2020 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)

How do friendship and advice ties emerge? A case study of graduate student social networks

Aditya Tyagi
Northwestern University
Evanston, USA
adityatyagi2020@u.northwestern.edu

Diego Gómez-Zará
Northwestern University
Evanston, USA
dgomezara@u.northwestern.edu

Noshir S. Contractor
Northwestern University
Evanston, USA
nosh@northwestern.edu

***Abstract*—In this paper, we analyze the factors that are most likely to explain the formation of friendship and advice ties among 44 students from a professional STEM graduate program. To answer our research questions, we investigate how students' characteristics influence the formation of their friendship and advice networks using descriptive network analysis, community detection, and Exponential Random Graph Models. The results show that the formation of friendship and advice ties is mostly driven by demographic homophily and prior group activities. Our findings also suggest that female students were more constrained in their friendship and advice networks than male students. We discuss the implications of these results for how graduate students' social networks form at the beginning of their program.**

***Index Terms*—Homophily, personality, community detection, exponential random graph models, network analysis, minorities**

I. Introduction

Starting a graduate program marks a significant transition in many adults' lives. Students must adapt to a new environment, often in a new country, take on more responsibilities, and become more autonomous [1]. Moreover, students have to develop and maintain new friendships, as well as find new sources of advice during their studies. What are some of the key factors that influence whether two students become friends? Or ask each other for advice? Prior studies of social networks have identified two factors that strongly influence the likelihood of forming a relationship between individuals: homophily [2] and activity focus [3, 4]. Homophily describes interacting with others who are similar on various dimensions (e.g., age, education, gender) and activity focus refers to sharing prior group activities (e.g., working in a team together, attending the same party). Because of differences among students, those from small subgroups of the student population can be disadvantaged when it comes to forming friendships, getting advice, and influencing others [5].

Although the literature on friendship and advice networks is extensive, little is known about how STEM students' characteristics and prior group activities affect their friendship and advice networks. Furthermore, analyzing students' social relationships when they are starting their graduate programs can help us understand if and why segregation and inequalities emerge. Understanding the emergence of this problem is crucial because it can suggest strategies to increase cohesiveness, diversity, and exchange of ideas within student networks [6]. Exploring the gaps and groups that emerge among students can inform new strategies to leverage their social capital and reduce dropout rates, especially among underrepresented subgroups.

IEEE/ACM ASONAM 2020, December 7-10, 2020
978-1-7281-1056-1/20/$31.00 © 2020 IEEE

In this case study, we analyze how 44 students of a STEM professional master's program developed friendship and advice networks in their program. Based on their demographics, personality traits, and prior group activities, we explore the factors that are most likely to explain the formation of a friendship and advice relationship. We found that while some of these relationships were driven by homophily, others were driven by heterophily, that is, interacting with others who are different. We also found that prior group activities play a critical role in friendship and advice tie formation. Finally, we found that female students are at a disadvantage vis-a-vis male students: female students are less likely to reach more diverse groups than male students, and female students are less likely to be asked for advice than male students.

II. Related Work

In this section, we examine relevant studies about the formation of friendship networks and advice networks from management, network, and learning sciences literature.

Prior research has explored differences in people's motivations for establishing friendships and for giving and receiving advice. First, advice networks have been described as more instrumental than friendship networks since advice relationships depend on the seeker's task, and the desired outcome is knowledge [7]. Second, advice networks are more fluid than friendship networks over time because they depend on the nature of the information needed. In contrast, friendships are based on trust, emotion, and affection over time. They take more time to form and, once formed, are more difficult to break. Third, advice relationships tend to be asymmetric and non-reciprocating since the less knowledgeable individual is more likely to seek advice from a more knowledgeable person [8]. In contrast, friendship ties are more likely to be symmetrical, where both individuals recognize the relationship. Because of these differences between these two social networks, the identifying variation in factors that predict friendship or advice relationships has gained the attention of several scholars.

How do people's characteristics influence the formation of friendship ties and advice ties? Prior studies offer mixed insights about how gender and race affect the formation of ties in social networks. Morimoto and Yang [9] examined how race, gender, and age determine friendships among graduate students from a sociology terminal master's program. They found strong homophily based on gender and race, but also a directionality in friendship nominations towards white people: minority students were more likely to pick white students as friends; however, white students rarely reciprocated these friendship ties. Moreover, they found that female students were more likely to engage in homophilous friendships than were male students. They contrasted their findings with those of Ibarra and colleagues, which showed that in corporate settings, men tend to have larger and more homogeneous networks than women [10, 11]. Morimoto and Yang reasoned that men must prefer bonding in hierarchical settings like office environments, rather than in egalitarian settings such as graduate school. In their sample, they hypothesized that the male students must be bonding vertically (i.e., with advisors, professors who are predominantly male) than horizontally (i.e., with fellow students who are mostly female). Thus, men were able to obtain additional mentorship to advance their academic careers further.

Other studies highlighted how people's personality traits explain tie formation. Solomon et al. [12] studied egocentric networks of Twitter users, broken down into friends, colleagues, and family communities. Analyzing each user's personality, they found that homophily was not a universal principle; rather, homophily was moderated by personality trait and relation type. For example, they observed that while friend networks exhibit strong homophily, colleague networks exhibit weak homophily. Moreover, while extraverted users tended to make friends, mainly with other extraverted users, neurotic-conscientious friend pairings tended to form more often than neurotic-neurotic and conscientious-conscientious friend pairings. Fang et al. [13] explored how participants' personality traits influence network position in expressive (such as friendship) and instrumental (such as advice) networks. They analyzed how the two most studied structurally advantageous network positions, namely brokerage and in-degree centrality, relate to personality. They found that openness was negatively associated with in-degree centrality in expressive networks, while conscientiousness was positively associated with individuals' brokerage in both expressive and instrumental networks. These results demonstrate that there are systematic differences in the network position occupied by various personality types. Such differences in network position can explain differences in creativity, idea generation, and financial compensation between actors. Together, these differences imply that it is worthwhile to study the granular mechanics of how personality traits influence person-to-person ties, as we do in our study.

Prior studies have also explored how individuals from minority groups face obstacles in establishing friendship and advice ties in organizational contexts. In one study, Forret [14] identified factors that hold minorities back in corporate settings and provided suggestions to mitigate their impact. This study noted that for minorities working in companies, human capital is not enough, social capital is equally as important. Seibert et al. [15] found that the structure of one's network was related to longer-term career success outcomes. Specifically, the size of individuals' networks, the strength of ties, the pattern of ties, and the resources of the ties are critical in evaluating their social capital. In particular, the pattern of ties has been shown to be positively related to upward mobility [16] as well as greater managerial performance [17]. These findings provide a strong impetus to evaluate disadvantages (financial, career, or otherwise) from the perspective of structural holes and network constraint [18]. Many of the problems minorities face in building their social networks in companies can be attributed to the similarity-attraction paradigm, tokenism theory, and existing organizational structures [14]. However, each of these theories implies that women's diminished social capital vis-a-vis men stems from women's minority status and hierarchical organizational structures. By examining contexts where the above do not hold (i.e., non-hierarchical, gender-balanced) networks such as ours, we can expect to uncover additional factors that undermine women's social capital [19].

Although the literature is extensive about the effect of demographic factors such as age, gender, and race on friendship formation, the nuanced effects of personality and prior group activities in determining friendship ties have been left unexplored. Moreover, there is work left to be done to determine what impact friendship ties have on the flow of ideas through the advice network. Block and Grund [20] emphasized that studying friendship networks requires us considering multiple dimensions, such as individuals' characteristics and prior collaborations. Thus, our research questions are:

- **RQ1.** Who are the subgroups within the STEM graduate student networks that are potentially disadvantaged?
- **RQ2.** What is the basis for communities in STEM graduate student friendship and advice networks?
- **RQ3.** How do prior group interaction and personality traits together explain the formation of friendship and advice ties?

III. Methodology

To answer each of our research questions, we conducted a case study with a cohort of graduate students at a large university in the United States. Graduate students were members of a professional STEM Master's program spanning 15 months. In this cohort, there are 48 students evenly split between female and male. All students had undergraduate degrees in STEM fields and about 50% with at least some work experience. As part of the program, students take a total of 17 classes. Out of these 14 classes are taken together as a cohort, while the remaining 3 classes are practicum, capstone, and an elective respectively. All students took the same classes throughout their five academic quarters in the program as a cohort. However, one class was an elective class. As part of the program's goals, students have to assemble *practicum teams* to work as consultants with real clients. Each student was asked to rank his/her client preferences, and then practicum teams are assembled according to students' preferences. They were asked for their preferences within the first 2-3 weeks of the program. The practicum teams' size ranges from 4 to 5 students. Upon graduation, students in this program typically join in industry as analytics professionals at top tech companies. This provides an early-career basis to reason about later stage outcomes such as the lack of female participation in tech leadership roles.

Once participants completed their practicum teams' projects, we sent them a survey using Qualtrics. We provided monetary incentives to the students to get their responses, and we explained to them that their data would be kept private and de-identified for research purposes. The survey was completed by 44 members of the cohort (i.e., 91.67% response rate).

A. Measures

1) Demographics: For descriptive and comparative purposes, we asked a series of questions about the participants' demographics. Fifty percent of respondents were female, with an average age of 23.68 (SD=2.47). They were primarily Asian (77.1%), followed by White (22.9%). To avoid asking participants about their nationalities, we asked them: "What country do you most identify with?" Most participants identified with China (50.0%), followed by USA (29.2%), India (16.7%), and a European country (4.2%). Most participants lived close to campus (87.5%), while the remainder lived in nearby urban areas (12.5%).

2) Personality: Students answered the mini-IPIP scales [21], which assessed the Five-Factor Model attributes of agreeableness, conscientiousness, extraversion, neuroticism, and openness. Participants responded to 20-items (four per trait), and the items were then averaged for each trait.

3) Friendship network: In the survey, students selected the cohort members who they considered as friends and enjoyed spending time with socially. We collected their answers and constructed the students' friendship network in which each node represents a student, and a tie represents a friendship nomination from the person reporting it to another member of the cohort. In other words, a tie going from node *i* to node *j* indicates person *i* selecting person *j* as a friend.

4) Advice network: In the survey, students also selected the cohort members who they sought out for schoolwork-related help or professional advice. The survey reiterated that their responses did not necessarily need to overlap with those reported as their friends in response to the preceding question. We collected their answers and constructed the students' advice network in which each node represents a student, and a tie represents an advice nomination from the person reporting it to another member of the cohort. A tie going from node *i* to node *j* indicates that person *i* asks person *j* for advice.

5) Practicum teams: Finally, students selected the cohort members with whom they are working on their practicum projects. We used their responses to construct the group activity network, where each node represents a student, while an undirected tie between two nodes indicates membership on the same practicum team. We considered two students who were part of the same team as having a collaboration.

B. Descriptive network analysis

To answer our RQ1, we conducted network analysis to investigate how core network statistics varied with student attributes, and thus identify potentially disadvantaged sub-groups. We began with an exploratory network analysis of the friendship and advice networks computing global network summaries like density (i.e., the ratio between the number of edges observed and the number of ties possible in the network), centralization by degree (i.e., graph-level centrality score based on node-level centrality measure) and reciprocity (i.e., dyads that are mutually linked). We then analyzed each network visually, along with patterns in attribute values. Finally, as part of our exploratory analysis we evaluated how gender influenced network constraint, authority, and degree centrality in advice and friendship networks. We computed the following measures for this analysis:

- *In-degree centrality*: it measures a node's importance as the number of edges incoming to a node
- *Network constraint*: it measures the extent to which a node connects two others with complementary sources of information [22]. The higher the network constraint, the lesser sources of non-complementary information the node has access to. For unweighted networks, Borgatti [23] proposed a simple formula to calculate the size of an ego node's effective network, and hence constraint as $n - \frac{2t}{n}$, where n is the number of nodes in the egocentric network, t is the total ties in the egocentric network.
- *Authority*: The notion of authority scores comes from the HITS (Hyperlink Induced Topic Search) algorithm, a link analysis algorithm used to rate web pages [24]. When applied to social networks, HITS assigns two scores for each node: a hub score and an authority score. A hub score measures a node's tendency to point to other authoritative sources of information. Meanwhile, a high authority score signifies a node pointed to by many hubs [25]. Given the recursive nature of this definition, an iterative algorithm (HITS) is used to compute hub and authority scores, respectively. We applied HITS analysis to the advice network because we wanted to identify authoritative sources of knowledge (i.e., authorities) in the network.

C. Community detection analysis

To answer our RQ2, we conducted a community detection analysis to explore the group structure of the advice and friendship networks, and their potential demographic basis. Our rationale behind deploying community detection algorithms was to uncover what communities existed based solely on the vertex-edge structures in the network, and then compare to the various individual attributes to see if they matched up with our intuition. We used modularity maximization to determine the communities in our graph. Modularity maximization views the community detection problem as an integer linear program, with modularity as the objective function to be maximized [26]. We decided to use maximum modularity clustering because our graph was relatively small (44 nodes approx.), and this made the problem computationally tractable.

D. Exponential Random Graph Models

Lastly, to answer our RQ3, we utilized an Exponential Random Graph Model (ERGM) to evaluate the effect of personality traits in forming friendship ties and advice ties. ERGMs are particularly well suited to identify the individual, relational, and network-level variables included in our multi-theoretical, multilevel network hypotheses, and to explain the motivations behind who nominates whom for friendship and advice ties [27]. In addition to examining the effect of individual level attributes, this statistical model estimates the likelihood of the observed friendship and advice network structures emerging from all possible network configurations generated based on certain hypothesized self-organizing principles among the ties as well as from other network ties, such as shared work-group or collaboration. This distribution in ERGM as:

$$P(Y = y|\theta) = \frac{\exp(\theta^T s(y))}{c(\theta)}$$

Where Y is a network, s(y) is a vector of network statistics, θ is a vector of coefficients, and $c(\theta)$ is a normalizing constant. The log-odds probability of a tie occurring in the network is represented as: $logit(Y_{ij}|Y^C) = (\theta)'\delta s(y)_{ij}$ where Y_{ij} represents a dyad in the network Y, Y^C is the rest of the network, and $\delta s(y)_{ij}$ is the change in the vector of network statistics caused by setting Y_{ij} to 1.

Similar to logistic regressions, ERGM uses the Maximum Likelihood Criterion to estimate the network statistics' coefficients. Positive and significant coefficients indicate that the corresponding independent variable is more likely to influence a tie occurring than by chance. Negative and significant coefficients indicate that the independent variable is less likely to result in a tie occurring than by chance alone. We use Markov-Chain Monte Carlo (MCMC) to identify maximum likelihood estimates (MLE) for parameter values. MCMC simulates thousands of random networks fitting the model's quantifiable properties, rather than attempting to count the impossibly large number of possible network's edge permutations. We selected network statistics based on driving factors previously studied in social network formation [28]. Once the ERGM statistics and its coefficients are estimated, we test whether the observed network is likely to be observed within the distribution of simulated networks. We conducted this analysis using R 3.6.0 and the "ergm" package from *statnet* [29].

For the friendship and advice network, we specified an ERGM for each network that include the following terms:

1) Structural effects: We controlled whether endogenous effects influenced the formation of ties in their friendship and advice networks: the number of ties, to what extent participants tended to reciprocate these ties, and to what extent participants tend to close triads. To measure these properties, we included the terms *edges*, *mutual*, and *dgwesp* from the *ergm* package, respectively. A positive and significant effect means that the respective structural signature is more likely to occur in the friendship (advice) network than by chance.

2) Individual attributes: We controlled whether individuals' attributes were likely to explain the formation of ties in their friendship and advice networks. We separated the effects by the sender (i.e., out-link) and receiver (i.e., in-link).

3) Edge-covariates: We represented the practicum team's network as an independent variable for predicting ties in friendship and advice networks. Since friendship and advice networks were highly bound, we included the friendship (advice) network to see whether it influenced the formation of a tie in the advice (friendship) network. For each one of these prior networks, we added these edge-covariate terms using *edgecov* from the *ergm* package. A positive and significant effect means that prior ties are more likely to form an advice tie (friendship) than by chance. In contrast, a negative and significant effect means that prior ties are less likely to form an advice tie (friendship) than by chance.

4) Categorical homophily effects: We controlled homophily among categorical variables (i.e., gender, race, and identified-country) using the *nodematch* term from the *ergm* package. This term counts how many nodes connected in the network share the same value for that categorical attribute. As a result, a positive and significant effect means that participants were more likely to form a tie with another individual with the same characteristics. In contrast, a negative and significant effect means that participants are more likely to form a tie with individuals who have different values in that attribute. Since these comparisons are among categorical values, we set "Female" as the base for gender, "Asian" as the base for race, the "European country" as the base for identified-country.

5) Numerical homophily effects: We controlled for homophily among numerical variables (i.e., age, openness, conscientiousness, extraversion, agreeableness, and neuroticism) using *absdiff* from the ergm package. In contrast to *nodematch*, this term measures the absolute difference of an attribute between two participants. A smaller difference means that participants with similar values among that attribute are more likely to form a tie, whereas a big difference means that participants with different values among that attribute are more likely to form a tie. As a result, a negative and significant effect means that participants are more likely to form a tie when they have similar scores in that attribute, whereas a positive and significant effect means that participants are more likely to form a tie when they have different scores in that attribute.

IV. Findings

We start by presenting descriptive data to characterize our 44 participants' friendship and advice networks. The density of the friendship network was 25.9%, and the density of the advice network was 38.3%. We calculated both networks' centralization scores using degree centrality and found that the advice network (43.4%) was substantially more centralized than the friendship network (21.3%). We found high

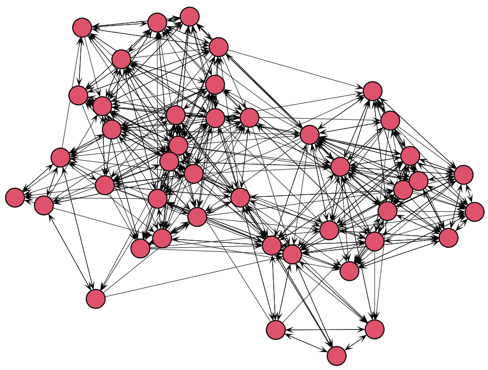

Fig. 1. Friendship network

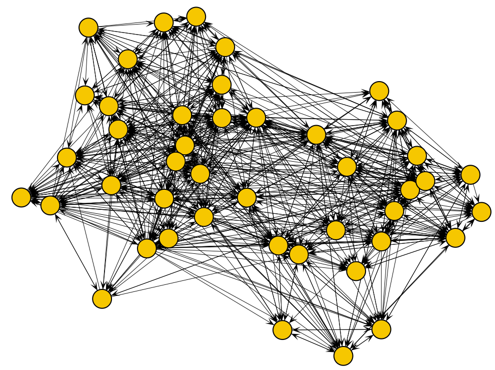

Fig. 2. Advice network

reciprocity levels in both the friendship network (68.6%) and advice network (61.5%). Figure 1 represents participants' nominations for the Friendship network, and Figure 2 represents participants' nominations for the Advice network.

A. RQ1: Disadvantaged Subgroups

Our network analysis indicates that gender is a key distinguishing factor in friendship and advice networks. We found that female students are disadvantaged vis-a-vis male students in both the friendship and advice networks. This was concluded on the basis of a one-sided, two-sample, (unpaired) *t*-test comparing male and female students. Female students tend to occupy high network constraint positions and male students tend to occupy low network constraint positions in the friendship network ($m_{female} = 0.1851 > m_{male} = 0.1677, t = 1.2618, p > 0.05, df = 45.415$) and advice network ($m_{female} = 0.1294 > m_{male} = 0.1158, t = -2.52, p < 0.001, df = 44.625$). This indicates that females report friendship ties with others who report friendship ties with one another. In the advice network, female students tend to have lower indegree centrality than male students ($m_{female} = 16.08 < m_{male} = 19.96, t = 3.64, p < 0.001, df = 45.821$). Thus, female students are less likely to be sought out for advice than male students. As for Authority scores, we find that female students are less likely to have high authority scores than male students in the advice network ($m_{female} = 0.67, m_{male} = 0.77, t = 3.15, p < 0.001, df = 45.957$). We did not find any meaningful differences based on race and identified-country. In other words, there was no discernable disadvantage in terms of race and nationality.

B. RQ2: Groups Driven by Country Homophily

In the friendship network, we were able to detect three communities driven by the top-3 identified-countries among participants: China, USA, and India (Modularity = 0.39). Figure 3 shows in red the cluster formed by students identified with China, followed by the blue cluster formed by students identified with the USA, and the green cluster formed by students identified with India.

In the advice network, we detected two communities driven primarily by the top-2 identified-countries among participants (Modularity = 0.22). The blue cluster is formed by students who are not identified with China and the red cluster is formed by students identified with China. Figure 4 shows the clusters formed in the advice network. In other words, friendship and advice communities were mostly explained by students' identified-countries. We did not find communities driven by race or gender.

C. RQ3: Personality and prior collaboration predict friendship and advice ties

We found that the magnitude and direction of effects varied for both the friendship and advice networks. Table I shows the ERGM results of the friendship network and advice network. In the case of the friendship network, reciprocity ($\beta = 1.911, p < 0.001$) and triadic closure ($\beta = 0.524, p < 0.001$) were more likely to occur than by chance. We found that working on the same practicum team ($\beta = 0.794, p < 0.001$) and asking each other for advice ($\beta = 1.891, p < 0.001$) were strong predictors of a friendship between two students. Female students ($\beta = -0.469, p < 0.01$), students with higher levels of extraversion ($\beta = 0.009, p < 0.05$) and neuroticism ($\beta = 0.008, p < 0.05$) were more likely to be mentioned as friends. Male students ($\beta = 0.617, p < 0.001$), older students ($\beta = 0.167, p < 0.001$), and students identified with India ($\beta = 1.214, p < 0.05$) were more likely to mention others as friends. Similarly, students with high levels of openness ($\beta = 0.013, p < 0.001$), conscientiousness ($\beta = 0.011, p < 0.05$), and neuroticism ($\beta = 0.008, p < 0.05$) were likely to mention others as friends. Regarding the homophily effects, the ERGM results show that students who identified with the same country ($\beta = 1.305, p < 0.001$), had similar agreeableness levels ($\beta = -0.008, p < 0.05$), dissimilar openness levels ($\beta = 0.008, p < 0.05$) and dissimilar extraversion levels ($\beta = 0.005, p < 0.10$) were more likely to form ties. Conscientiousness and neuroticism were found to be insignificant in determining friendship ties.

The ERGM results of the advice network indicate that reciprocity ($\beta = 0.907, p < 0.001$) and triadic closure ($\beta = 1.241, p < 0.001$) were more likely to occur than by chance. However, while reciprocity levels of the advice network were lower than in the friendship network, the triadic

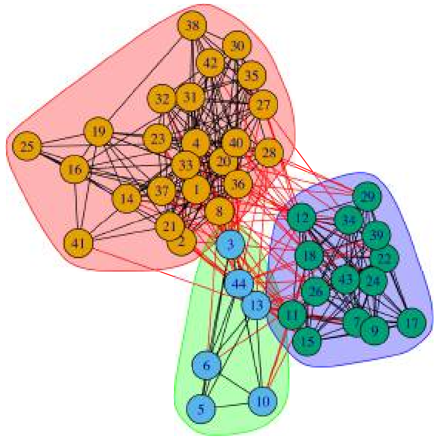

Fig. 3. Community detection for the friendship network

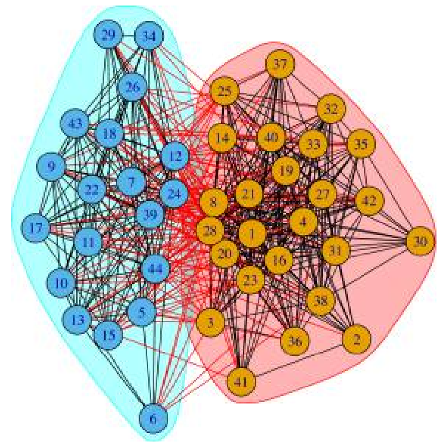

Fig. 4. Community detection for the advice network

closure levels were higher in the advice network. Asking or receiving advice was more likely to occur between friends ($\beta = 1.941, p < 0.001$), but not necessarily between practicum teammates ($\beta = 0.356, p < 0.10$). Unlike the friendship network, male students were more likely to be asked for advice, although this coefficient was not statistically significant ($\beta = 0.228, p < 0.10$). Compared to the students who identified with European countries, students identified with China were less likely to be asked for advice ($\beta = -1.416, p < 0.001$). We found that students who identified with China ($\beta = 3.165, p < 0.001$) and white students ($\beta = 0.708, p < 0.01$) were more likely to ask for advice. Also, students who were more likely to ask for advice had lower levels of openness ($\beta = -0.011, p < 0.001$) and neuroticism ($\beta = -0.021, p < 0.001$), and higher levels of extraversion ($\beta = 0.007, p < 0.05$) and agreeableness ($\beta = 0.026, p < 0.001$). Regarding homophily effects, we found that students who identified with the same country were more likely to establish a tie ($\beta = 0.791, p < 0.001$). Finally, the results show that students with different agreeableness levels ($\beta = 0.01, p < 0.01$) and similar extraversion levels ($\beta = -0.005, p < 0.10$) were more likely to establish an advice tie.

V. Discussion

The main goal of this case study was to explore how early-career STEM professional master's students develop friendship and advice networks during their programs. Our findings show that students' friendship and advice relationships were driven mostly by similarity among demographic and personality attributes, as well as having prior group activities together.

RQ1 asked about the subgroups within the STEM graduate student networks that are disadvantaged. Our results suggest that female students were likely to be more constrained than male students in friendship networks. In other words, female students were less likely to be exposed to novel information beyond what is already circulating in their closed cohesive group of friends. In contrast, male students were more likely to act as brokers and have friends in different social circles. In the advice networks, female students were also more constrained than male students. In our sample, male students were able to receive and provide advice from different students, whereas female students were more likely to receive and provide advice within closed social circles. Female students' advice was also overall less solicited than male students. These findings reveal that even when a cohort is gender-balanced, social networks of female and male students could evolve unequally. Prior literature confirms this disadvantage between female students and male students. In the field of sociology, female students tend to build largely homophilous networks among students and are not able to bond with largely male faculty [9]. The formation of mentoring relationships is one of three critical factors in advancing women and minorities in organizations; the other two being networking and network groups [14]. Our findings confirm that women are indeed disadvantaged; however, we provide a novel explanation. Instead of attributing women's disadvantage to a largely male faculty body like [9] or to women's minority status in hierarchical organizations [14], we attribute their disadvantage to the high network constraint positions that they tend to occupy in the graduate student network. As a result, they can be less likely to get access to novel information needed to produce innovative ideas. It is reasonable to expect that this phenomenon will replicate over time once students begin their career as tech professionals. Thus, our work provides one possible, early career explanation for a lack of female representation in tech leadership roles. Our findings also have implications for business leaders, sociologists, and policy-makers as they seek to develop diversity initiatives. As our analysis shows, even with perfect gender parity, network effects can create 'invisible inequity' as early as graduate school.

RQ2 asked about the basis for communities to emerge in STEM graduate student friendship and advice networks. Our results show that participants were likely to form friendship and advice groups based on their nationalities. Rather than

	Friendship ERGM	Advice ERGM
Structural terms		
Edges	-10.386 (1.589)***	-7.862 (1.48)***
Reciprocity	1.911 (0.219)***	0.907 (0.171)***
Triadic closure	0.524 (0.098)***	1.241 (0.2)***
Prior relationships		
Practicum teammate	0.794 (0.19)***	0.365 (0.191)†
Giving/receiving advice	1.891 (0.149)***	
Friends		1.941 (0.143)***
Individuals' attributes (in-links)		
Gender (Male)	-0.469 (0.173)**	0.228 (0.134)†
Race (White)	-0.386 (0.305)	-0.171 (0.264)
Identified-country (China)	-2.682 (0.47)***	-1.416 (0.419)***
Identified-country (India)	-2.077 (0.476)***	-0.97 (0.422)*
Identified-country (USA)	-1.774 (0.394)***	-1.045 (0.34)**
Age	0.074 (0.04)†	0.018 (0.035)
Openness	-0.004 (0.003)	0.002 (0.003)
Conscientiousness	-0.007 (0.004)†	0.005 (0.003)
Extraversion	0.009 (0.004)*	0.002 (0.003)
Agreeableness	-0.006 (0.005)	-0.005 (0.004)
Neuroticism	0.008 (0.004)*	0.005 (0.003)
Individuals' attributes (out-links)		
Gender (Male)	0.617 (0.168)***	0.109 (0.136)
Race (White)	0.586 (0.313)†	0.708 (0.259)**
Identified-country (China)	0.633 (0.563)	3.165 (0.464)***
Identified-country (India)	1.214 (0.56)*	2.728 (0.47)***
Identified-country (USA)	0.633 (0.483)	2.511 (0.397)***
Age	0.167 (0.039)***	-0.026 (0.034)
Openness	0.013 (0.003)***	-0.011 (0.003)***
Conscientiousness	0.011 (0.004)*	-0.004 (0.003)
Extraversion	0 (0.003)	0.007 (0.003)*
Agreeableness	-0.017 (0.005)***	0.026 (0.004)***
Neuroticism	0.008 (0.004)*	-0.021 (0.003)***
Categorical homophily effects *(nodematch)*		
Gender	0.146 (0.118)	0.002 (0.105)
Race	0.261 (0.166)	0.061 (0.165)
Identified-country.	1.305 (0.167)***	0.791 (0.15)***
Numerical homophily effects *(absdiff)*		
Age	-0.092 (0.031)**	-0.006 (0.03)
Openness	0.008 (0.003)*	-0.001 (0.003)
Conscientiousness	-0.003 (0.004)	0.001 (0.003)
Extraversion	0.005 (0.003)†	-0.005 (0.003)†
Agreeableness	-0.008 (0.004)*	0.01 (0.003)**
Neuroticism	0 (0.003)	0.002 (0.003)
Null Deviance	3,127.48 on 2,256 d.f.	3,127.48 on 2,256 d.f.
Residual Deviance	1,379.35 on 2,220 d.f.	2,013.24 on 2,220 d.f.
AIC	1,451.34	2,085.23
BIC	1,657.31	2,291.20

Significance codes: *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$, † $p < 0.10$

TABLE I
ERGM RESULTS FOR FRIENDSHIP AND ADVICE NETWORKS. STANDARD DEVIATION IN PARENTHESES.

gender or race, it was their identified-country that influenced how students established friendship and advice circles. For example, Asian students who identified with the US tended to form ties with students who also identified with the US, rather than with students who identified with China. We checked whether other attributes were likely to explain students' social groups. Neither race nor gender was likely to explain these subdivisions among STEM graduate students. The community detection analysis indicates the hardwired tendencies to seek advice and establish friendships with people from the same identified-country. Similar results were reported by [30], who found high levels of homophily in nationality on research co-authorship networks. It would be reasonable to extrapolate that such nationality-based communities will continue to appear even when these students begin their careers as tech professionals. Thus, one possible solution for managers looking to foster more innovative, high-value teams would be to encourage team formation with individuals from diverse national backgrounds. One study found that CEO's with highly heterogeneous in terms of nationality social networks had higher firm valuations, due to enhanced social learning [31]. One possible explanation for forming friendships and advice relationships with people of the same country is the low cultural dissonance. People from the same country are more likely to share the same language, social norms, and references than people from different countries [32]. Lastly, RQ3 asked whether prior group interaction and personality traits together explain the formation of friendship and advice ties. In the case of the friendship network, our results show that homophily and prior collaborations were most likely to explain a friendship tie between two students. Students who nominated another person as a friend have similar levels of extraversion and had a prior group activity together. However, people with different levels of conscientiousness and agreeableness were more likely to establish friendships. In the advice network, we found that giving and receiving advice were more likely to happen between students who identified with the same country, who lived in the same apartment complex, were from the seam practicum team, and who already identified themselves as friends. We did not find any significant effects between students' personality traits and their advice relationships. One possible explanation is the instrumental nature of advice networks. Rather than seeking advice from people who are compatible in terms of personality, students sought advice from other students who are familiar to them in terms of prior experiences and identified countries, which ultimately promote trust among individuals. Living nearby in the same apartment building also played a role in facilitating advice ties. Additionally, ERGM results show a high multiplexity between friendship and advice networks, in the sense that one tie is likely to predict the emergence of the other. Since we cannot infer any causality with these results, future longitudinal work should explore the causal effect between friendship and advice ties.

There are some limitations to our study that can serve as a foundation for future work. First, we relied on self-reported data to analyze students' friendship and advice networks. Students could have forgotten to nominate other students as friends, as well as those from whom they sought advice. Second, students' understanding of what constitutes friendships may not be consistent. To address the above two factors, future extensions could study whether students' online interaction patterns on social networks such as Twitter, LinkedIn, Instagram mirror our findings in real-life. In particular, an analysis of student collaboration patterns on enterprise social media such as Slack would add a new light to our findings. Third, we did not assess whether students' prior collaborative experience working in groups with other students was positive or not. Future studies may build on our work by evaluating the individual experiences (positive/negative) in working with their practicum team, and how this affects friendship/advice relationships. Fourth, neither African-American nor Hispanic students were part of this sample, which are the largest minority groups in the US. Future case studies should consider a good representation of these minorities in their analyzes. Finally, our survey was conducted at a singular moment. Future studies could benefit from longitudinal analysis to study how friendship and advice networks evolve over time.

VI. Conclusion

In this paper, we studied the most likely factors that explain the formation of friendship and advice relationships in a professional STEM master's program. Our case study has shown that students' race, gender, identified-country, and personality are key factors determining whether friendship and advice ties form between students. However, the effect is not the same for all personality traits. In addition, having prior collaborations has a positive influence on building ties in both networks. Finally, we show that there is a strong gender effect in favor of males in terms of network position. Sociologists, policy researchers, and management scientists can use our findings at the initial graduate school level to motivate a interdisciplinary, longitudinal analysis of female attrition in tech roles over time - commonly referred to as the leaky bucket problem [33]. Our results provide empirical evidence regarding the social bias in favor of interacting with similar people, and the social challenges that female students must face when they start a graduate program.

References

[1] B. McEwan and L. K. Guerrero, "Maintenance behavior and relationship quality as predictors of perceived availability of resources in newly formed college friendship networks," *Communication Studies*, vol. 63, no. 4, pp. 421–440, Sep. 2012.

[2] M. McPherson, L. Smith-Lovin, and J. M. Cook, "Birds of a feather: Homophily in social networks," *Annual Review of Sociology*, vol. 27, no. 1, pp. 415–444, 2001. [Online]. Available: https://doi.org/10.1146/annurev.soc.27.1.415

[3] S. R. Corman and C. R. Scott, "Perceived Networks, Activity Foci, and Observable Communication in Social Collectives," *Communication Theory*, vol. 4, no. 3, pp. 171–190, 03 2006. [Online]. Available: https://doi.org/10.1111/j.1468-2885.1994.tb00089.x

[4] S. L. Feld, "The focused organization of social ties," *American Journal of Sociology*, vol. 86, no. 5, pp. 1015–1035, 1981. [Online]. Available: https://doi.org/10.1086/227352

[5] L. Leszczensky and S. Pink, "Ethnic segregation of friendship networks in school: Testing a rational-choice argument of differences in ethnic homophily between classroom- and grade-level networks," *Soc. Networks*, vol. 42, pp. 18–26, Jul. 2015.

[6] I. J. Raabe, Z. Boda, and C. Stadtfeld, "The social pipeline: How friend influence and peer exposure widen the STEM gender gap," *Sociol. Educ.*, vol. 92, no. 2, pp. 105–123, Apr. 2019.

[7] J. Nebus, "Building collegial information networks: A theory of advice network generation," *AMRO*, vol. 31, no. 3, pp. 615–637, Jul. 2006.

[8] K. M. Carley and D. Krackhardt, "Cognitive inconsistencies and non-symmetric friendship," *Soc. Networks*, vol. 18, no. 1, pp. 1–27, Jan. 1996.

[9] S. A. Morimoto and S. Yang, "What friendship entails: An empirical analysis of graduate students' social networks," *Sociol. Spectr.*, vol. 33, no. 2, pp. 99–116, Mar. 2013.

[10] H. Ibarra, "Homophily and differential returns: Sex differences in network structure and access in an advertising firm," *Adm. Sci. Q.*, vol. 37, no. 3, pp. 422–447, 1992.

[11] ——, "Paving an alternative route: Gender differences in managerial networks," *Soc. Psychol. Q.*, vol. 60, no. 1, pp. 91–102, 1997.

[12] R. S. Solomon, S. Pykl, A. Das, B. Gamback, and T. Chakraborty, "Understanding the Psycho-Sociological facets of homophily in social network communities," *IEEE Comput. Intell. Mag.*, vol. 14, no. 2, pp. 28–40, May 2019.

[13] R. Fang, B. Landis, Z. Zhang, M. H. Anderson, J. D. Shaw, and M. Kilduff, "Integrating personality and social networks: A Meta-Analysis of personality, network position, and work outcomes in organizations," *Organization Science*, vol. 26, no. 4, pp. 1243–1260, Aug. 2015.

[14] M. L. Forret, "The impact of social networks on the advancement of women and racial/ethnic minority groups," *Gender, ethnicity, and race in the workplace*, vol. 3, pp. 149–166, 2006.

[15] S. E. Seibert, M. L. Kraimer, and R. C. Liden, "A social capital theory of career success," *AMJ*, vol. 44, no. 2, pp. 219–237, Apr. 2001.

[16] J. M. Podolny and J. N. Baron, "Resources and relationships: Social networks and mobility in the workplace," *Am. Sociol. Rev.*, vol. 62, no. 5, pp. 673–693, 1997.

[17] S. Rodan and C. Galunic, "More than network structure: how knowledge heterogeneity influences managerial performance and innovativeness," *Strat. Mgmt. J.*, vol. 25, no. 6, pp. 541–562, Jun. 2004.

[18] R. S. Burt, "Structural holes and good ideas," *Am. J. Sociol.*, vol. 110, no. 2, pp. 349–399, 2004.

[19] D. J. Brass, "Men's and women's networks: A study of interaction patterns and influence in an organization," *Academy of Management Journal*, vol. 28, no. 2, pp. 327–343, 1985. [Online]. Available: https://doi.org/10.5465/256204

[20] P. Block and T. Grund, "Multidimensional homophily in friendship networks*," *Network Science*, vol. 2, no. 2, pp. 189–212, Aug. 2014.

[21] M. B. Donnellan, F. L. Oswald, B. M. Baird, and R. E. Lucas, "The mini-IPIP scales: tiny-yet-effective measures of the big five factors of personality," *Psychol. Assess.*, vol. 18, no. 2, p. 192, 2006.

[22] R. S. Burt, "Structural holes versus network closure as social capital," in *Social capital*. Routledge, 2017, pp. 31–56.

[23] S. P. Borgatti, "Structural holes: Unpacking burt's redundancy measures," *Connect.*, vol. 20, no. 1, pp. 35–38, 1997.

[24] J. M. Kleinberg, "Authoritative sources in a hyperlinked environment," *Journal of the ACM (JACM)*, vol. 46, no. 5, pp. 604–632, 1999.

[25] C. D. Manning, P. Raghavan, and H. Schütze, *Introduction to Information Retrieval*. USA: Cambridge University Press, 2008.

[26] U. Brandes, D. Delling, M. Gaertler, R. Gorke, M. Hoefer, Z. Nikoloski, and D. Wagner, "On modularity clustering," *IEEE Trans. Knowl. Data Eng.*, vol. 20, no. 2, pp. 172–188, Feb. 2008.

[27] G. Robins, T. Snijders, P. Wang, M. Handcock, and P. Pattison, "Recent developments in exponential random graph (p*) models for social networks," *Soc. Networks*, vol. 29, pp. 192–215, 2007.

[28] M. Zhu, Y. Huang, and N. S. Contractor, "Motivations for self-assembling into project teams," *Social Networks*, vol. 35, no. 2, pp. 251 – 264, 2013, special Issue on Advances in Two-mode Social Networks. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0378873313000166

[29] D. R. Hunter, M. S. Handcock, C. T. Butts, S. M. Goodreau, and M. Morris, "ergm: A package to fit, simulate and diagnose exponential-family models for networks," *J. Stat. Softw.*, vol. 24, no. 3, p. nihpa54860, 2008.

[30] M. Vaanunu and C. Avin, "Homophily and nationality assortativity among the most cited researchers' social network," in *2018 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM)*, Aug. 2018, pp. 584–586.

[31] Y. Fang, B. Francis, and I. Hasan, "Differences make a difference: Diversity in social learning and value creation," *Journal of Corporate Finance*, vol. 48, pp. 474–491, 2018.

[32] M. Rhee, D. Yang, and T. Yoo, "National culture and friendship homophily in the multinational workplace," *Asian Business & Management*, vol. 12, no. 3, pp. 299–320, Jul. 2013.

[33] S. Gupta and S. Shrotriya, "The leaky bucket of female talent in the indian it industry."