately, faculty success. Using network analysis provides a mechanism for determining if the strategies are producing results.

Social Network Analysis

Social networks are composed of relationships within and between individuals, groups, and organizations. Social network analysis (SNA) is a set of techniques used to identify and measure those relationships by studying both the whole network and connections between individuals (Honeycutt, 2009). Wasserman and Faust (1994) explain network analysis as focusing on both the interchange of resources and the patterns of interactions between persons. In SNA terminology, interactions are viewed in terms of "nodes" and "edges"—nodes being the individual actors within the network, and edges being the interactions between and among actors. This research utilizes the techniques of SNA to answer the posed research questions around how the overall structure of the Network is changing, how collaborative groups are forming, and what impact these connections are having on student retention and faculty productivity.

SNA is applied within education research in many diverse ways. For example, Miller (2020) used SNA to assist in detecting culturally defined organizational categories to clarify the definition of comprehensive institutions. Baker (2018) used SNA examine students' consideration of groupings of majors. His research provides guidance for colleges and universities about factors to consider when forming guided pathways and meta majors at their institution. Lukacs and David (2019) considered how students' personal networks became unstable in the process of college transition. They used social network analysis to study a group of Roman students and found significant differences between students in their reliance on certain groups in the process of academic adjustments.

Gonzalez Canché (2019) recently advanced the use of SNA in education research by bridging geographical and social network analysis to statistically model structures with education data and showed how to reveal meaningful structures in qualitative data with these methods. Gonzales Canché (2018) used geographical network analysis to study student migration in U.S. higher education and found that states that attract more non-resident students tend to export their own resident students to other states and that high school students are less likely to migrate out of state the more financial aid they receive within their own state. Gonzales Canché and Aguilar (2015) applied SNA to institutional data from Calizona Community College to study the effects of peers and credit attainment on underrepresented minority students in community colleges. They found that male Latino and male African American students benefited from interacting with peers in the same racial/ethnic group with higher amounts of credits accumulated.

Several researchers have also used SNA methods to analyze student retention. Eckles and Stradley (2012) determined relationships in a network by using archived data. They found that the retention of students' friends had a greater impact on their retention than did the performance variables commonly believed to be associated with retention. Almeida et al. (2019) utilized SNA to study social capital in first-generation students' academic success. They found that, for this set of students, social capital with faculty and staff predicted grade point average. Poldin, Valeeva and Yudkevich (2016) studied how the achievements of students are influenced by the achievements of peers in their social network. They discovered that this peer influence happens chiefly through relationships, such as study partners that share knowledge, and not as much through mere friendship connections. Another group (Berthelon et al., 2019) that studied peer networks found that peer quality improves student performance and that the breadth and cohesion of a student's network positively affects a student's outcomes.

This research seeks to use SNA to analyze the network structure over time with multiple statistics to study how the Network is growing and changing. This study explores who is involved, how they are connected, and how this involvement may impact faculty careers or student persistence.

METHODS

The methods section will explain how the data for this study was collected and analyzed.

Data Collection

The Network research team developed a survey to study the composition of the network each year. The Network's research and leadership team assessed the survey for its face validity before revision and redistribution to the entire network. Surveys were distributed by email to any individual involved in the Network. Data were collected for five consecutive years, from 2018 to 2022. Surveys were completed online through the Qualtrics survey application. An IRB (institutional review board) approved consent form was provided at the beginning of the survey. Respondents then submitted basic demographic information such as name, organization, and role (student, faculty, administrator, etc.). Lastly, they were asked to name other individuals that they collaborated with on projects related to the Network. Because "collaboration" can mean very different things to different people, and the degree to which two people collaborate on projects differ, a numeric classification for levels of collaboration was provided to help participants understand the meaning of each term. The following scale developed by Hogue et al. (1995) and Borden and Perkins (1999) was used:

1) Networking: Aware of organization, little communication, loosely defined roles, independent of decision making.
2) Cooperation: Share information, formal communication, somewhat defined roles, independent decision making.
3) Coordination: Share information frequently, defined roles, somewhat share decision making.
4) Coalition: Frequent communication, shared resources, shared decision making.
5) Collaboration: Frequent communication, shared resources, mutual trust, coordination on most or all decision making.

Over the five years of data collection, 249 individuals either responded or were named by someone who responded to the survey. One limitation of social network data collected through virtual surveys is that responding to the survey is entirely voluntary and moderately time consuming. In order to overcome these limitations it was decided to include any individual named by any respondents in the analysis even if they did not fill out the survey themselves. The researchers and project leaders felt this method provided a more honest depiction of the network than only using people who filled out the survey and cutting off any of their named collaborators who did not fill out the survey.

Network members consisted of undergraduate students, graduate students, student advisors, industry members, K-12 teachers, university staff members, faculty members, researchers, and administrators. The organizations that these network members represented were also very diverse including colleges and universities, companies, state-level educational agencies, county school systems, nonprofits, and state research organizations. In addition to the survey data, student persistence data was collected from the largest institution in the Network. A list of publications related to the Network was also collected to form a publication network.

Network Construction

To explore the structure of connections between Network members, survey responses were converted into an adjacency matrix. The adjacency matrix turns network members into nodes and the connections between members of the network into edges. The reported strength of collaboration forms the edge weight in the network. Individual cells in the adjacency matrix, a_{ij}, represent a weighted edge between two collaborators i and j. If two survey respondents name each other in the survey, but they respond with a different strength of connection, then the average reported strength is used in both cell a_{ij} and cell a_{ji}, as such, the resulting matrix is square and symmetric, and the corresponding graph is undirected. Consequently, if one person names a network member that did not fill out the survey, or did not name that person in their survey, then the reported strength of connection is averaged with 0. This allows for network members who did not fill out the survey to still be accounted for in the analysis, but with a lesser weight than two people who both filled out the survey. Using the exact reported strength of connection would create a directed adjacency matrix and would have the potential to reveal more salient and complex information about the network structure but would require a higher survey response rate for the direction of the edges to be meaningful.

Another network was constructed from the list of academic publications written by Network members about results generated by Network activities. Individuals who appeared in one or more publications were included as nodes and edges were formed between two individuals if they were coauthors on at least one of the papers. This resulted in an undirected graph. Edges were weighted by the number of papers two individuals collaborated on.

Network Statistics

Yearly network statistics were calculated after the adjacency matrix was formed. The number of nodes (people), the number of edges (collaborative connections), the graph density (the ratio of the number of edges in the network to the maximum possible number of edges), and the number of surveys completed was calculated.

Because a network is composed of a set of nodes and edges with any number of configurations, when nodes or groups of connected nodes are not connected to other parts of the network, they form isolated "islands." These islands are called components. The component with the largest number of nodes is called the giant component. The size of the giant component and the number of components are both reported for this work. For a network changing in time, these two statistics can provide an indication of the growth of the main body of the network and how/if the network is splintering into smaller groups.

Two different centrality metrics were calculated in order to measure the connectedness of individuals to the larger network. These two metrics are strength (weighted degree) and betweenness centrality. Strength is a local centrality measure that shows how connected a node is to its immediate surroundings within a network. Strength is found by adding the weights of all edges connected to an individual node. Betweenness centrality is a global centrality measure that accounts for a node's position relative to each other nodes' positions in the network. Betweenness centrality for a node v is found by computing the shortest path between each pair of nodes, finding the fraction of shortest paths that include the node v for each node pair, then summing this fraction over all node pairs (Freeman, 1977). Betweenness centrality scales with the number of node pairs in a network, so the statistic is normalized by dividing by the number of node pairs not including v itself; $\frac{(N-1)(N-2)}{2}$ for undirected networks. Including both a local and a global measure should capture most of the relevant information about the significance of each node in the structure of the network. Each node has its own strength and betweenness centrality and these are reported for the top actors in the network. The average strength and betweenness centrality are also computed for the entire network. A person with a higher betweenness centrality is someone who is the bridge between unconnected network members, whereas a person with a higher strength is someone who simply has many connections. Betweenness can also be indicative of information flow; information tends to flow through nodes that bridge otherwise unconnected nodes. People with higher strength tend to have higher betweenness centralities, and vice versa.

Maximal clique analysis was used to determine the close-knit groups forming in the network. A clique is a group of directly connected individuals within the larger graph, such that each pair of individuals in the clique has an edge connecting them. A clique is maximal if it includes the largest subgroup of individuals where everyone in the subgroup is connected. People can belong to more than one clique. The Bron and Kerbosch Algorithm (Bron & Kebosch, 1973) can be used to find all maximal cliques of each possible size along with the cliques' members. A clique strictly measures groups of individuals that each reference each other as a collaborator.

Because subgroup structures are often more complicated than cliques, another group measurement, community detection, was also employed to look at subgroups. For example, an institution could have a committee of ten faculty members. In a group of that size, there is a reasonable chance that not every member of the group would collaborate with every other member, and they would not be counted as a group in a clique analysis, even though they could meaningfully be defined as a collaborative group. A community detection algorithm (CDA) was used to determine the grouping structure of the network beyond the strictly defined, well connected cliques. A community is a set of nodes such that pairs of nodes in the set are more likely to be connected if they are both members of the same community than if they were members of different communities. Since the clique analysis allowed for overlap, a CDA was chosen that did not allow for overlap in communities to determine if the network divided naturally into groups that were more connected internally than they were connected externally. The communities were then analyzed for similarities between network members. The fast-greedy CDA was used for this work (Clauset et al., 2004).

RESULTS

In this section, the Network's evolution using several SNA tools is examined. Student persistence and faculty productivity of Network members are also considered. It should be noted that there is no suggestion of causality, but only presentation of information about persistence and productivity.

Network Structure and Evolution

The Network changed much over the five years. The changes are shown and discussed using basic statistics, centrality measures, community structures, and network attrition.

Basic Network Statistics

Table 1 shows several statistics associated with the Network. The number of individuals responding to the survey increased every year, however the response rate to the survey decreased as discussed in Section 3.1. In 2020, a website was added for the project and the number of people who signed up for the website also grew each year from 584 (2020), 736 (2021), to 934 (2022). While many people were interested in the project, not all of them actually engaged in the activities of the project. From the survey responses, people were added to the "active" Network in two ways, either they filled out the survey, or they were named by someone who filled it out. Table 1 demonstrates that the number of connected members (nodes) also grew each year and likewise, the total number of members to date continued to grow. From 2018 to 2021, the network was growing at relatively commensurate rates in all metrics, but in 2022 the number of connections (edges) drastically decreased. In 2022, network members were reporting fewer connections with lower collaboration levels relative to previous years. The average strength of network members is one of the most robust ways to test the connectivity of a network because it is less susceptible to the N^2 effect of edge density calculations, where smaller networks tend to have higher densities. The density, the ratio between the actual number of edges and the total possible number of edges, is included to illustrate this point.

From 2018 through 2021, the evolution of the Network was dominated by student growth, as shown in Figure 1. This changed in 2022 when the number of students in the network marginally increased, but the number of other types of network members substantially increased. The growth of active network members has been relatively constant, with an average of about 19 new members joining the Network each year. Each year, new members were added while older members left or disengaged. Once a person was an active part of the network and a number was assigned to them, they kept this number even if they became inactive and were not part of the node count in subsequent years. From 2019 to 2021 the number of non-student active network members changed negligibly. However, many non-student network members were joining or leaving the network, leading to an equilibrium of active non-student members.

Centrality Measures

Figure 2 shows the Network colored by members' self-reported roles. The position of each node in Figure 2 (and in all future network graphs) is related to the strength of connection to other nodes in the network. The Fruchterman-Reingold force-directed layout algorithm

(Fruchterman & Reingold, 1991) is utilized to place nodes closer together that have higher edge weights. Survey respondents are prompted by the question "Which of the following roles most accurately describes your role in the network?" They were provided with many options for their potential role, but for visualization purposes, the options in the survey were combined to the four roles displayed in Figure 2. Network members sometimes reported different roles from year to year, depending on which role most accurately described them that year. For example, for the first three years, node 26 was part of the "Government/Industry contact role" because they worked for a state research organization, but by 2021 they identified as a Faculty Member/Lecturer/Teacher, and by 2022 they were put in the other category because they reported that they were now an administrator.

In Figure 2 nodes are sized proportional to network members' strength of connections, but the scaling factor is small to show the change in the network relative to the most prominent Network members. These graphs make it clear that although students make up the dominant growth in the network in terms of new members, the large-scale structure of the network is dominated by non-student members. These networks also show some of the structural changes that occur when very connected network members leave the network. The loss of members like 27 and 82 could, to some extent, help account for the decreasing network metrics in 2022.

Table 2 illustrates the impact on network structure that members 27 and 82 had relative to other top network members. S indicates strength, SR strength rank, B betweenness, and BR betweenness rank. Strength rank and betweenness rank represent the rank of an individual for a given centrality metric for a given year. For example, SR of 1 means that the individual had the highest strength for the entire year; BR of 5 means that individual had the fifth highest betweenness value for that year. The other top network members, 8, 26, and 23 were chosen by their strength rank; from 2018-2022 they consistently had some of the highest ranks, usually top 5. Top network members were chosen by SR and not BR because the top members by betweenness had more year-to-year variance. Table 2 makes clear that individuals who control the most unique information flow between connected groups (high BR) are often not the same individuals who simply have the most connections (SR). Cells in Table 2 are labeled NA when a network member was not in the Network that year, either because they had not joined yet, or because they left the Network entirely.

Figure 3 is colored similarly to Figure 2, but it is sized by betweenness to show network members that are pivotal in retaining the structure of the network. Losing a network member with high betweenness is likely to split the network into smaller components, taking much of the possible information transfer with them. This graph clearly shows that the network is held together by members with various roles. Figure 2 emphasizes the dominance of faculty members in terms of the total number of connections in the network. Figure 3 shows that students, government/industry contacts, and other members (i.e., administrators and non-profit employees) keep the network together as generally as a single component rather than separate, internally well-connected components. Network members with high betweenness act as major information distributors within the network. For example, students like 61 in 2019, 99 in 2021, and 146 in 2022 are connected to many other students who are not connected to anyone else, so they seem to distribute information to these otherwise unconnected students.

Community Structure

Referring to Table 1, it can be seen that the giant component size of the network was identical to the total network size until 2020. In 2020 and 2021 the network split into four components. In 2020 all three of the new components were of size two, while in 2021 one component was size four, one was size three, and one was size two. In 2022 many isolated sub-groups split off from the main network, resulting in 11 components, all of which, outside of the giant component, were of size two. The number of components and the size of the giant component relative to the number of nodes in the network give some indication of the possibility of information flow in a network. Individuals who are part of the isolated components are less likely to be a part of, or even know about, many of the Network's activities and events.

The giant component and the communities resulting from the application of the fast-greedy community detection algorithm are shown in Figure 4. Nodes are colored by community membership and a shaded region is included over each community. Nodes are roughly sized by strength, with a minimum size threshold to make node color visible. Shaded regions overlap because of the force-directed graph node placement algorithm, but community membership does not actually overlap. Edges between communities are colored red and edges within communities are colored black. The CDA was applied only to the giant component because the maximum size of any other component in any year was four, so each component was its own community, and the community detection provided no new information for these smaller components. Communities that resulted from the application of the CDA to the networks were analyzed for institutional homogeneity. This gives a broad picture of the network, indicating which years had greater levels of inter-institutional collaboration and which years had collaborations forming within individual institutions.

In 2018, out of the six communities identified, the four smallest were almost entirely institutionally homogeneous, while the two largest communities had very little overlap in institutional membership. Out of the six communities in 2019, the smallest two communities were mostly composed of members from one or two institutions, while the other four communities were much more heterogeneous than 2018. In 2020, except for the largest community, the others were institutionally homogeneous. In 2021 the largest community, shaded red and colored orange, actually included members mostly from one institution, but was less central in the network than other smaller communities that were much more institutionally diverse. In 2022, all eight of the communities, except for the largest and most central one, were each composed of members from single institutions. In the years 2018, 2020, and 2022, the CDA identified groups of individuals from the same institution that were more connected to each other than to other institutions. In 2019 and 2021, the grouping structures identified by CDA seemed to be much less dependent on the community member's institution, indicating a higher level of cross-institutional collaboration, and information transfer, during those two years.
with them, so many of the connections that exist each year were made that year. The way the survey is administered and the data is analyzed helps reflect the true membership of the active Network. Thus, the turnover in the Network better reflects people who are not active during the year and not just people who forget to fill out the survey.

Network Attrition

Many different individuals join and leave the network each year. The composition of the network changes by approximately 30-50% every year. However, over 50% of individuals from each previous year remain in the network. Over time this flow of network members does lead to the network changing drastically, with only about 35% of the individuals that were in the network in 2018 remaining in the network in 2022. The change in network members is directly proportional to the change in network connections; when network members leave, their connections are taken.

Student Retention

Over the course of the five years of the project, institutional data was available for 52 students of the 107 total students. These students were from the largest institution in the Network and this data were not readily available for students from the other institutions. The institutional data were compared to the Network data. Student retention was examined one year after each student participated in the network survey. Thirty-nine students entered the network with a STEM major declared and were retained in a STEM major one year later. Two students entered the network with a STEM major declared and switched to a STEM adjacent major; both of which switched to the Environmental, Soil, & Water Science major. Eleven students entered the network with a non-STEM major. Ten of these students were still not within a STEM major one year later, two of these 10 students started STEM-adjacent and remained STEM-adjacent one year later, and one student switched to a STEM major from a non-STEM major within a year of entering the Network. One hundred percent of students who entered the network with a STEM or STEM adjacent major were retained in STEM or STEM adjacent majors one year later. At this same institution on average, 86% of students are retained in STEM one year after declaring a STEM major.

Figure 5 shows connections between student members in the Network from 2018 to 2022. Students who were only connected to non-student members are included as isolated nodes. Node size is proportional to the student's strength in Figure 2 to show which students were more connected to other students and which were more connected to non-student network members. For example, node 3 in 2018 is larger than the nodes in 2018 that actually have connections because node 3 is connected to more non-student members than the other students in 2018.

Early in the project, student connections relied heavily on connections to faculty members. In 2020 and beyond, students seemed to form larger intra-institutional groups. The large drop in connectivity evident in the total network in 2022 is less evident here. It seems that student connections in 2022 either decreased or increased by institution, with the light blue and burgundy institutions having smaller student subgroups, and the red institution having larger student subgroups. On average there is still a decrease in average student connectivity, which can be seen in the number of tiny, isolated nodes in 2022 compared to 2021. These individuals are connected only to faculty members.

Faculty Collaboration and Productivity

Maximal clique analysis was applied as another way to examine the structure of the network to quantify the number of closely-knit collaboration groups. These cliques are formed from members who are all connected to each other. These close-knit groups consisted mainly of faculty members and, in the early years, were developing around the working group structures that had been set up by the Network. The cliques grew in number from 2018 through 2021 and the average size of a clique grew each year during this time, showing that more people were working more closely in larger groups (see Table 3). It should be noted that the number of cliques can be larger than the number of nodes because they represent different combinations of connected nodes. The clique analysis is similar to that of the communities as it shows close-knit groups that are forming; however, the difference is that in clique analysis, groups are studied where everyone is connected to everyone else. These groups tend to be more like working groups than the communities do. Recall, a community is a set of nodes where pairs of nodes are more likely to be connected if they are both members of the same community than if they were members of different communities.

In 2018, the largest cliques were of size six and there were 10 cliques of this size. All these largest cliques in 2018 were made up of individuals from different organizations, with either six, five, or three different organizations represented in various cliques of size six. In 2019, the largest clique was size seven and there were three this size and 29 cliques of size six. The largest cliques in 2019 were made up of people from seven different organizations. In 2020, the number of large cliques of size six and seven again increased to 46 of size six and seven of size seven. The large cliques were very diverse again, containing people from seven organizations. These cliques were mostly made up of faculty from different colleges and universities but also contained staff of research organizations or state education organizations. The pandemic years did not have much effect on the growth of the cliques, since most groups met on an online platform already because they were from all over the state. In 2021, there continued to be an increase in close-knit groups. There were 92 size-six cliques, 20 size-seven cliques, and two size-eight cliques. The largest cliques in 2021 were again made up of five and six organizations.

In 2022, the Network structure changed, and the cliques dropped off dramatically, both in number and in size. Examining the cliques for 2022, the largest are size five and there are only two of these. The two large cliques were made up of only project leadership and not composed of faculty at different institutions. The size-five cliques went from 241 in 2021 to two in 2022. The size four cliques went from 384 in 2021 to 24 in 2022 and even the size three cliques went from 403 in 2021 to 99 (around 25 percent) in 2022. Of all the measures of the Network, the number and sizes of cliques reveal changes in the group structure of the network.

An alternate measure of the connectivity of the Network was developed by examining the network of academic publications, see Figure 6. This graph represents 25 publications (specifically written about Network related topics) and 44 authors, with 30 faculty and graduate students from the Network and the other 14 faculty and graduate students from outside the Network. There are 148 edges, indicating a minimum of 148 instances of publication-based collaboration. In the figure, the Network members are colored in blue, while non-network members are in red. The nodes represent people, and the connections represent whether they are coauthors on a paper. The thickness of the edges corresponds to the number of co-authorships shared between a pair of nodes. The size of each node corresponds to the number of publications of each individual.

The publication network naturally divides itself into several distinct clusters. The clusters are also connected to other clusters, showing that some members work with different groups. There is one isolated node, indicating a sole author publication. The non-members were typically graduate students who were working in a group with a faculty member who was part of the Network. These members also show that there exist individuals outside of the network that are influencing the Network. Non-network members participating in publications using data from the network or from students who participate in the network are influencing the structure of the main network by influencing network leaders’ policy decisions with their data analysis. It is clear that some of the members of the Network (22, 43, 118) are very productive (have many publications) and that they are working with others in the Network to coauthor papers. The numbers on these nodes correspond to the numbers in the overall Network graph, so when examining the nodes in the publication network, it can seen that node 8 and node 26 also appear in Table 2, indicating that they were both also prominent members of the main Network.

Precisely quantifying the edge overlap between the publication network and the main set of networks is difficult because the main network breaks down connections by year while the publication network includes papers published any year during the network's existence. However, any edge that exists in the publication network between two network members also exists at some point in time in the main network because publishing is a type of collaboration. Figure 6 is not included to show unique connections between faculty members not captured through the main network, it is included to reveal the nature of a single type of connection within the network. Publications represent particularly active network members who are collaborating through the peer-reviewed publication process rather than through informal networking or formal committees within the Network.

For the individuals with the highest degree in the network of publications, there was an examination of their strength and betweenness scores in the Network in 2022. For the top ten individuals in the network of publications, seven had Network scores in the 2022 network, the others had either left the Network by 2022, did not take the survey, or were not named. The average strength for this group was 25 and the average betweenness score for this group was 582. This was much higher than the overall average in 2022, which was a strength of 7 and an average betweenness of 186. This shows that these individuals in the network of publications are highly connected and influential in the main Network.

DISCUSSION

This section will answer the three main questions posed in the introduction.

Research Question 1 – How is the overall structure of the network changing?

The overall structure of the Network changed from year to year but had stable growth with more people joining than leaving each year; the number of active members grew every year. The number of connections and the average strength grew each year from 2018 to 2021; overall, Network members were gaining connections and/or strength of connections. There was a reduction of both the number of edges and strength in 2022, which appear to be the result of changing the structure of the working groups. The giant component grew every year, which implies that network members were not isolated in small groups but were more centralized. This could be due to the Network holding bi-annual conferences for all members and also having leadership meetings where many people were invited to present monthly.

The Network composition was dominated by student growth from 2018 until 2021. In 2022, network growth was dominated by non-student members and student numbers remained relatively constant. In the first few years of the network, student connections were primarily to faculty members. However, since 2020, robust groups of student clusters formed, largely within individual institutions, mainly due to practices that the Network put in place including student campus clubs and student leadership groups. The drop off in network size and connectivity in 2022 did not impact student clustering in the Network as much, likely because these connections were made within individual institutions.

Although students played a major role in increasing the number of yearly Network members, non-student members were consistently the most connected and influential members of the network. It was also found that information flow in the network was not necessarily dominated by the most connected members, but by a mix of well-connected members and members who acted as bridges between well connected groups in the network. Students, faculty members, government/industry contacts, and other roles like administrators or K-12 educators were all found as important bridges connecting less connected groups in the network. Regardless, when particularly connected members left the network, their absence was clearly felt in the Network structure. The loss of a few very connected members from 2021 to 2022 likely had at least some effect on the decrease of network statistics in 2022.

From the inception of the Network, different groups had varying roles. The faculty members' role was mainly to develop, collaborate, and disseminate best practices of engagement and retention for STEM students. The students were to provide insight into student needs through student voice on all major committees and groups. The non-profit (other) partners provided oversight and management of student resources, assisting with paying students for any work being undertaken by them and keeping them engaged. The government labs and industry members were to provide input about workforce needs in STEM-related organizations and provide connections to agencies and companies. Thus, as the connections between these groups in Figures 2 and 3 are considered, it can be seen that the faculty members are the leaders in strength (connections) and in betweenness (being on more paths between pairs of node) for the first four years, until the policy changes and then the non-profit (other) members and undergraduate students appear as leaders in these areas in 2022. There was one major faculty member who left the network in 2021 (node 27) and this departure along with the policy changes seemed to facilitate change in the appearance and distribution of the graph (Figure 2 and 3).

One interesting aspect of network change that is not obvious from the structure alone is that Network members join and leave each year. The total composition of the network changes by about 30-50% each year. By 2022, only 35% of Network members remained that started in the network in 2018, but the total number of Network members increased every year. This highlights the flexibility of many of the interactions and characteristic roles that Network members take; collaboration still increased even when many Network members were replaced each year.

Research Question 2 – What collaborative groups are forming? What are their compositions?

The communities resulting from the application of the community detection algorithm revealed some interesting features in the Network. In general, the largest, most connected communities included the most connected individuals in the Network and included a diverse number of Network members with many different roles and from many different institutions. The smaller communities were less centrally connected to the Network and were most frequently composed of Network members from one or two institutions. The CDA identified a natural grouping structure in 2018, 2020, and 2022, where, more often than not, individuals in the same community were from the same institution, whereas in 2019 and 2021 that grouping structure appears much less frequently. Grouping by institution is reasonable in 2018 because the network was so new that most individuals that worked together already knew each other from their own institutions. The less central, more institutional grouping structure of 2020 could be seen as an effect of COVID, where there was some sense of necessity in institutional grouping to figure out how to deal with the pandemic. In 2022 this grouping structure could be attributed to the development of institutional teams, which encouraged working within the institution to implement some of the strategies and programs developed by the Network in the previous four years.

When looking at groups from the clique analysis perspective, from 2018 through 2021, there were working groups formed around the key goals of the project including faculty/student engagement, industry connections, student readiness for college, and summer immersive research experiences to come up with ideas to make improvements in education in the state and in different colleges and universities. During these years, the Network stressed working groups around the core goals of the network and developing best practices to share across organizations.

In 2022, the last year of a five-year grant, the leadership of the Network started focusing on the sustainability of changes at each organization. They discontinued most working groups and changed the focus and funding to institutional teams at each college and university. They also had leadership teams established around different areas such as sustainability and improvement science to support institutional team research efforts. Until 2022, collaboration was growing among network members and between network organizations, but due to some conscious policy changes, the inter-institutional working groups were replaced by institutional teams and thus the collaboration between institutions was reduced in 2022. The Network made a conscious decision to change the way groups worked together in 2022. They moved away from working groups around project goals, where working group chairs were receiving funding, to working groups within institutions, where institutions were getting funding, and the groups focused on sustaining the work at their institution beyond the initial funding of the grant. These decisions had an impact on the structure of the network. These changes strongly affected the community structure and the number and sizes of the cliques. In the clique analysis, it was found that from 2018 to 2021, the largest groups were composed of inter-organizational groups. However, in 2022, most of the fully connected groups were smaller and everyone in the group was from the same organization.

Research Question 3 – What impact do these connections have on student retention and faculty productivity?

When examining student retention, it can be seen that the Network provides at least three out of four of Tinto's formal academic systems for student network members. Many students interact personally with faculty and staff at their institution that would not have done so otherwise. The Network also provides formal and informal social interactions with peers through various student groups. Aligning with Astin's theory of student involvement, students involved in the Network actively engage in extracurricular activities through student events. They also form stronger bonds with faculty members that are part of the Network. Students are encouraged to participate in these activities, but their involvement is often entirely voluntary. Students' voluntary output of energy participating in and making connections through these programs leads to greater personal development and an enhanced learning experience throughout college.

One hundred percent of the 52 students at the largest institution were retained in STEM or STEM adjacent majors one year after joining the network, compared to an 86% one-year retention rate for the average student after they declared their STEM major at the same institution. This high rate of retention could be attributed to a number of factors. The Network provided small monetary stipends for students participating in research activities. Also, along with the research requirements with the stipends, students were required to attend Network meetings, meetings with their campus clubs, and present their research or present about the Network to their home high schools, to their home institution, or to the state legislature. These requirements kept the students in close contact with other students taking similar courses, or who had already taken similar courses, and with similar backgrounds (often rural/first-generation). The sense of community developed from being in contact with other students on their campus through the Network, as well as the stipend provided with their research likely gave the students more motivation and support to stay in their STEM major.

When examining faculty productivity, it can be seen that the Network facilitated communication and networking among its faculty members allowing them to build social capital. Aligning with Hampton and Parker (2011) who promoted face-to-face interaction as being vital to success, the Network held bi-annual meetings, bringing together faculty, government, and industry members to work and discuss projects and methods for STEM persistence. This seems to suggest, as Hampton and Parker (2011) found, that faculty members with more connections are more productive. The facilitation of working groups and regular times to meet during the year leads to groups that are able to share ideas and collaborate at a higher level. The Network also facilitated monthly meetings of key research groups. They invested in virtual platforms for meeting and sharing information. Leadership on the project consisted of people from many organizations and they invited people from all agencies to monthly leadership meetings to share and provide input. Considering the methods of Lybarger et al. (2014) provided a way to show evidence of faculty productivity (based on publication record) within the Network. By examining the publications generated over the five years of the project it was observed that the Network faculty members who were the highest contributors in the network of publications also had high strength and betweenness scores in the main network. These individuals were collaborating more and positioned between many members of the Network; thus, they were in the communication flow of the Network.

By examining the cliques and communities that were forming within the Network it was noticed that the cliques and communities continued to grow and change from 2018-2021. The cliques were growing in size and in number during these years as faculty members were forming inter-institutional working groups. In 2021, there were 241 size five cliques, meaning there were many close knit groups that were collaborating. The community structures also showed similar growth up through 2021 indicating cross-institutional collaboration. These structures within the graph should be indicative of information transfer between different programs. Starting in 2022, the number and size of cliques declined and the community structure became more intra-institutional. The shift in 2022 could indicate that the main collaborations were occurring within institutions and could be indicative of institutions attempting to implement strategies and programs developed by the Network.

IMPLICATIONS

Over its five years of activity, leaders of the Network constantly studied and adapted Network policy to find the best methods to promote collaboration across the state, which, in turn, bolstered student STEM retention. One of the most successful policies that was adopted early on was including student voice to all activities from major committees to planning and attending Network convenings. They also facilitated student clubs and funded student research as a means to keep students engaged. The early use of Networked Improvement Communities arranged around major goals of the project helped many strongly connected subgroups and communities to form within the larger network. These close-knit groups provided a means for students to stay connected and become leaders and a means for faculty to work productively on research topics. The use of an analysis tool, such as a social network analysis, has been an effective tool for project leaders to provide insight into how the structure of the Network was growing and changing and to capture how the students and faculty members were connecting across the years.

CONCLUSIONS

In this study, data were collected from or about 249 individuals (faculty, students, statewide educational organization staff, research organization staff, and industry staff) over a five-year period about the network connections forming within a project in a rural state to promote the retention of students in STEM majors. Social network analysis showed that the structure of the Network changed over time. For the first four years, the network connections increased in number and strength (more edges with higher weight on them) (see Figure 2) and the close-knit groups (cliques) grew in number and size (see Table 3). The betweenness score also increased for key individuals, showing the emergence of key leaders in the project who were disseminating information to and collaborating with many others (see Figure 3). The fifth year brought many changes to the project and, consequently, changes to the Network. The structure of the Network changed as leaders began to transition the project to sustainability within state academic organizations. The Network's edges, density, average strength, average betweenness, and number and size of cliques all decreased. The number of components increased, indicating a small fracturing of the central network.

For the first four years, the project focused efforts on Networked Improvement Communities (Russell et al., 2019) within the Network as part of the overall plan for using an iterative process to implement, study, and revise change ideas in key areas addressed by the Network. During this time, the project invested in forming social relationships. Through the lens of Social Capital Theory (Lin, 2017), the Network was putting resources in areas to enhance the outcomes of the project. The payoff of the investments (funding NICs and statewide conferences) was the growth in the connections formed throughout the state. The Network provided many opportunities through online working groups, leadership meetings, and bi-annual face-to-face conferences (when possible) to facilitate networking opportunities for faculty. The findings of Hampton and Parker (2011) about providing face-to-face opportunities to help groups form and synthesize information and the findings of Schweer et al. (2011) related to using networking as a way of fostering inclusion of underrepresented groups are supported by the positive effects of the many networking opportunities given to faculty from both large and very small institutions throughout the state. There is evidence provided by the publication network (see Figure 6) that faculty productivity appeared to be happening through collaborations formed within the Network.

Tinto's Model of Institutional Departure (Tinto and Cullen, 1973; Tinto, 1973) and Astin's Student Involvement Theory (Astin, 1984) provide a context when examining student data. Tinto claims that the primary reason college students withdraw from their institution is that they fail to become integrated into communities, they have difficulties with academics, and they have difficulties with their career choices. Rural students may place a higher value on being in a community than their non-rural peers (Howley et al., 2014; Wright 2012). The Network has provided a way to connect to a community of other students, professors, and even statewide industry partners. This provides students with a community at their institution and in the state, a way to navigate academic challenges, and a path into their chosen STEM fields. According to Astin's Student Involvement Theory (Austin, 1984) the quantity and quality of energy that students invest into their college experience is directly proportional to the amount of personal development and learning a student experiences in college. He suggests that student involvement should be the primary metric by which effective educational practices or policies are measured. The Network provided multiple ways for students to be involved and to take leadership positions.

REFERENCES

Almeida, D. J., Byrne, A. M., Smith, R. M., & Ruiz, S. (2019). How relevant is grit? The importance of social capital in first-generation college students' academic success. *Journal of College Student Retention: Research, Theory, & Practice, 23*(3), 539-559. https://doi.org/10.1177/1521025119854688

Astin, A. (1984). Student involvement: A development theory for higher education. *Journal of College Student Development, 40*, 518–529.

Astin, A. (1997). How "good" is your institution's retention rate? *Research in Higher Education, 38*, 647–658. https://doi.org/10.1023/A:1024903702810

Baker, R. (2018). Understanding college students' major choices using social network analysis. *Research in higher education*, *59*(2), 198–225. https://doi.org/10.1007/s11162-017-9463-1

Baker, S., & Pomerantz, N. (2000). Impact of learning communities on retention at a metropolitan university. *Journal of College Student Retention: Research, Theory & Practice*, *2*(2), 115–126. https://doi.org/10.2190/62p5-cq2u-ntuw-dm1c

Berthelon, M., Bettinger, E., Kruger, D. I., & Montecinos-Pearce, A. (2019). The structure of peers: The impact of peer networks on academic achievement. *Research in Higher Education*, *60*(7), 931– 959. https://doi.org/10.1007/s11162-018-09543-7

Boggs, G. R., Dukes, C. M., & Hawthorne, E. K. (2022). Addressing the STEM Workforce Shortage. U.S. Chamber of Commerce Foundation. https://www.uschamberfoundation.org/education/addressing-stem-workforce-shortage

Borden, L. M., & Perkins., D. F. (1999). Assessing your collaboration: A self-evaluation tool. *Journal of Extension*, *37*(2), 67–72. https://archives.joe.org/joe/1999april/tt1.php

Bowman, N. A., & Culver, K. (2018). When do honors programs make the grade? Conditional effects on college satisfaction, achievement, retention, and graduation. *Research in Higher Education*, *59*, 249–272. https://doi.org/10.1007/s11162-017-9466-y

Bron, C., & Kebosch, J. (1973). Algorithm 457: Finding all cliques of an undirected graph. *Communications of the ACM*, *16*(9), 48–50. https://doi.org/10.1145/362342.362367

Burke, A. (2019). Student retention models in higher education: A literature review. *College and University*, *94*(2), 12–21.

Burt, R. S. (1992). Structural Holes: The Social Structure of Competition. Harvard University Press.

Clauset, A., Newman, M. E. J., & Moore, C. (2004). Finding community structure in very large networks. *Physical Review E*, *70*(6). https://doi.org/10.1103/physreve.70.066111

Coleman, J. S. (1988). Social Capital in the Creation of Human Capital. American Journal of Sociology, 94(Supplement: Organizations and Institutions: Sociological and Economic Approaches to the Analysis of Social Structure), S95-S120.

Cunningham, A. F., Erisman, W., & Looney, S. E. (2008). Higher education in Michigan: Overcoming challenges to expand access. *Institute for Higher Education Policy*.

Digby, J. (2005). *Peterson's smart choices: Honors programs and colleges (4th ed.)*. Thomas Peterson's.

Eckles, J. E., & Stradley, E. G. (2012). A social network analysis of student retention using archival data. Social Psychology of Education: An International Journal, 15(2), 165–180. https://doi.org/10.1007/s11218-011-9173-z

Freeman, L. C. (1977). A set of measures of centrality based on betweenness. *Sociometry*, *40*(1), 35–41. https://doi.org/10.2307/3033543
Fruchterman, T. M. J., & Reingold, E. M. (1991). Graph drawing by force-directed placement. *Softw. Pract. Exp.*, *21*, 1129-1164. https://doi.org/10.1002/spe.4380211102
Gonzalez Canche, M.S. & Rios-Aguilar, C. (2015). Critical Social Network Analysis in Community Colleges: Peer Effects and Credit Attainment. *New Directions for Institutional Research.* 2014:75-91. https://doi.org/10.1002/ir.20087
González Canché, M.S. (2018). Geographical Network Analysis and Spatial Econometrics as Tools to Enhance Our Understanding of Student Migration Patterns and Benefits in the U.S. Higher Education Network. *The Review of Higher Education* 41(2), 169-216. https://doi.org/10.1353/rhe.2018.0001
González Canché, M.S. (2019). Geographical, Statistical, and Qualitative Network Analysis: A Multifaceted Method-Bridging Tool to Reveal and Model Meaningful Structures in Education Research. In: Paulsen, M.B., Perna, L.W. (eds) Higher Education: Handbook of Theory and Research. *Higher Education: Handbook of Theory and Research*, vol 34. Springer, Cham. https://doi.org/10.1007/978-3-030-03457-3_12
Hampton, S. E., & Parker, J. N. (2011). Collaboration and productivity in scientific synthesis. *BioScience*, *61*(11), 900–910. https://doi.org/10.1525/bio.2011.61.11.9
Herman, A. (2019). America's STEM crisis threatens our national security. American Affairs, 3(1), 127-148. https://americanaffairsjournal.org/2019/02/americas-stem-crisis-threatens-our-national-security/
Hogue, T., Perkins, D., Clark, R., Bergstrum, A., Slinski, M., & Associates. (1995). Collaboration framework: Addressing community capacity. https://www.thenationalcouncil.org/resources/collaboration-framework-addressing-community-capacity/
Honeycutt, T. (2009). Making connections: Using social network analysis for program evaluation. *Mathematica Policy Research.* https://www.mathematica.org/publications/making-connections-using-social-network-analysis-for-program-evaluation
Howley, C., Johnson, J., Passa, A., & Uekawa, K. (2014). College enrollment and persistence in rural Pennsylvania schools. rel 2015-053. *Regional Educational Laboratory Mid-Atlantic*.
Howley, C., Johnson, J., Passa, A., & Uekawa, K. (2015). College enrollment and persistence in rural Pennsylvania schools.
Lin, N., Cook, K. S., & Burt, R. S. (2017). Building a network theory of social capital. *Social Capital: Theory and Research* (pp. 3–28). Routledge Taylor & Francis Group. https://doi.org/10.4324/9781315129457-1
Luk´acs J, A´., & D´avid, B. (2019). Roma undergraduates' personal network in the process of college transition. A social capital approach. *Research in Higher Education*, *60*(1), 64–82. https://doi.org/10.1007/s11162-018-9503-5
Lybarger, M., Royal, K., Akers, K., & Zakrajsek, T. (2014). Using social network analysis to evaluate research productivity and collaborations. *The Journal of Faculty Development*, *28*(1), 49–58.
McQueen, H. (2009). Integration and Regulation Matters in Educational Transition: A Theoretical Critique of Retention and Attrition Models. *British Journal of Educational Studies*, 57(1), 70–88. https://doi.org/10.1111/j.1467-8527.2008.00423.x

Miller, G. N. (2020). I'll know one when I see it: Using social network analysis to define comprehensive institutions through organizational identity. *Research in Higher Education*, *61*(1), 51–87. https://doi.org/10.1007/s11162-019-09552-0
Pascarella, E., & P. Terenzini. (1991). *How college affects students; Findings and insights from twenty years of research.* Jossey-Bass.
Peri, G., Shih, K., & Sparber, C. (2015). STEM workers, H-1B visas, and productivity in US cities. *Journal of Labor Economics, 33*(S1), S225-S255.
Poldin, O., Valeeva, D., & Yudkevich, M. (2016). Which peers matter: How social ties affect peer-group effects. *Research in Higher Education*, *57*(4), 448–468. https://doi.org/10.1007/s11162-015-9391-x
Reagans, R., & McEvily, B. (2003). Network Structure and Knowledge Transfer: The Effects of Cohesion and Range. Administrative Science Quarterly, 48(2), 240-267.
Russell, J.L., Bryk, A.S., Peurach, D., Sherer, D., Khachatryan, E., LeMahieu, P.G., Sherer, J.Z., & Hannan, M. (2019). The social structure of networked improvement communities: Cultivating the emergence of a scientific professional learning community. *American Educational Research Association Annual Meeting*.
Saw, G. K., & Agger, C. A. (2021). STEM Pathways of Rural and Small-Town Students: Opportunities to Learn, Aspirations, Preparation, and College Enrollment. *Educational Researcher, 50*(9), 595-606. https://doi.org/10.3102/0013189X211027528
Schweer, M., Assimakopoulos, D., Cross, R., & Thomas, R. J. (2011). Building a well-networked organization. *MIT Sloan Management Review*. https://sloanreview.mit.edu/article/building-a-well-networked-organization/
Sims, L. R., & Ferrare, J. J. (2021). "Since I am from where I am from": How rural and urban first-generation college students differentially use social capital to choose a college major. *Journal of Research in Rural Education*, 37(6), 1–21. https://doi.org/10.26209/jrre3706
Snyder, J., & Cudney, E. A. (2017). Retention Models for STEM Majors and Alignment to Community Colleges: A Review of the Literature. Journal of STEM Education: Innovations & Research, 18(3). https://www.jstem.org/jstem/index.php/JSTEM/article/download/2194/1871/
Tierney, W. G. (1992). An Anthropological Analysis of Student Participation in College. *The Journal of Higher Education*, 63(6), 603–618. https://doi.org/10.2307/1982046
Tinto, V. (1975). Dropout from Higher Education: A Theoretical Synthesis of Recent Research. *Review of Educational Research*, *45*(1), 89–125. https://doi.org/10.3102/00346543045001089
Tinto, V., & Cullen, J. (1973). Dropout in Higher Education: A Review and Theoretical Synthesis of Recent Research.
U. S. Bureau of Labor Statistics (2023) Employment Projections. https://www.bls.gov/emp/tables/stem-employment.htm
van den Hurk, A., Meelissen, M., & van Langen, A. (2019). Interventions in education to prevent STEM pipeline leakage. *International Journal of Science Education, 41*(2), 150-164.
Warner, E. T., Carapinha, R., Weber, G. M., Hill, E. V., & Reede, J. Y. (2016). Faculty promotion and attrition: The importance of coauthor network reach at an academic medical center. *Journal of general internal medicine*, *31*, 60–67. https://doi.org/10.1007/s11606-015-3463-7

Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications.* Cambridge University Press. https://doi.org/10.1017/CBO9780511815478

Wright, C. J. (2012). Becoming to remain: Community college students and post-secondary pursuits in central Appalachia. *Journal of research in rural education, 27*(6). https://doi.org/10.18113/p8jrre2706

APPENDIX

Table 1 Network Statistics

Statistic	2018	2019	2020	2021	2022
Survey Responses	25	30	44	62	83
Nodes (Active Members)	48	67	81	105	122
Total Members to Date	48	85	127	182	249
Edges	146	183	215	304	211
Density	0.129	0.083	0.066	0.056	0.029
Average Strength	10.85	10.57	11.62	12.50	7.00
Average Betweenness	0.041	0.037	0.038	0.031	0.032
Giant Component	48	67	75	96	102
Number of Components	1	1	4	4	11
Number of Communities	6	6	6	7	9

Table 2 Yearly change in strength and betweenness of top actors in the network. S stands for strength, SR for strength rank, B for betweenness, and BR for betweenness rank.

ID	2018				2019				2020			
	S	SR	B	BR	S	SR	B	BR	S	SR	B	BR
8	53	1	0.261	1	56.5	1	0.302	1	58	1	0.046	18
27	17.5	11	0.075	11	46	2	0.177	4	54.5	2	0.156	6
26	26.5	5	0.006	20	31	8	0.027	23	52	3	0.019	30
23	36	3	0.103	9	45	4	0.123	8	42	5	0.047	16
82	NA	NA	NA	NA	NA	NA	NA	NA	38	6	0.028	25
ID	2021				2022							
	S	SR	B	BR	S	SR	B	S				
8	65.5	2	0.05	14	53	1	0.17	4				
27	76	1	0.428	1	NA	NA	NA	NA				
26	55	3	0.019	30	30	4	0.208	2				
23	45.5	4	0	60	36	2	0.042	14				
82	40	6	0.001	57	NA	NA	NA	NA				

Table 3 The yearly clique structure for cliques of size 3 or greater. The table shows the year along the top and the size of the cliques along the left side. The numbers within the table show how many of each size cliques there were for that year. For example, there were 195 cliques of size 3 in 2018.

Clique Size	2018	2019	2020	2021	2022
3	195	242	269	403	99
4	152	215	243	383	24
5	162	111	140	241	2
6	10	29	46	92	0
7	0	3	7	20	0
8	0	0	0	2	0
Mean	3.73	3.89	3.98	4.08	3.22

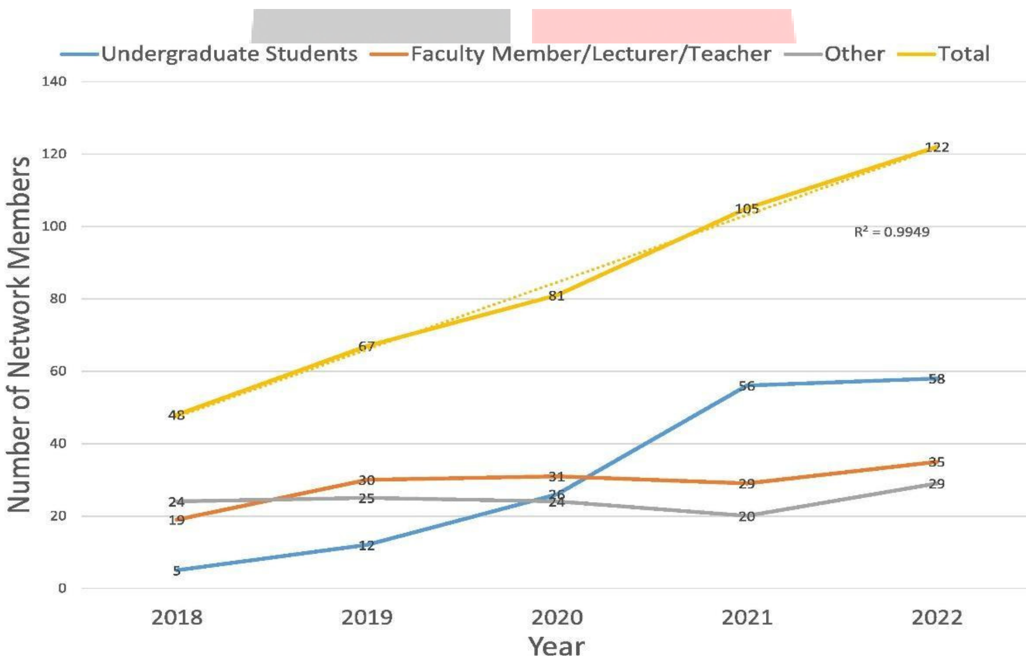

Fig. 1 Network by Category

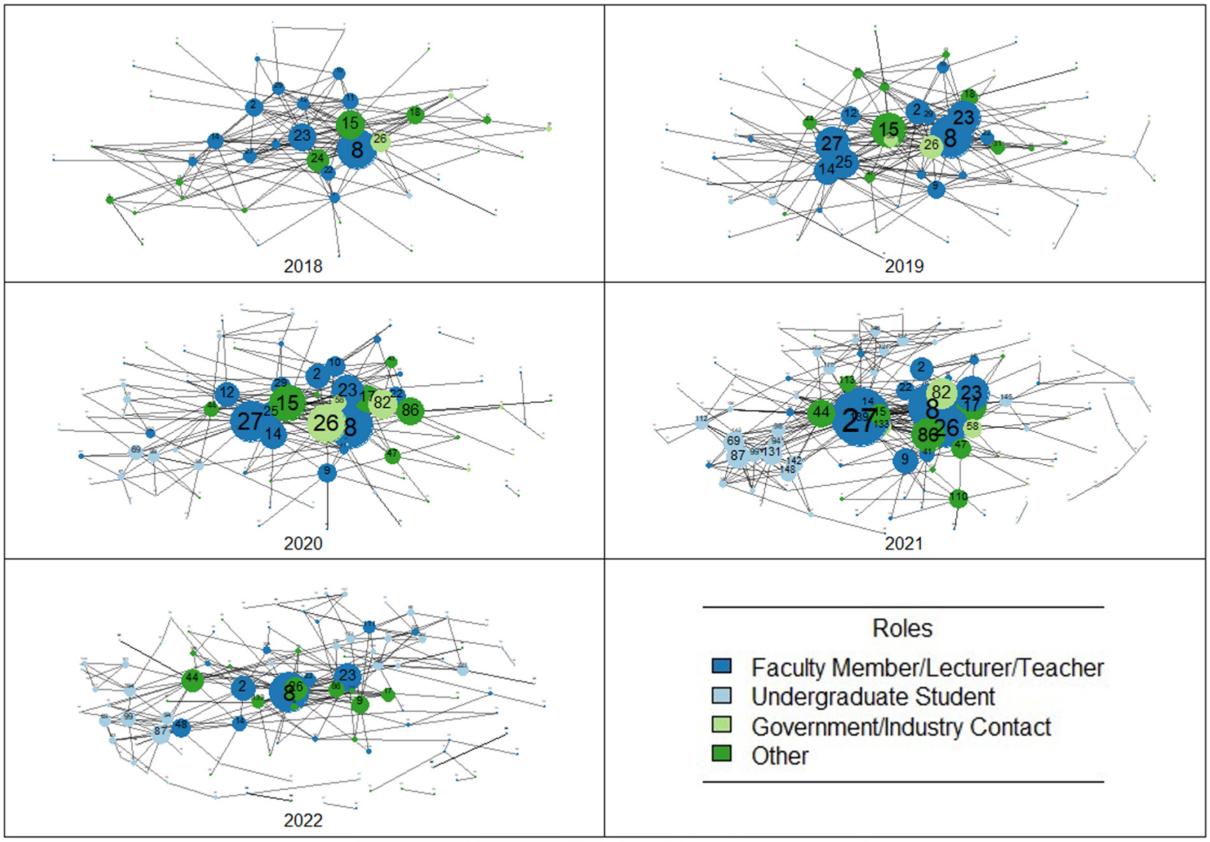

Fig. 2 Network by role, nodes sized by strength

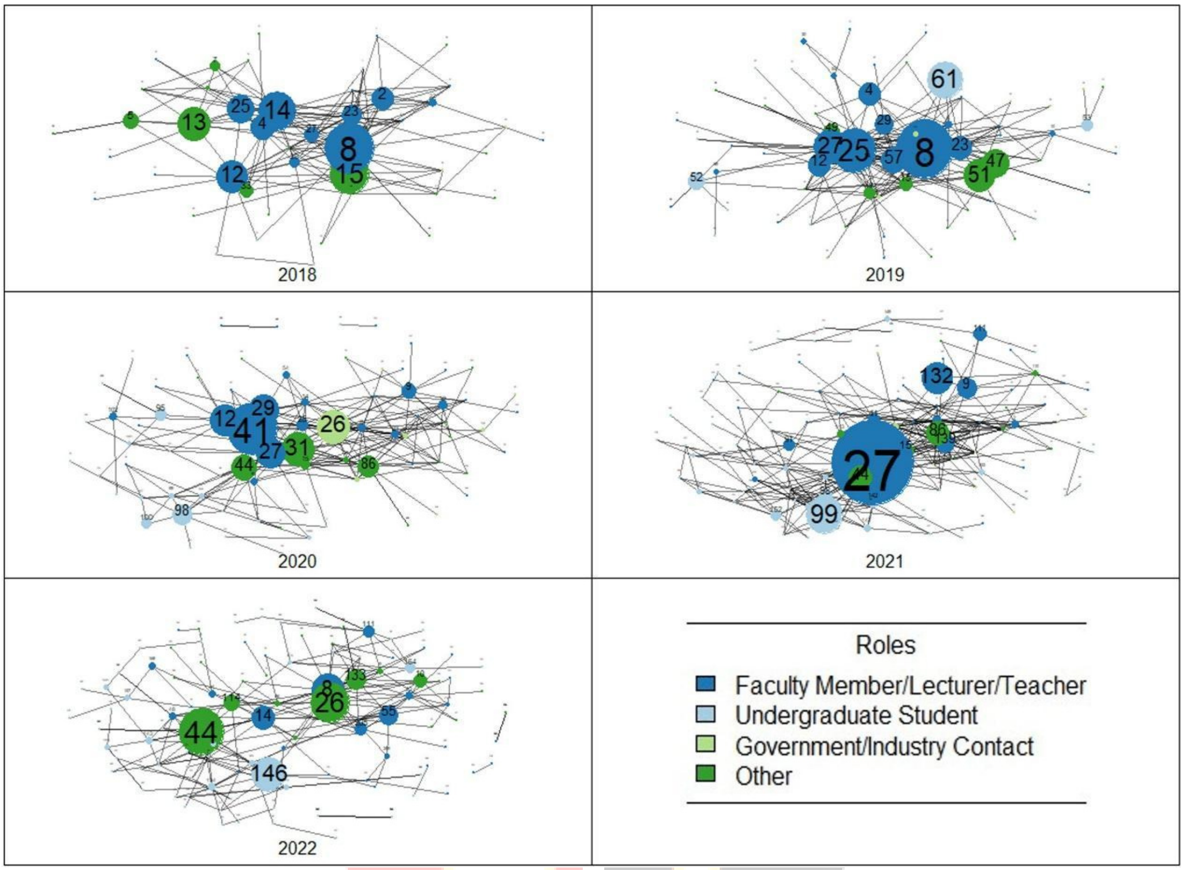

Fig. 3 Network by role, sized by betweenness

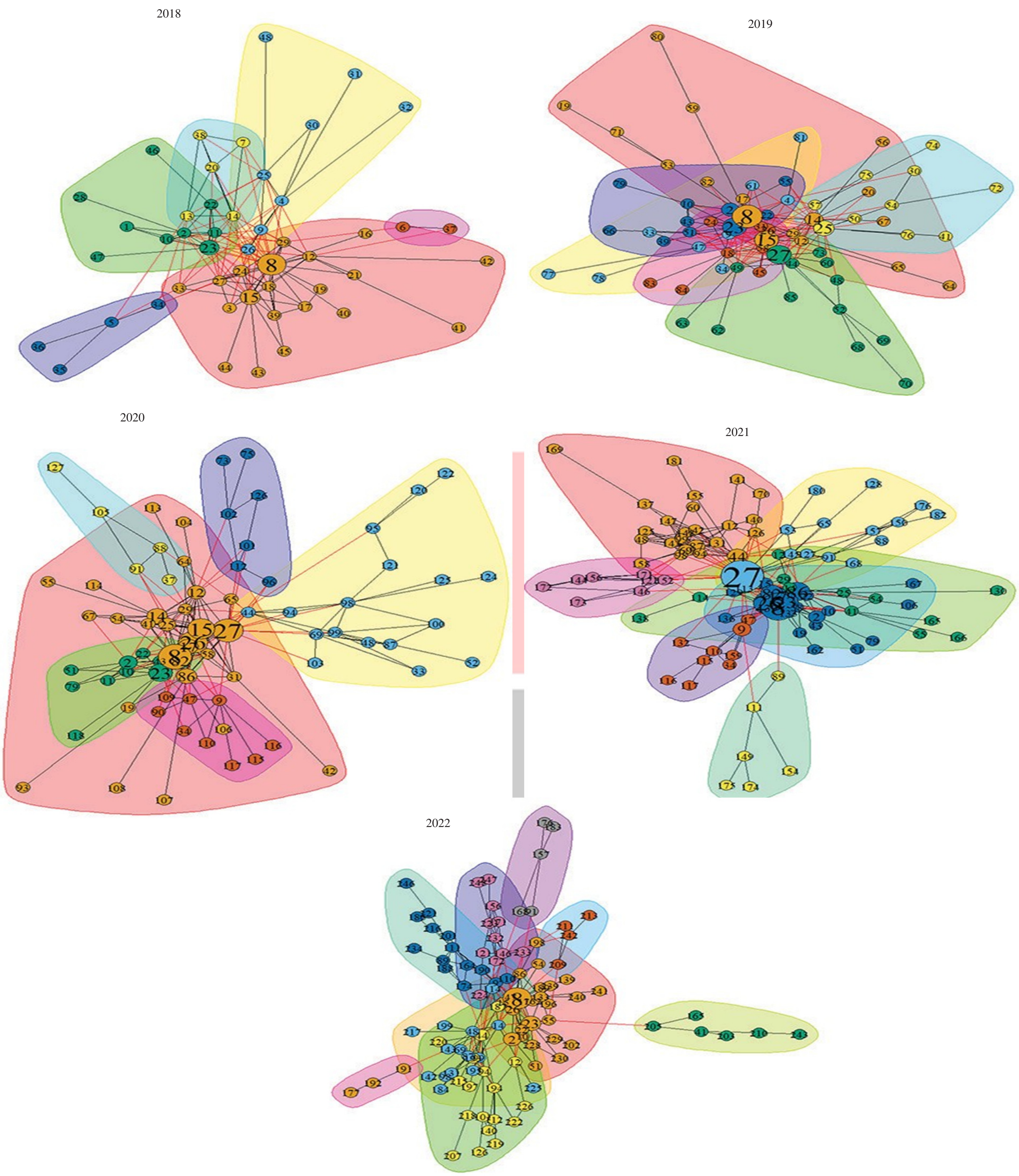

Fig. 4 Communities identified in each year’s giant component

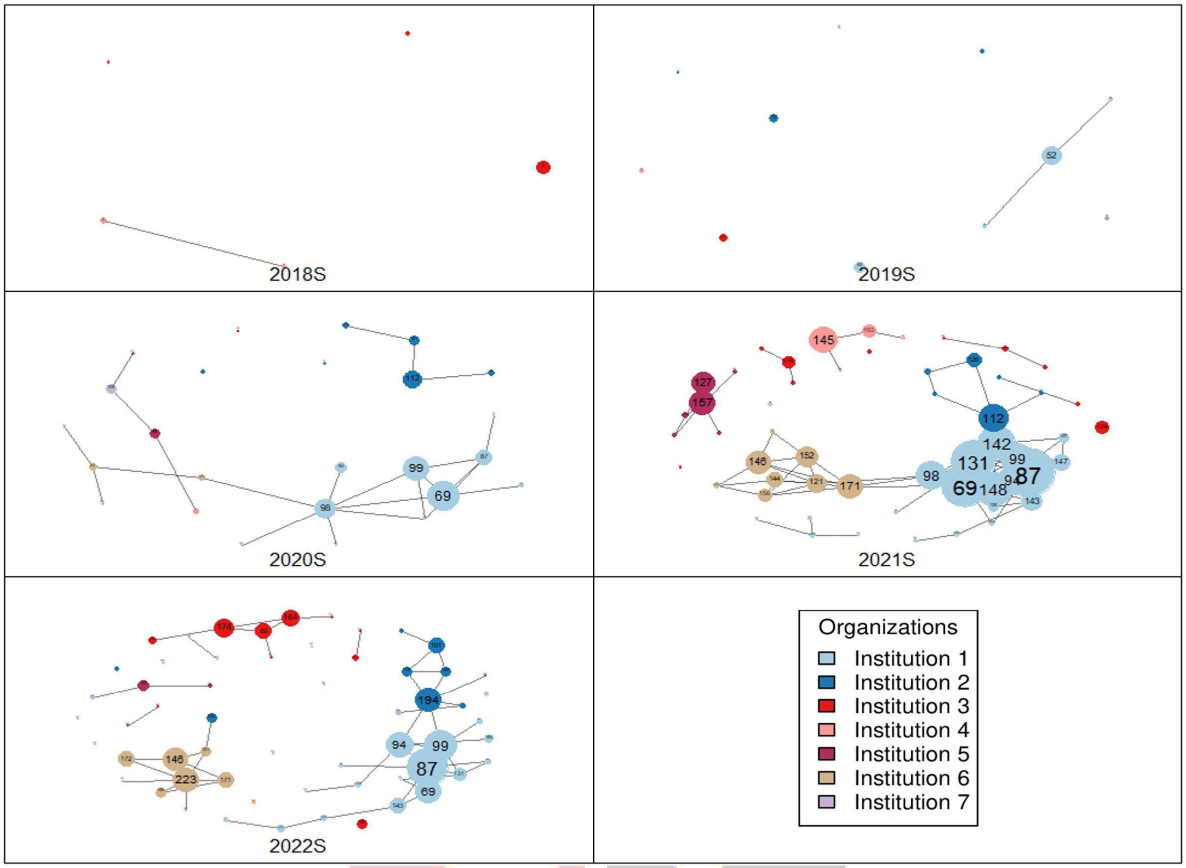

Fig. 5 Graph of student-to-student connections

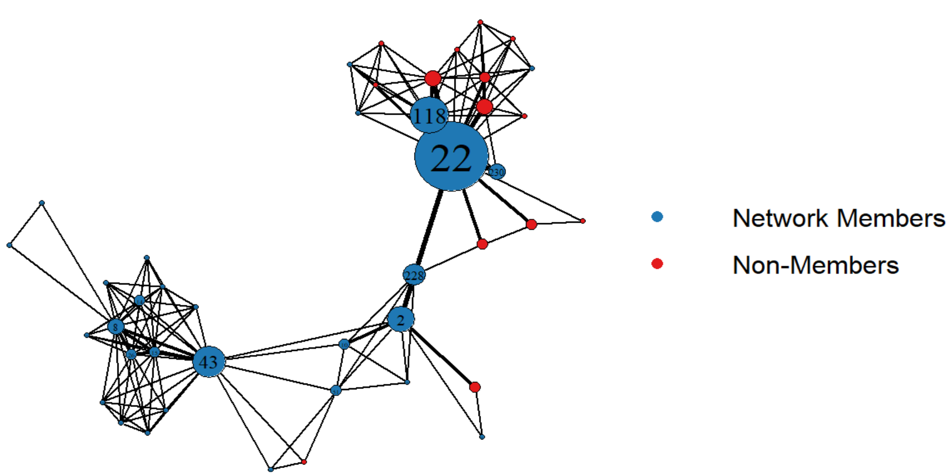

Fig. 6 Network of publications