Modeling Participation Behaviors in Design Crowdsourcing Using A Bipartite Network-Based Approach

Zhenghui Sha*
Assistant Professor
Department of Mechanical Engineering
University of Arkansas
Fayetteville, Arkansas 72701

Ashish M. Chaudhari
Graduate Research Assistant
School of Mechanical Engineering
Purdue University
West Lafayette, Indiana 47907

Jitesh H. Panchal
Associate Professor
School of Mechanical Engineering
Purdue University
West Lafayette, Indiana 47907

Abstract This paper analyzes participation behaviors in design crowdsourcing by modeling interactions between participants and design contests as a bipartite network. Such a network consists of two types of nodes, participant nodes and design contest nodes, and the links indicating participation decisions. The exponential random graph models (ERGMs) are utilized to test the interdependence between participants’ decisions. ERGMs enable the utilization of different network configurations (e.g., stars and triangles) to characterize different forms of dependencies, and to identify the factors that influence the link formation. A case study of an online design crowdsourcing platform is carried out. Our results indicate that designer, contest, incentive, and factors of dependent relations have significant effects on participation in online contests. The results reveal some unique features about the effects of incentives, e.g., the fraction of total prize allocated to the first prize negatively influences participation. Further, we observe that the contest popularity modeled by the alternating k-star network statistic has a significant influence on participation, whereas associations between participants modeled by the alternating 2-path network statistic does not. These insights are useful to system designers for initiating effective crowdsourcing mechanisms to support product design and development. The approach is validated by applying the estimated ERGMs to predict participants’ decisions and comparing with their actual decisions.

1 INTRODUCTION

In design crowdsourcing, a contest sponsor specifies a design problem to a crowd of engineers or designers. Participants decide how much effort and resources to expend working on solutions, and submit their solutions to the contest sponsor. The contest sponsor evaluates all solutions and picks the best solutions. The contest sponsor awards monetary prizes or gifts with comparable value to the participants with winning solutions. Such an execution of design crowdsourcing is called a crowdsourcing contest.

For the effective use of crowdsourcing contests in engineering design applications, it is necessary to understand how the characteristics of the design problem and payment structure influence participants’ decisions. The existing game-theoretical studies of crowdsourcing contests predict that contest-related factors such as the number of prizes and entry fee may increase or decrease participant effort and the quality of solutions depending upon participants’ risk preferences and homogeneity in their problem-specific knowledge [1–4]. The results of our prior studies on design crowdsourcing contests indicate that participants’ decisions depend not only on contest-related factors, such as the number of prizes, payment structure, and prize amounts, but also on problem-related factors, such as design problem, problem description, cost of resources and evaluation criteria [5–8].

Although existing studies are useful in predicting the effects of various factors on contest outcomes, there is a lack of knowledge about how inherent relations between participants or between contests influence participation decisions. On online crowdsourcing platforms such as GrabCAD [9]
and Amazon Mechanical Turk [10], multiple crowdsourcing contests are open for participation at the same time, and a participant also makes decisions about which contests to participate in. In such cases, the external connections between participants and the stated characteristics of contests are likely to generate dependencies. If two participants know each other through a common workplace, their individual participation decisions may be dependent (peer effect). If two participants are aware of each others’ submissions in past contests, their decisions to participate in the same contest in the future are related (association effect). A participant may decide to participate in a contest because many other participants are participating in the same contest (popularity effect). Two contests with the same sponsor, similar design problems, or similar payment structure may attract participation of the same set of participants (homophily effect).

The primary objective of this study is to understand whether and to what extent the aforementioned participant and contest dependencies influence the participation in design crowdsourcing contests. Understanding the role of dependencies in participation will help in identifying a set of the most critical factors that drive participation decisions, establishing better predictive models of participation, and for creating design guidelines for future contests. We also evaluate the effects of problem- and contest-related factors on participation with and without incorporating dependencies. Although there are various ways to evaluate the contest outcomes [11], we focus on participation because the diversity and the quality of solutions depends on the number of people participating in a contest. For instance, more number of participants implies a greater number of solutions, and a potential for greater diversity. For the design problems where specific designs need to be improved, a greater number of participants means higher chances of achieving a better solution.

In this study, we propose a bipartite network-based approach to modeling participation decisions in crowdsourcing contests between multiple participants and multiple contests. Each link in a network represents an instance of participation between a participant and a contest. For a scenario with a single contest and multiple participants, a star network is a natural representation, where the focal node denotes a contest while the connected nodes denote its participants [12]. On the other hand, a bipartite network is more suitable for the a design crowdsourcing scenario with multiple contests and participants. A bipartite network representation divides nodes into two groups, called nodes. Each node is the focal node of a star network connecting it to a set of nodes in the other mode (see Figure 1). The bipartite network representation assumes that a link exists between nodes from different modes, so that a connection between two participants or two contests is irrelevant when representing participation decisions.

With a bipartite network setting, we employ the Exponential Random Graph Model (ERGM) to model participation decisions as the link formation process in network. The ERGM models the probability of a link formation in terms of the statistics of node attributes and the assumed network configurations. The coefficients associated with these statistics represent the effects of problem- and contest-related factors, and the effects of dependencies in the formation process representing participants’ decision-making in a crowdsourcing contest. This approach is demonstrated in a case study on participation decisions in GrabCAD, an online design crowdsourcing platform [9].

The rest of the paper is structured as follows. Section 2 provides a review of literature on crowdsourcing studies as well as the applications of ERGM in engineering design. Section 3 presents theoretical details of ERGM and the network configurations adopted in this study. Section 4 introduces GrabCAD contests, its participants and the corresponding dataset. Section 5 discusses the implementation of ERGMs on the GrabCAD dataset, and the results of the analysis. Section 6 presents evaluation and validation of the ERGMs estimated from the GrabCAD dataset. Finally, Section 7 summarizes the paper’s findings, limitations and future work.

2 REVIEW OF RELEVANT LITERATURE

In this section, a review of the literature on crowdsourcing studies and ERGM research in engineering design are presented, and the research gaps are highlighted.

2.1 Crowdsourcing studies in engineering design

In the theoretical studies of crowdsourcing contests, game-theoretic models have been developed to model the effects of prize structure on participants’ effort and the resources (cost) they invest. The authors’ previous work summarizes such models and insights from them [6][8]. The results in these models are based on the assumptions placed on participant attributes. For example, the models by Szymanski and Velentti [13] and Clark and Riis [14] show that a single prize motivates greater effort than multiple prizes of the same total amount if all participants are homogeneous, i.e., they incur the same cost for the same effort and have risk-neutral preferences. However, if participants have non-homogeneous costs, then, according to Sheremeta [4] and Sisak [3], multiple prizes are better than a single prize of the same amount. Other models have been developed to
compare fixed prize contests and auctions where participants bid on a prize that they want to be paid for their submissions \[2,15,16\].

Due to the current lack of models to account for all the factors influencing participant behaviors in crowdsourcing, many studies employ empirical data. The empirical studies \[17\]–\[23\] investigate the effects of intrinsic motivation, such as enjoyment, curiosity, developing individual skills, and extrinsic motivation, such as monetary rewards, improving job prospects, and gaining peer recognition, on participation and effort in contests. Apart from the effects of motivation, the empirical studies also evaluate the effects of problem-specific factors such as task granularity, problem complexity, specificity, and solvability on solution quality \[24\]–\[26\], as well as sponsor-specific factors such as reputation, trust in sponsor and sponsor feedback on participation and solution quality \[27\]–\[29\]. Some studies \[30\]–\[32\] also investigate how participants’ expertise, experience, and familiarity with the problem affect contest outcomes. The domains of contest problems in these studies vary from graphic design, logo design, CAD design, product design, and software engineering to e-commerce.

Despite a growing interest and various research efforts, the existing crowdsourcing studies analyze participation and effort decisions based on the characteristics of a single contest, or when analyzing multiple contests, they assume that participation decisions are independent. We address this gap by quantifying dependencies using network structures and explicitly model the effects of dependencies on participation decisions in crowdsourcing using a network-based approach.

### 2.2 ERGM in engineering design

The exponential random graph model (ERGM) has been a popular approach in social science for modeling the social network formation \[33\]–\[35\] due to its capability to model complex social relations as network topologies. A rich literature exists on modeling conditional dependence between network links using ERGM specifications such as dyadic independence assumption \[35\], the Markov assumption \[37\], and the social circuit assumptions \[38\]–\[39\], both in bipartite settings \[40\] and in multilevel settings \[41\].

In recent years, there has been a trend towards using ERGMs in engineering design literature for modeling complex sociotechnical systems and understanding customer decision-making behaviors. The extension of ERGMs to engineering design is natural because it enables incorporating social interactions and human factors in design process to better model system architectures. In addition, the unique strength of ERGMs in modeling cross-level network interconnections \[40\]–\[41\] has made it a powerful tool for studying hierarchical structures and multilevel relations among different information layers in complex systems. ERGMs have been used in engineering design from different aspects. For example, Wang et al. \[42\] used a multidimensional ERGM to study customer preferences in vehicle markets by considering both social interactions and product associations.

Sha and Wang studied products’ co-consideration relations with a unidimensional ERGM \[43\], and compared its predictive performance with network-based logistic regression model \[44\]. Fu et al. \[45\] adopted the bipartite network setting of ERGM and studied the choice behaviors in support of engineering design. These studies have inspired us to further extend the capabilities of ERGM in engineering design, and to evaluate the feasibility of network configurations and the interdependence assumption in analyzing participation behaviors in design crowdsourcing contests.

### 3 MODELING INTERDEPENDENCE AMONG PARTICIPATION BEHAVIORS

In this section, we provide an overview of exponential random graph models. Then, the network configurations used to represent dependencies as well as their mathematical expressions are introduced in Section 3.2.

#### 3.1 Overview of exponential random graph models (ERGM)

An ERGM, first introduced by Frank and Strauss \[37\], interprets the global network structure as a collective result of various local network effects, called network configurations. The key idea is that an ERGM considers an observed network, \(y\), as one specific realization from a set of possible random networks, \(Y\), following the distribution in Equation (1). A network \(y\) can be represented by an adjacency matrix of binary values \(y_{ij}\), where 1 indicates the existence of a link between nodes \(i\) and \(j\), and 0 indicates the absence of the link. The probability of a random matrix \(Y\) equal to the adjacency matrix of a particular network \(y\) is:

\[
Pr(Y = y) = \frac{\exp(\theta^T G(y))}{\kappa(\theta)} ,
\]

where \(\theta\) is a vector of model parameters. \(G(y)\) is a vector of statistics for network configurations and nodal attributes, and \(\kappa(\theta)\) is a normalizing coefficient to make sure the probability has value between 0 and 1. Equation (1) suggests that the probability of observing any particular network is proportional to the exponent of a weighted combination of network characteristics.

With such a modeling framework, the objective of a researcher is to formulate an appropriate model structure that represents real-world systems through the identification of network configurations as well as the attributes of nodes and links. In ERGM, the attributes of nodes are represented through the nodal covariates \[46\], whereas network configurations are represented through network statistics. Such model structure provides a plausible and theoretically principled hypothesis for the network generation process. Once a model structure is determined, the parameter \(\theta\) can be readily estimated with training data using maximum likelihood estimates or Markov Chain Monte Carlo simulations \[47\].
The contest k-star can be used to model the popularity effect of a design contest. The underlying reason for the formation of k-star (why a contest is popular) does not necessarily have correlations with its own attributes (e.g., the prize amount) because popularity is a social phenomenon.

The k-star statistic $S_k(y)$, $k = 2, \ldots, n - 1$ denotes the number of k-stars in the graph $y$. For example, the k-star statistic for participants quantifies the number of participants who have participated in $2, 3, \ldots, n - 1$ contests, respectively. Recent work by Snijders et al. [39] considers a more refined model based on the k-star statistics. It is defined as the scalar-valued network statistic using the following equation:

$$g^{Aks}(y; \lambda_s) = S_3(y) - \frac{S_3}{\lambda_s} + \ldots + (-1)^{n-3} \frac{S_{n-3}(y)}{\lambda_{n-3}^{n-3}}$$

where $\lambda_s$ is a positive scalar quantifying the decaying effect of large $k$ on this network statistic, i.e., a larger $k$ has a lower effect on the overall k-star statistic than a smaller $k$. $\lambda_s$ can be treated as a constant defined by researchers or treated as a parameter, similar to $\theta$ in Equation (1), estimable from empirical data. Because of the alternating signs in front of the $S_k(y)$ terms, Snijders et al. [39] call $g^{Aks}(y; \lambda_s)$ an AKS statistic.

b) Alternating 2-paths configuration (A2P): The A2P configuration is used to account for the interdependence arising from the shared events. In design crowdsourcing, the inclusion of this effect allows us to answer the following question: if two participants have one challenge in common (meaning a shared event), are they more likely than expected by chance to have a second challenge in common, and a third one and so on? It tests whether the distribution of observed non-random shared events is driven by pairs of participants. A possible reason for the formation of shared events is that two participants may have common interests. It is important to note that such a dependency may be implicit to participants and not explicitly known. Knowing the precise common interests, however, may not be of interest to researchers as long as the effect of that dependency is understood, which can be realized by A2P network configuration.

A k-two-path is a set of $k$ distinct paths of length two joining the same pair of nodes (see Figures 2(c) and 2(d)). A 1-two-path is the same thing as a 2-star [39]. However, a 1-two-path is usually associated with its endpoints whereas a 2-star is usually associated with its center. If the number of k-two-paths with participants as endpoints in the network $y$ is $P_k(y)$, where $k$ may be any value from 1 to $n - 2$, then the A2P statistic is [48].

$$g^{A2P}(y; \lambda_p) = P_3(y) - \frac{P_3}{\lambda_p} + \ldots + (-1)^{n-3} \frac{P_{n-2}(y)}{\lambda_{n-3}^{n-3}}$$

where $\lambda_p$ has the similar effect as $\lambda_s$ does in Equation (3). This configuration can help quantify the potential effects of the relations among contests (e.g., sub-contest relations) and the relations among participants (e.g., friends) on the participation decisions.

Fig. 2. Two types of network configurations for modeling interdependence of participation decisions.

The final model structure is often achieved through an iterative process by trials of different combinations of network configurations supported by domain expertise or feature selection algorithms in data mining.

ERGM is used to describe different kinds of interdependence using various network configurations, and moreover, it can quantitatively assess the importance of such interdependence during the network formation process. ERGM is used to describe different kinds of interdependence with network configurations $(\theta g_s)$: A $g_s$ is a k-star $k$-star configuration $(AkS)$: The $AkS$ statistic $A_2P$ statistic is [48]...
4 FIELD DATA FROM GRABCAD CONTESTS

With the ERGM and those identified network configurations discussed above, in this section, we demonstrate our approach of modeling participation behaviors in a case study on GrabCAD – an online design crowdsourcing platform.

4.1 Introduction to GrabCAD

GrabCAD [9] is a community of individuals who contribute to an online library of computer-aided-design (CAD) files. This library is open to online crowds, and CAD files are available to everyone for free. Over 4 million members have contributed more than 2 million CAD files. Members of the community hail from various regions across the world. Members’ areas of expertise in design are also wide ranging. They include interior-, graphic-, industrial-, and product-design.

Manufacturers and GrabCAD sponsors leverage the large crowds on GrabCAD to solve problems, create new designs or promote their brands using design contests. The contests are conducted over a period of few days, and consist of problem descriptions, solutions from community members in the form of CAD files, and prize information for top solutions. Each design challenge starts with posting a problem statement, supplementary files if the problem statement involves using or improving existing designs, prize amounts to be awarded to top solutions, and a description of requirements upon which solutions are evaluated. During the active phase of a challenge, the GrabCAD community members upload solutions as STEP, IGES files, or rendering images along with any analysis results required by the problem statement. All submissions are accessible by the public to view and download during the time when the challenge is active. Members, at the time of uploading their solutions, are therefore able to know the solutions that others have submitted already. At the end of a challenge, a panel of judges from the sponsor company and/or GrabCAD staff evaluate all submissions, and determine finalists to award and recognize their submissions. The finalists are made public by posting weblinks to their profile and web-links to their submission.

4.2 Data collection

We analyzed all contests hosted on GrabCAD platform between 2011 and 2018. From this dataset, we filtered out contests that involved promotional tasks with no specific problem statement or prizes, e.g., community contests with no prizes hosted by GrabCAD to promote crowd’s involvement. We removed contests that have crowd submissions set to private. Such contests are dropped from our analysis because participants’ actions in them are either unknown, or most likely caused by intrinsic factors which can not be observed. Finally, we selected data of 117 contests which provides measurable information on influencing factors and outcomes of contests. 96 contests conducted between year 2011 to 2016 were used for training ERGMs, while the remaining 21 contests were used for validation. Scrapy [49], a Python-based web crawler, was used to capture the visible data in accordance with regulations of GrabCAD.

The dataset for each challenge includes the description of the sponsor company, the number of monetary prizes and gifts, corresponding prize amounts (worth price for gifts), the number of entries/submissions, the unique identity of each submission and corresponding participant, and the identity of the winners. Based on participant identities and their submissions, the number of participants for a challenge is derived after removing repetitive and blank submissions so that each participant is counted only once. The number of prizes is counted as the sum of number of monetary prizes and number of gifts. Based on the unique identity of every participant, we estimate the number of times a participant is mentioned as a finalist. The winning rate for a participant is the ratio of the number of times he/she is a finalist to the total number of contests in which he/she participated.

4.3 Descriptive analysis of GrabCAD network

With the collected GrabCAD data, we construct a bipartite network (see Figure 3). We first perform descriptive network analysis to have an overall understanding of the network and the participation behaviors in GrabCAD contests. The degree indicates a node’s connectivity to other nodes which is an important measure of network properties [50]. The average degree of participant nodes in the GrabCAD network is 1.87. This indicates that on average one participant participated in about two design contests. The average degree of contest nodes is 67.47. It means that one contest has about 65 participants on average.

Figure 4 shows the degree distribution of participant nodes and contest nodes, respectively, on log-log scale. The results indicate certain aggregate-level participation behaviors. For example, it is observed that a few contests attracted a large number of participants: 300 participants submitted total of 637 entries to the GE Jet Engine Bracket Challenge, 297 participants submitted 461 entries to the NASA Handrail Clamp Assembly Challenge, and 227 participants submitted 390 entries to the Ultimaker 3D Printer Toy Design Challenge. These three design contests are shown in Figure 3. Among 96 design contests, there are 25 contests each of which has attracted more than 100 participants.

In addition, we observe that most people (2661 out of 3462) just participated in one design contest. Only a few participants participated in many contests. These participants can be regarded as active GrabCAD players. For example, there are 159 people out of 3462 participated in 5 design contests. And only two people participated in more than 30 design contests. As of 2016, the top 3 designers have participated GrabCAD design contests 42, 37, and 29 times, respectively.

From the degree analysis, we observe that on GrabCAD, some of the contests are very popular, and some of the participants are very active and motivated. These characteristics indicate that the existing mechanism of GrabCAD indeed influences participants’ behaviors and the network structure...
observed in Figure 3. But the descriptive analysis in itself is not sufficient to answer questions about how quantitatively the popularity effects and effects of other factors affect participation decisions. In the next section, we use ERGMs to establish a statistical inference model to answer these questions and provide insights on the participation behaviors on GrabCAD.

5 IMPLEMENTATION OF ERGMs

In this section, we discuss the ERGM implemented on the GrabCAD network with different settings. Specifically, in Section 5.1 we specify the assumptions for incorporating various contest and participant attributes as well as the network configurations into the ERGM. Section 5.2 presents the results and discussion.

5.1 Model settings for the GrabCAD network

Table 1 lists the attributes we consider for the ERGM implementation. Our prior analysis of the GrabCAD data identifies that task complexity, clarity of problem specification, prize amounts, and the number of prizes influence the participation [8]. In this paper, with the network representation of the GrabCAD data, we evaluate the effects of additional attributes such as a participant’s past success (e.g., winning rate) and past recognition (e.g., the frequency of being a finalist), and dependencies such as contest popularity and associations.

a) Attributes of design contests. The first attribute is the type of design task in a contest. If a task requires the design of a single component, we refer to it as a part-design task. If a task involves the design of multiple components and a mechanism of interaction between them, we refer to it as a system-design task. For example, Autodesk Robot Gripper Arm Design Challenge [51] involved a design of a robotic arm where the objective was to minimize the weight of the arm. Since the task was to design a single component, i.e., the arm, it is a part design task. In contrast, the Dirt-bike Tire Changing Tool Challenge [52], for instance, observed a task to design a tire-changing tool with multiple parts such as tool
frame, lever, and gears. The solution required contestants to design a mechanism for interaction between the tool, its parts and tire. This task is therefore a system-design task. We use a variable, isPart, to indicate whether the task is a part design task (isPart = 1) or a system design task (isPart = 0).

Furthermore, we categorize the GrabCAD contests into ideation- and expertise-based types based on whether the problem requirements and judging criteria are clearly specified. Following the definitions by Terwiesch and Xu [53], the contests with high uncertainty in evaluation are categorized as ideation contests, whereas the contests with low uncertainty in evaluation are categorized as expertise-based contests. For example, the Alcoa Airplane Bracket Bearing Design Challenge [54] asked the participants to optimize an existing bracket design. The contest sponsors used well-defined requirements such as the ratio of ultimate strength and weight to evaluate the submissions. Also, the contest required participants to use topology optimization methods, because of which the technical uncertainty was low. Thus, this contest was an expertise-based challenge. On the other hand, in the Robot Gripper Test Object Challenge [55], participants were asked to come up with ideas for test objects to evaluate a robot gripper, but the evaluation criteria were not well-defined. Any implementation of test object can be verified using simulation or prototypes. However, unclear problem requirements result in uncertainty in comparing any two objects. This challenge, therefore, fits into the category of ideation-based contests. We use another variable, isIdea, to indicate whether a contest is an ideation-based design problem (isIdea = 1) or an expertise-based design problem (isIdea = 0).

We use an additional attribute, the number of entries \(N_e\), as an indicator of a contest’s overall attractiveness. \(N_e\) counts multiple entries (submissions) by the same participant as well as the blank entries, i.e., only web pages but with no files attached.

b) Attributes of designers. The designer attributes include two continuous variables, the winning rate \(R_w\) and the number of times that a designer’s solution was mentioned in the finalists of a challenge \(N_f\). These two variables reflect the overall skills and problem-solving ability in design crowdsourcing. The number of wins is not adopted because there is a potential correlation between the number of wins and the number of participation. So people who participated many times may win more number of times. Using frequency in this situation may overestimate the skills of a participant. The adoption of percentage is especially suitable when the occurrence likelihood of an event is low, such as the winning event in this context.

However, there is also a downside accompanied with adopting percentage. The downside arises when the occurrence likelihood of an event is high. For example, in the GrabCAD, \(N_f\) is counted as the number of times a design is ranked in top ten. We observe from the data that there are many cases where a person participated once and she/he was in the finalist. This result indicates that this person is extremely skilled, leading to an unfair comparison to a really skilled participant who may have been a finalist 50% of times out of a large number of contests where she/he may have participated in. Therefore, using percentage in this situation may overestimate the skills of a participant. Other attributes pertaining to designers’ demographics may have an impact on the participation behaviors but are not considered in this paper, mainly due to the data unavailability.

c) Incentive structure. The incentive mechanism is a critical component that impacts the success of crowdsourcing. In the literature on tournament compensation [56], researchers commonly predict that status, power, or money incites competition and differentiation, which are beneficial to crowdsourcing. However, as indicated by Bloom and Michel [57], “research has yielded mixed results about what amount of pay dispersion is optimal. In some cases, more dispersed pay distributions have been positively related to performance outcomes. In other cases, greater dispersion has been negatively related to performance outcomes”. Therefore, the role of monetary incentives in design crowdsourcing is not completely understood. In this study, we create a set of variables to study the effects of monetary incentive in design crowdsourcing.

The structure of monetary incentive is often studied in terms of the difference between prizes. Harrison and Klein [58] suggest two indices to quantify the difference, the coefficient of variation (COE) and Gini coefficient. We adopt the COE in our model. In most contests, the first prize is an important factor that influences participation [59]. To account for the effect of the first prize, we include both the fraction value \(F_{p1}\) as well as the absolute amount \(M_{p1}\) of the first prize in our model.

In the GrabCAD contests, prizes are awarded in two ways, through monetary prize or through gifts (e.g., iPad). To capture the impact of the prize type, we create a categorical variable prize level, \(L = \{L1, L2, L3, L4\}\). In this study, prize level \(L1\) includes contests that have neither the monetary prize nor gifts; prize level \(L2\) includes contests with only have gifts; prize level \(L3\) includes the contests which have only monetary prizes; while prize level \(L4\) includes the contests which have both monetary prize and gifts. In summary, we use the four variables, the COE, \(F_{p1}\), \(M_{p1}\), and prize levels, to model the monetary incentives in design crowdsourcing.

d) Network configurations. As introduced in Section 3, we consider the alternating k-star statistics \(g^{AkS}\) for participant k-stars and \(g^{AkS}\) for contest k-stars and the alternating 2-path statistics \(g^{A2P}\) for contest k-stars to model interdependence between links. The former enables the assessment of the contest popularity effect (contest AkS) and the level of participant’s engagement in GrabCAD contests (participant AkS) on the formation of participation links. The latter enables the investigation of significance of shared events/partners in the GrabCAD network. Since the current ERGM does not support the two-mode analysis of A2P, we use the one-mode version to capture the overall effect of the shared events regardless of the modes of nodes. In ERGM, a network configuration, called edge, is included to control the number of links in a network to ensure the regenerated networks have the same density as the observed ones. The edge estimates
Table 1. Summary of the variables and their statistics. For continuous variables, minimum and maximum along with the average and the standard deviation are reported. For categorical or binary variables, the number of instances is reported.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Descriptions</th>
<th>Statistics</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Contest</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of entries ($N_e$)</td>
<td>The number of solution entries observed for a challenge (may include multiple entries by the same designer or blank entries).</td>
<td>Min:2; Max:638; Avg:113.90; Std:112.36</td>
</tr>
<tr>
<td>Ideation-based or expertise-based (isIdea)</td>
<td>Whether a contest is an ideation-based design problem or an expertise-based design problem.</td>
<td>Ideation-based: 59; expertise-based: 37</td>
</tr>
<tr>
<td>Part design or system design (isPart)</td>
<td>Whether a contest is a part design problem or a system design problem.</td>
<td>Part design: 46; System design: 50</td>
</tr>
<tr>
<td><strong>Participant</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of times being in finalist ($N_h$)</td>
<td>The number of times a designer’s solution are mentioned in the finalists of a challenge.</td>
<td>Min:0; Max:5; Avg:0.10; Std:0.36</td>
</tr>
<tr>
<td>Winning rate ($R_w$)</td>
<td>The percentage of wins of a participant regardless of the level of prize.</td>
<td>Min:0; Max:1; Avg:0.03; Std:0.15</td>
</tr>
<tr>
<td><strong>Incentive structure</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1st prize ($M_{p1}$)</td>
<td>The absolute total value of the first prize including the value of gifts.</td>
<td>Min:0; Max:$10000$; Avg:$1575.97$; Std:$1649.28$</td>
</tr>
<tr>
<td>1st prize fraction ($F_{p1}$)</td>
<td>The percentage of the value of the first prize including the value of gifts.</td>
<td>Min:0; Max:1; Avg:0.55; Std:0.25</td>
</tr>
<tr>
<td>Coefficient of variation (COE)</td>
<td>The disparity between different levels of prize.</td>
<td>Min:0; Max:4.80; Avg:0.68; Std:0.75</td>
</tr>
<tr>
<td>Prize level L1</td>
<td>Neither do the contests have monetary prize nor gifts.</td>
<td># of instances: 3</td>
</tr>
<tr>
<td>Prize level L2</td>
<td>The contests only have gifts.</td>
<td># of instances: 14</td>
</tr>
<tr>
<td>Prize level L3</td>
<td>The contests only have monetary prize.</td>
<td># of instances: 65</td>
</tr>
<tr>
<td>Prize level L4</td>
<td>The contests have both monetary prize and gifts.</td>
<td># of instances: 14</td>
</tr>
</tbody>
</table>

The likelihood that a participant chooses a design contest randomly if no knowledge about the participant or design contest attributes is provided. It is analogous to the intercept term in a linear regression model.

5.2 Results and discussion

With the model settings introduced in Section 5.1, we define two ERGMs based on Equation (1). Model 1 assumes that the links are formed independent of each other, thus does not incorporate network configurations. Model 2, on the other hand, incorporates the network configurations introduced in Section 3.2 to model dependencies. So, Model 2 includes the variables $g_{\text{MAK}}, g_{\text{AKS}},$ and $g_{\text{AP}}$, but Model 1 excludes them. Table 2 presents the model parameters estimates and relevant performance metrics calculated using the statnet package in R program [46].

As shown in Table 2, regardless of the differences in network configurations, both models agree on certain participation behaviors from GrabCAD data. Participants are more likely to participate in design contests with part design problems than ones with system design problems. They are more likely to participate in ideation contests than in expertise-based contests. For example, Model 1’s results indicate that the probability of an individual participating in a part design contest is about 1.5 (i.e., $e^{0.407}$) times that in a system design contest. The probability of participating in an ideation-based design contest is 1.16 times that in an expertise-based contest.

For given design contests, links are more likely to be formed with the participants who have higher winning rate or have been finalists more often. In other words, the participants who have won or been recognized have a higher probability of participation. Since all the attributes of design contests and participants are normalized to $[0, 1]$, the estimated parameters allow comparison of effects across different variables. Accordingly, the number of times being a finalist is the most influential among all the independent variables (excluding the network configurations).

The effect of the 1st prize amount is not statistically significant for the GrabCAD network. However, the estimated effect for the 1st prize fraction is statistically significant and negative. This indicates that higher the fraction of the 1st prize, lower is the probability of link formation between a designer and a design contest. These two results imply that, in GrabCAD contests, participants more likely to participate in contests that have lower allocation to the 1st prize. The larger the allocation to the 1st prize, the smaller is the probability of winning for a participant. With a lower percentage of the 1st prize, the total prize is distributed across multiple prizes, and the probability of winning prize amount through lower places such as the second or third prizes is higher.

The parameter estimates for prize level indicate that the type of prize affects participation. With no prize and no gift (Prize level L1) as a baseline, the positive signs of the three prize levels in Table 2 show that participation is higher in
Table 2. Estimated parameters of two ERGMs for the GrabCAD network. Model 1 assumes link independence (i.e., no network configurations are included), while Model 2 takes link dependence into account (i.e., three network configurations are included).

<table>
<thead>
<tr>
<th>Explanatory Var.</th>
<th>Model 1</th>
<th>Model 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Est. θ</td>
<td>Std. Err.</td>
</tr>
<tr>
<td></td>
<td>Mean</td>
<td></td>
</tr>
<tr>
<td>Number of entries (N_e)</td>
<td>2.234 ***</td>
<td>0.074</td>
</tr>
<tr>
<td>isPart = 1</td>
<td>0.407 ***</td>
<td>0.038</td>
</tr>
<tr>
<td>isIdea = 1</td>
<td>0.145 ***</td>
<td>0.031</td>
</tr>
</tbody>
</table>

Attributes of participants

<table>
<thead>
<tr>
<th></th>
<th>Number of times being in finalists</th>
<th>Winning rate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(N_e)</td>
<td>(R_w)</td>
</tr>
<tr>
<td></td>
<td>3.913 ***</td>
<td>0.092</td>
</tr>
<tr>
<td></td>
<td>0.326 ***</td>
<td>0.092</td>
</tr>
</tbody>
</table>

Incentive structure

<table>
<thead>
<tr>
<th></th>
<th>1st prize amount (M_1)</th>
<th>1st prize fraction (F_1)</th>
<th>Coefficient of variation (COE)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(M_1)</td>
<td>(F_1)</td>
<td>(COE)</td>
</tr>
<tr>
<td></td>
<td>0.130 **</td>
<td>-0.273 ***</td>
<td>1.131 ***</td>
</tr>
<tr>
<td></td>
<td>0.070</td>
<td>0.067</td>
<td>0.016</td>
</tr>
<tr>
<td></td>
<td>0.025</td>
<td>-0.185 **</td>
<td>0.121 ***</td>
</tr>
<tr>
<td></td>
<td>0.082</td>
<td>0.067</td>
<td>0.015</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Prize level L2</th>
<th>Prize level L3</th>
<th>Prize level L4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.141 **</td>
<td>1.295 **</td>
<td>1.263 ***</td>
</tr>
<tr>
<td></td>
<td>0.200</td>
<td>0.196</td>
<td>0.198</td>
</tr>
<tr>
<td></td>
<td>0.234 **</td>
<td>0.370 **</td>
<td>0.344 **</td>
</tr>
<tr>
<td></td>
<td>0.134</td>
<td>0.127</td>
<td>0.129</td>
</tr>
</tbody>
</table>

Network configurations

<table>
<thead>
<tr>
<th></th>
<th>edge</th>
<th>Participant alternating k-star (g_AkS)</th>
<th>Contest alternating k-star (g_AkS)</th>
<th>Alternating 2-paths (g_A2P)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-6.164 ***</td>
<td>-0.190</td>
<td>-0.080</td>
<td>0.00092</td>
</tr>
<tr>
<td></td>
<td>0.134</td>
<td>0.079</td>
<td>0.366</td>
<td>0.00048</td>
</tr>
</tbody>
</table>

Model fit

<table>
<thead>
<tr>
<th></th>
<th>Null deviance</th>
<th>BIC</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>460738</td>
<td>59127</td>
</tr>
<tr>
<td></td>
<td>460738</td>
<td>59064</td>
</tr>
</tbody>
</table>

Significance codes: ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1

In affecting participants’ behaviors. The more participants

We do not observe significant effects of the participant

In the actual implementation, the network statistics gwdegree (called geometrically weighted degree, equivalent to AkS) and gwdsp (called geometrically weighted dyadwise shared partner, equivalent to A2P) were used. The interpretation to the sign of the estimated coefficients is opposite. Please refer to [42] for more details.
recognized and popular, or the sponsor engaged in providing feedback to participants \[8][28].

As to the overall model fit, Model 2 is not greatly improved compared to Model 1. This is indicated by the Bayesian information criteria that is the lower, the better. Model 2’s BIC (59064) is only slightly lower than Model 1’s (59127). To better evaluate the models’ performance and validate the capabilities of those dependence settings, further investigation on their predictive power is performed in the following section.

### 6 MODEL EVALUATION AND VALIDATION

In this section, we analyze the predictive power of Model 1 and Model 2 for predicting links in the GrabCAD network. We evaluate model predictions using the training data of 96 contests, and cross-validate using the test data of 21 contests.

We follow a two-step approach for model evaluation and validation. First, the model estimates given in Table 2, the data of the explanatory variables and the observed GrabCAD network are used together with the ERGM model given in Equation (1) to calculate the probability of whether a link will be formed. Based on the probabilities of individual link formation, synthetic networks are generated. Second, the regenerated networks are compared with the observed links in the real network and the prediction accuracy can be quantified by various measurements such as Precision, Recall (a.k.a. Sensitivity), Specificity, and Accuracy. The confusion matrix [61] in Table 3 summarizes these measurements.

We mainly use Precision and Recall to assess the models’ predictive performance, because the GrabCAD network is imbalanced. The number of all possible links between the participants and the contests is 3462 × 96 (i.e., 332352), out of which the number of observed links is 6432 and 325920 links are negatives. Since a model is more likely to predict negative cases under the imbalanced scenario (the random probability is about 98%), counting the predicted positive cases using Precision and Recall better reflects the predictive capability of the model [62].

Table 4 shows the results of model evaluation and prediction. With each model, 200 synthetic networks are generated in the first step. In the second step, at the threshold of the probability 0.5, Model 1 yields the Precision and Recall at 0.0037 and 0.36, respectively. Model 2 yields the Precision and Recall at 0.0044 and 0.43, respectively. The results show that among the observed 6432 cases of participation, Model 1 predicts 36% of them while Model 2 successfully predicts 43% of them indicating that Model 2 has a better predictive performance than Model 1. Note that the Precision values in both models are very small, because at the threshold of 0.5, only 67 and 65 links are predicted to exist in the network with the two models, respectively. Among these predicted links, Model 1 correctly predicts 24 cases, whereas Model 2 correctly predicts 28 cases. That is also why Model 2 has higher recall value than Model 1. As a baseline for reference, the random network model generated by randomly forming links at the threshold of 0.5 has Precision and Recall of 0.00031 and 0.056, respectively.

Alternatively, we can vary the threshold for linking probability such that the number of links of a predicted network is closest to that of the real network, i.e., 6432 links.\[4\] With this setting, Model 1 has both Precision and Recall of 0.1, Model 2 delivers the same performance compared to Model 1, while the random network model produces both Precision and Recall of 0.045.

In addition to the evaluation on the regenerated GrabCAD networks, we perform a cross-validation using the test network of 21 GrabCAD contests that were conducted after the month of July 2016 till August 2018. The test data includes a total of 1486 participants and 1950 number of participation instances (i.e., the number of links in the network is 1950). The results of cross-validation are presented in Table 4. The precision and recall are not obtainable for Model 2 and the random network model with the fixed probability threshold of 0.5, because the link formation probabilities are

\[3\] BIC is a commonly used statistical metric for evaluating model fit, calculated by the formula, $BIC = -2\log L + \log(n)p$, where $L$ is the log likelihood, $p$ is the number of explanatory variables, and $n$ is the sample size.

\[4\] Model 1 produces 6474 links with the probability threshold of 0.085. Model 2 produces 6537 links with the threshold of 0.085. And, the random network model produces 6429 links with the threshold of 0.185.
very small and no links are predicted. In the other scenario where the number of links is kept closest to that of the real network, Model 2 outperforms both Model 1 and the random network model.

The predictive performance of both models still has room for improvement. The limitation mainly comes from two sources. First, the network is a highly imbalanced meaning a significant higher number of negative cases than positive cases. So predicting the positive cases (i.e., the existing links) is naturally challenging. This leads to a small number of correctly predicted links, thus a low precision. Second, the models developed in this study consider a small number of factors to explain the participants’ decisions. Many other explanatory variables need to be explored further. For example, other factors of dependent relations, such as social ties between more than two participants, could play a role. But the lack of appropriate network configurations makes the modeling of those relations challenging at this stage. Despite its limitations, this work demonstrates how network structures are used to represent different forms of dependencies and study their effects on participation decisions in crowdsourcing contests. Further, the purpose of the prediction analysis using the models is not to support the forecast of participation decisions per se. It is to compare performance of the model with dependencies assumed (Model 2) against the benchmark model without dependent settings (Model 1), and against the random network model. The results support our hypothesis in this regard.

7 CONCLUSIONS AND FUTURE WORK

In this paper, we study the participation behaviors in online design crowdsourcing contests assuming that there exists dependencies among participants’ decisions to participate. We propose a network-based approach which uses the ERGM in a bipartite setting to represent participation decisions. This approach enables a novel combination of advanced network modeling techniques and real crowdsourcing data for behavioral analysis. The results indicate that contest popularity, modeled by alternating k-star network statistics, significantly influences whether a participant would choose a design contest or not. Associations among participants or among contests, modeled by alternating 2-path network statistics, do not have a significant influence on the participation.

In addition to the findings related to the network configurations, we obtain following insights into the design of crowdsourcing contests:

1. Participants are more likely to participate in contests related to part-design, compared to system-design. Therefore, sponsor companies may consider first decomposing a system design problem into several integrated part design problems, and then using GrabCAD to solicit potential solutions for each component.

2. Participants are more inclined to participate in ideation-based design contests, and, therefore, companies may analyze the level of expertise required before posting them online to estimate the number of participants.

3. The significant effects of participant attributes (e.g., the winning rate) indicate that participants are significantly motivated by recognition. The more the number of wins, the higher is the probability of participation in future contests. Therefore, contest designers should implement a certain accomplishment or reputation mechanism, such as the level-based or badge-based recommendation system, to attract more participation.

4. Contest designers should avoid allocating a high percentage of the total prize amount to the 1st prize. Instead, maintaining a certain variation among different prizes will likely improve overall participation. Also, having multiple forms of prizes, such as both monetary award and gift in a contest, will likely attract greater participation.

The results indicate that the model which accounts for dependencies (Model 2) is a better predictor of participation decisions than the model without dependencies incorporated (Model 1), as manifested in a higher prediction accuracy measured by precision and recall. The imbalanced nature of the network data poses challenges in achieving a successful rate of prediction. Despite this, the ERGM with dependencies is capable of achieving the prediction accuracy of 43%, which is significantly higher than 5.6% of random predictions.

In the future work, more experiments are required to collect datasets to validate the conclusions. Also, other types of dependent relations need to be modeled using different network configurations and tested on crowdsourcing data. For example, if the data about the social relations between participants are available, the ERGM for multidimensional bipartite networks can be used to probe into the underlying factors influencing interdependent decisions. With the estimated ERGMs in Table 2, further prediction of how potential participants make decisions given the attributes of design contests can be carried out. This is useful for contest designers in planning decisions and anticipating the possible number of solutions that can be generated from a contest. Such a prediction at aggregate level would enable the forecast of the evolution of the entire crowdsourcing system, which may benefit crowdsourcing system designers to develop new incentives to attracting more participants.

Acknowledgements

The authors gratefully acknowledge the financial support from the U.S. National Science Foundation (NSF) CMMI through grant # 1400050. An earlier version with preliminary results was presented at the ASME 2018 Computers and Information in Engineering Conference in Quebec City, Canada [63].
References


