HOW ONLINE DESIGN COMMUNITIES EVOLVE OVER TIME: THE BIRTH AND GROWTH OF OPENIDEO

Mark Fuge
Berkeley Institute of Design
Department of Mechanical Engineering
University of California
Berkeley, CA 94709
Email: mark.fuge@berkeley.edu

Alice Agogino
Berkeley Institute of Design
Department of Mechanical Engineering
University of California
Berkeley, CA 94709
Email: agogino@berkeley.edu

ABSTRACT

While companies are turning to online communities of outside designers to bring new ideas into their product development process, several questions remain unanswered: How do design communities form, evolve, and die out over time? What integrates newcomers into the community? How can one grow community without impeding idea inspiration? This paper explores these questions by analyzing how OpenIDEO, an Open Innovation design platform, has evolved from conception to present day. We find that OpenIDEO possesses a stable core who frequently collaborate with transient members, and that large single communities have evolved into smaller but denser communities over time. Moreover, OpenIDEO’s use of community managers and incentives promotes an efficient network for generating new ideas, while fostering cohesive collaboration groups. By viewing design communities as an evolving network, we can guide future design communities to become sustainable and efficient—ultimately unlocking their potential to accelerate human development.

INTRODUCTION

Companies are increasingly turning to distributed groups of volunteers to help them design their new products and services. This trend, sometimes referred to as Open Innovation or Crowd Design, takes place across industries and scales, including large corporations, design consultancies, technology licensing companies, and even government agencies. These communities leverage the creative power of thousands of volunteers, often creating novel design solutions at unprecedented speeds.

However, building and maintaining an effective design community is no short order: we need to understand how they develop and evolve over time if we wish to create mechanisms that support their growth and effectiveness. While researchers have studied collaboration networks, few have explored the growth of these recent design networks.

To that end, this paper analyzes the growth and evolution of OpenIDEO, a successful online open innovation community centered around designing products, services, and experiences that promote social impact. It addresses how the design network and its members have changed over time by using network analysis techniques on collaboration data from OpenIDEO’s inception to present day. Our analysis covers multiple scales: the design network as a whole, the community structure within that network, and how actions of individual members contribute to its overall behavior.

After providing some background on network evolution and design networks, this paper presents how we use collaborations on OpenIDEO to model its social evolution. This leads our main...
results:
1. It took around nine months for the OpenIDEO’s network properties to stabilize.
2. Single, large, centralized communities have given way to smaller, more numerous groups over time.
3. Network membership is highly transient, with many members only participating in a single challenge. There are a small core of members who provide extensive design feedback across multiple challenges.
4. Consecutive challenges see higher across-challenge retention rates than those that have a temporal gap between them.
5. A major difference between those who participate in single challenges versus multiple challenges lies in the latter group giving more comments to others.

Lastly, this paper concludes with a discussion about the implications for managing distributed design communities and presents possible directions for future research.

BACKGROUND ON DESIGN NETWORK EVOLUTION
This paper integrates two threads of research: Network Evolution and Design Networks. From the network evolution literature, we adopt techniques for modeling social networks over time, particularly that of Palla et al. [1]. From the design network literature, we note that past work has either focused on technological networks or on static models of social structure. Through this integration, this paper introduces new groundwork in describing the evolution of real-world collaborative design communities.

Network Evolution
The literature on complex networks is vast, and covered in detail by relevant textbooks by Newman and Scott [2, 3]. The portion relevant to this paper concerns how complex networks form over time, and what that behavior means for properties of interest: the case information diffusion, robustness to network attack, etc.—the reviews in Holme and Saramäki [4] and Jackson et al. [5] present good complementary introductions. This past work falls roughly into two styles: empirical studies of how real-world collaboration networks have evolved, and statistical models that attempt to capture real-world behavior. We focus on the former, given this paper’s empirical nature, though we direct interested readers who want relevant reviews of statistical network models to Csermely et al. [6] and Castellano et al. [7].

Two types of empirical networks lie closest to this paper: scientific co-authorship networks and open source software development. Though they have their differences, these two types of collaborative networks share several similarities with OpenIDEO: individual actors collaborate with each other through a formal feedback processes that results in a shared arti- fact. Barabasi et al. [8] present a representative empirical analysis of how scientific collaborations evolve over time; they note how collaborations follow a familiar power-law distribution (also present in OpenIDEO [9]) and that the clustering of the community tends to decrease over time. In Open Source Software development, Saraf et al. [10] demonstrate the social evolution of developers over stages in project lifecycles; notably, they find that assortative mixing or “homophily” [11]—where members form new links with those most similar to them—increases over time. Likewise, social network evolution (e.g., Kossinets et al. [12]) also displays a growth in homophily. In contrast, our results show that OpenIDEO’s design network has the opposite behavior—the most gregarious members tend to seek out less well-known collaborators.

Design Networks
Within design networks, this paper builds off work in two main research areas: studies that focus on technological design networks or computer simulations of collaboration, and studies that observe small design networks under experimental conditions to determine network effects of idea generation.

Prior research by Panchal and various collaborators is the closest application area to that considered by this paper. In their empirical work, they study the network structure of both Open Source Software and Hardware [13] networks—the fundamental difference with this work being that they address technological networks, while we address the social evolution of the design community. Panchal does consider social evolution in [14, 15], but bases the findings primarily on computer-simulated agents—in contrast, this work provides empirical evidence from a large real-world design community.

Many of our later implications build off of work that links the network structure of a design team with its idea generation potential. Both Mason et al. [16] and Stephen et al. [17] independently found that higher local clustering (i.e., when all your neighbors are also neighbors with each other) around a node reduces their idea generation ability, due to “complex contagion”—the tendency to copy your neighbors when they all say similar things. Essentially, they demonstrate via human experiments that if your immediate collaborators are also connected to each other (high clustering), bad ideas can fixate the entire group on a poor solution. In contrast, networks with high efficiency (low average distance between all nodes) but low clustering do not suffer from this “group think,” while still being able to spread good ideas rapidly throughout the network.

In our previous static study of OpenIDEO’s community [9], we discussed several possible mechanisms by which the network achieves desirable efficiency and clustering properties. In contrast to our previous work, this paper demonstrates how OpenIDEO’s structural properties evolved over time and provides insight into events that shaped the course of OpenIDEO’s
development. It includes a deeper behavioral analysis of the individuals within OpenIDEO and introduces a time-dynamic model based on the work of Palla et al. [1]. In relation to other design networks, OpenIDEO differs from previous studies of collaborative design networks in the following: it focuses on applying human-centered design to both product and service design; the users come from a variety of design backgrounds, including industrial design, engineering, business, and arts; and the challenges focus on large-scale social problems, rather than specific technical challenges seen in engineering-centric competitions (i.e., DARPA’s FANG challenges). We direct readers to the review by Lakhani et al. [18] for more information about OpenIDEO’s governance, such as how it runs the challenges, the type of design challenges it hosts, how it operates within IDEO, and the type of participants using the platform.

STUDYING THE EVOLUTION OF OPENIDEO

Analyzing OpenIDEO’s evolution from inception to present day required several non-trivial methodological choices: 1) what aspects of the OpenIDEO network constitute the nodes and edges in the network; 2) how does one model time-dependent quantities, like link strength, using only discrete time events; and 3) how did we determine community definitions within the network.

Defining OpenIDEO’s Collaboration Network

To translate OpenIDEO’s community into a mathematical network, we use a similar data representation to our prior work [9]—each individual represents a node in a graph, and we add a directed edge from user A to user B when either of the following events occur: user A comments on user B’s submitted concept, or user A replies to user B’s comment on any site content. The content of these comments can be positive, negative, or neutral in tone, and do not have to relate strictly to submitted designs, since our goal here is to model the social transactions between individuals (though a more qualitative study of the comment content would be an interesting extension of this work). These connections create a weighted, directed graph whose general structure looks approximately like those in Fig. 3—a central core of users with many connections surrounded by a periphery of users with few connections.

Adjusting for Temporal Data

A major difference between our prior work [9] and this paper lies in how we adjust the specific edge weights over time to account for temporal events. To handle temporal events in the OpenIDEO network, we use the approach of Palla et al. [1], who model each edge weight as having a decay factor:

\[ w_{A,B}(t) = \sum_i w_i \exp(-\lambda |t - t_i|/w_i) \]  

(1)

where \( i \) indexes an event between user A and user B. Essentially, this treats the edge weight as a sum of exponentially decaying functions—if the users interact regularly, they will have high edge weight, but if they go months without interacting, their edge weight approaches zero. We treat all events as equally important by setting \( w_i = 1 \), and set the decay rate \( \lambda \) such that \( w_i \) is 1% of its original strength after 100 days of inactivity—we provide sensitivity analyses and experiment code on our companion website\(^5\) that demonstrate the robustness of this particular choice.

With this added temporal model, we can now represent the OpenIDEO network at any given timepoint by summing up all prior user events, and using Eqn. 1 to appropriately decay the edge weights. To save computation and induce sparsity on the graph structure, all edges below a certain cutoff threshold are discarded—we describe how we set this threshold in the next section.

Community Definitions

Before we can present our results, we need to discuss one last important methodological element: how do we define or detect communities within the network? In this paper, we construct design communities on social terms: communities are groups of designers that communicate among each other, as defined through social interactions such as commenting. Community detection on graphs is still an active area of research, and for this paper we adopt the Clique Percolation Method [19]. It is a widely-used community detection method that can identify an unspecified number of communities where \( k \) specifies how interconnected the community should be—higher \( k \) values mean smaller, more densely connected communities, while lower \( k \) values would create larger more loosely connected communities.

To set \( k \) and the edge weight cutoff, we take the approach described in Palla et al. [1] of finding the highest \( k \) such that large communities are still able to form, and then reducing the cutoff threshold until it is barely above the value needed to preserve community structure. We conducted sensitivity analyses around both \( k \) and the cutoff value, but since the overall results did not change we omitted those results for readability—interested readers can view these supplemental sensitivity analyses and download our full experiment code on our companion website.\(^5\) In physical terms, this cutoff procedure defines a minimum communication frequency that two individuals must overcome to still be considered “connected” to each other. Much like in other social communities, if two individuals stop collaborating the strength of their connection decays over time and an individual would leave a former community once his or her existing ties to that community atrophy. This minimum level of atrophy and the tightness of the resulting community are what the cutoff and \( k \) values represent, respectively.

\(^5\)http://www.markfuge.com/openideo
RESULTS

With a temporal version of the OpenIDEO collaboration network, we now present our results, starting at the whole-network level, moving down to community level, and finally to the level of individual users. The results highlight when the network-level disassortative mixing, efficiency, and clustering behavior observed in our prior work [9] started to occur; how community structure develops over time; how users enter and participate in the system; and what particular actions users take. We used the NetworkX library [20] to store and process our graph data, and we are making our experiment code available for those who wish to replicate our work.5

How the Entire Network Evolves

Motivated by the work of Mason et al. [16] and Stephen et al. [17], Figure 1 plots the assortativity (1a), efficiency (1b), and clustering (1c) of the OpenIDEO network over time. We find that all three properties varied over the entire lifetime of the network, with the biggest variations occurring during the first nine months of OpenIDEO’s growth. The network starts and remains disassortative over its lifetime, which means that frequent users collaborate more often with infrequent users than with each other.

How the community structure changes over time

To visualize how OpenIDEO’s community structure has changed over time, we first summarize the size and number of communities in Fig. 2, and then provide visual snapshots of the community structures at various timepoints in Fig. 3.

Figure 2 plots a point for each community on the x-axis with its corresponding size on the y-axis. For example, early in OpenIDEO’s development, only a few small communities existed (Fig. 3a: Dec. 10th, 2010). However, around the start of Challenge 3, the user base grew, expanding the communities (Fig. 3b: Mar. 23rd, 2011) until they eventually merged into a single, core community (Fig. 3c: Feb. 7th, 2012). After a year, the size of the largest community began to decrease and split into several smaller communities who share some common nodes (the red nodes in Fig. 3). As Fig. 3g demonstrates, the decrease in community size is not due to lack of membership or participation—the core is still actively participating, but has started developing smaller communities within the core, as evidenced by Fig. 3h (Jul. 2nd, 2013).

The Lifetime of Community Members

Having looked at the network and community levels, our frame of reference now focuses on the individual users: How long do they stay with the site? How do they spend their time when participating? What makes long-term, multi-challenge participants different than single-challenge participants?

To answer the first question about user lifetimes in OpenIDEO, we record the difference between the date a user joins the site and their last date of activity on the site (i.e., when they last submitted a concept or left a comment). Aggregating the data for all 5753 users, Fig. 4a shows a log-scaled histogram of the number of days between joining and last activity; it demonstrates the long-tail of participation, with only a small number of users remaining through several challenges. Figure 4a does not account for the fact that certain users joined later than other users, biasing the histogram towards lower participation times. To address this, Fig. 4b divides the number of days a user has been active by the total number they could have been active, based on their join date. The story remains the same: the vast majority of participants are transitory visitors, with a central core of committed members.

To further explore how users participate in the design community, Fig. 5 shows the three different possible user actions (joining, submitting a concept, or giving a comment to someone) for each user as a function of time. (We remove the 1794 users who joined but did not participate further, leaving a total of 3959 users). The y-axis represents a particular user id, where the users have been sorted by the date they joined OpenIDEO. It demonstrates not only the user growth pattern over time, but also the general behavior of most users: after joining, the users participate in a frenzy of activity that includes both concept submissions and feedback. Once the challenge ends, user participation drops dramatically, with reduced participation in subsequent challenges. Challenges that begin on or before a previous challenge ends correlate with higher user retention in the subsequent challenge;
(a) The network remains disassortative (negative assortativity) over its lifetime, unlike most other social networks.
(b) Efficiency fluctuates over time.
(c) Average clustering fluctuates over time, with a minor decrease as the network grows older.

FIGURE 1: It took about about 9 months to a year for the system properties to equilibrate.

FIGURE 3: A series of community snapshots of the OpenIDEO collaboration network over time. The colored polygons and nodes represent different communities, with the bright red nodes representing nodes that straddle different communities. Grey nodes are not part of any community.

Cases where there is a significant gap between challenges (e.g., mid-May to July, 2011) did not see much return participation.

Figure 6 normalizes Fig. 5, by shifting everyone by their join date, making the x-axis equivalent to the number of days a user has been on the site. This figure presents a clearer picture of the exact user progression, showing the concentration of activity around the initial challenge followed by a sharp drop-off in participation for most users. Since the y-axis is ordered in time, we can also discern that users have behaved similarly since around June, 2011 to present—there is little difference in activity level or distribution of activities after around user 750 and later.

Figure 6 shows an identical plot to that of Fig. 5, except that all the users have been sorted on the y-axis by the total number of actions they have performed on the site (essentially ranking them by activity level). This provides a sense of relative size and activity level: there are many more single-challenge participants compared to multi-challenge participants, and only a small fraction of the participants heavily contribute. This long-tailed
FIGURE 5: This figure captures user actions over time, including when they joined, when they submitted concepts, and when they commented on the concepts of others. Two things are evident: 1) Unsurprisingly, most activity takes place during challenges: user joins, submissions, and commenting activity all increase during challenges; and 2) user retention and participation across multiple challenges was higher when consecutive or simultaneous challenges were available.

The differences in user activities become more pronounced when we group users by those who have participated in only one challenge (2897 users = 73% of active users) and those who have participated in multiple challenges (1062 users = 27% of active users). Figure 7 records what happens next to the users after they perform a particular activity—essentially it is a “transition matrix” between activities. Comparing Fig. 7a and 7b, we find that those participants who participated in multiple challenges spent more time giving comments to others than they did submitting concepts or getting comments from others. In both cases, commenting formed a reinforcing cycle of giving and getting comments.

DESIGN IMPLICATIONS

Our main implications for managers of online design communities center around the following:

1. Initial design communities need time to settle, and managers should wait for the network properties to stabilize before evaluating the effectiveness of a design network.
2. Since community structure is dynamic and changes over time, managers should be aware of mechanisms for maintaining community structure, such as the usage of community managers and collaboration incentive systems.
3. Having a central core of users reach out to transient members is a good way to bring in fresh ideas and improve idea generation through higher information efficiency and lower clustering.
4. Spacing challenges so there is continuous involvement, while encouraging commenting through incentives, might increase retention among users.

In Figures 1, 2, and 3 we find that the design network changes its composition and structure throughout its lifetime, and that it can take time before it settles into a predictable range. In OpenIDEO’s case, it took around nine months and three design challenges before the network maintained any kind of regularity. We see this as an argument for allowing design communities the opportunity to “break themselves in” before trying to evaluate their particular quality.
FIGURE 6: This figure captures user actions over days since joining the community, when they submitted concepts, and when they commented on the concepts of others. Users are sorted by their total number of actions on the site, where the most frequent users are towards the bottom. Several things are evident: 1) Only a small portion of users participate regularly across many challenges—most users only temporarily participate. 2) Users initially start out with a flurry of activity that tempers in later challenges. 3) Among more frequent members, giving comments is more popular than submitting new concepts.

The transient nature of the user population is both a blessing and curse. On the one hand, having new members constantly joining increases idea diversity (and possibly novelty), but on the other hand having users leave after one challenge does not foster a strong sense of community feedback or knowledge retention across challenges. One can employ two complementary strategies here: increase user retention and make better use of the transient population. For the first, we recommend spacing challenges so that they are consecutive—the continuity of involvement appeared to correlate with participation in the subsequent challenges. We note that this link is not necessarily causal, though we do believe that social momentum and continuous engagement in the site will positively affect retention. For the second, OpenIDEO’s two strategies of using community managers to “cross-pollinate” [9] between new users’ ideas, along with providing incentives for giving comments, facilitates the disassortative, efficient, and lightly clustered network structure that Mason et al. [16] and Stephen et al. [17] recommend for product ideation. Further incentives for building off of existing ideas, rather than just commenting, would increase this benefit.

For establishing community, Fig. 2 and 3 both demonstrate the dynamic nature of how design communities evolve. Research has not yet investigated whether having a single larger community or several smaller connected communities provides a more conducive idea generation environment, so it remains to be seen what specific level of community a design network should strive towards. Despite the changes in community structure, the general structure of OpenIDEO remains similar over time: a central core of users collaborates heavily with temporary periphery members. We believe that this overarching structure is the more likely cause of the beneficial disassortativity, efficiency, and clustering seen in Fig. 1 and Fuge et al. [9].

We find the differences between Figs. 7a and 7b interesting, though not particularly surprising—it is easier to give someone a comment than to submit a new concept, so partaking in multiple challenges through commenting does not have a high barrier to entry. Regardless, our results do show a desirable “give a comment–get a comment” loop which OpenIDEO encourages through providing “collaboration points” to members for commenting.
A large number of users are transient—their activity on the site does not extend past a few days. 1794 users out of 5753 users (≈ 31%) join, but do not interact with the site further.

(b) Normalizing the lifetimes by the amount of time since the user joined, we see that user lifetime steadily decrease, as expected.

**FIGURE 4:** User lifetimes show a highly transient user population with a common core of long-term participants.

**CONCLUSIONS AND FUTURE WORK**

This paper presents an empirical network analysis of Open IDEO’s social evolution from concept to present day. The analysis spans both time and scale by considering the network structure as a whole, as well as communities within the larger network and the specific actions of users.

We find that OpenIDEO’s structure took several design challenges to stabilize, with its community structure becoming a mildly clustered core-periphery network. Members that participated in multiple challenges were more likely to give design feedback to others, with consecutive challenges engendering higher retention. We derive several implications for managing design communities: use core members to reach out to transient participants, space challenges sequentially or with overlap to promote continuous involvement, and use incentive structures to encourage giving feedback within communities.

(a) Users who only participated in a single challenge submitted around 1-1.5 concepts on average, and gave about 4 comments. They received around twice as many comments as they gave to others.

(b) Users who participated in multiple challenges submitted between 5-6 concepts on average over their lifetime, but gave a substantially higher percentage of comments to others.

**FIGURE 7:** A comparison of the transition states between single- and multi-challenge users. The numbers in the boxes represent the average number of times a user went from one activity to the next activity. Those users who participate in multiple challenges put more emphasis on giving comments.
For researchers, this paper provides a benchmark with which to compare the growth of other design communities and highlights the role of community growth. This provides several directions for further experimental or qualitative study: What are the range of factors that convince users to regularly participate? How could interventions, such as targeted collaboration reminders, alter the networks evolution over time to promote better ideation? What causes an individual to continue participating in the network when a challenge ends? What are appropriate computational methods for modeling this social interaction (e.g., Markov Reward Networks, as in Fig. 7)? Answering many of these questions requires a more controlled environment that our observational dataset allows, and would be a fruitful area of future research. In particular, we intend to use the research described herein to refine the design and structure of the DesignExchange as an interactive portal of design methods for the human-centered design community [21, 22]. By understanding how these design communities grow, evolve, and (eventually) die out, managers of online design communities can create environments that better support distributed idea generation, ultimately allowing these communities to guide the future of product development.

ACKNOWLEDGEMENTS

We would like to thank OpenIDEO whose work formed the basis of our dataset, as well as the reviewers for their helpful comments on our manuscript. We also thank Nathan Maton at IDEO for helpful discussions around the evolution of the community, and Kevin Tee for portions of the network analysis code. This work was supported by NSF CMMI-1334361 and the Department of Defense (DoD) through the National Defense Science and Engineering Graduate Fellowship (NDSEG) Program.

REFERENCES