Novelty and Diversity: Remixing with Human-based Search Algorithms

Short Paper

Yue Han
Le Moyne College
1419 Salt Springs Rd, Syracuse, NY
hany@lemoyne.edu

Jeffrey V. Nickerson
Stevens Institute of Technology
1 Castle Point Terrace, Hoboken, NJ
jnickerson@stevens.edu

Abstract

Remixing is used as a method to harness collective intelligence in many online innovation communities. In these remixing communities, users can reuse previous work to produce innovative outcomes. However, some users tend to make remix decisions based on what is immediately visible to them. Because of this they miss opportunities to further explore the design space. Can crowd members be guided to conduct collective exploration in a systematic way to better cover the design space? In this paper an exploratory study compares the collective exploration of ideas generated by four human-based remixing algorithms: deepening, widening, depth-first, and breadth-first. The distances between ideas that are generated by these remixing algorithms are calculated and compared. The results suggest that the deepening and hybrid algorithms can encourage users to better cover the design space. Potential ways to further improve the algorithms are discussed, with the goal of encouraging collective exploration through remixing.

Keywords: Remixing, digital innovation, innovation search, human computation, crowdsourcing, collective intelligence

Introduction

Remixing has been used to harness collective intelligence in many online innovation communities such as ccMixter, Scratch, and Thingiverse. The term remixing comes from music composition; it is used to describe the reinterpretation of already created song components. Today, the term has taken on a broader definition: an innovation process that generates new products by modifying or recombining existing products (Resnick et al. 2009; Stanko 2016). Remixing can be operationalized on a website as a built-in function that explicitly links new ideas (children) to source ideas (parents). The function can be extended to show a network of ancestors and descendants of an idea. Such remix networks illustrate how users collectively explore a design space, the space of all possible solutions (Nickerson and Yu 2012). More broadly, these networks provide valuable insights into innovation processes (Kyriakou et al. 2017).

Ideally, as users build upon each other's work, they should be able to collectively cover the design space and accelerate the discovery of novel ideas. Yet, in many online remixing communities such as Scratch, there is a trade-off between generativity and originality (Hill and Monroy-Hernández 2013). When users self-select what to remix, they make remix decisions based on what is immediately visible to them. What is visible is a small subgraph of the overall network, often just the more popular ideas. Because of this they miss the chance to remix relevant previous work that could further the exploration of the design space; instead, they end up generating similar ideas within a small area of the design space. As with many other online communities, online remixing communities have "temporariness" (Majchrzak and Malhotra 2016). The time and effort most users could spend on a task is limited. Therefore, finding interventions that could improve the efficiency of crowd work becomes important (Siangliulue et al. 2016), especially for organizations that are seeking to innovate through online communities. Given the same amount of crowd resource, is there a way to optimize the collective exploration? In other words, can we guide crowd members to conduct collective exploration in a systematic way that increases coverage of the design space? Increased coverage leads to an increased diversity of solutions, which increases the chances of useful discoveries.

Prior studies have examined the remix networks that is naturally formed by users in online remixing communities and discussed different remixing patterns observed in these communities (Nickerson 2015; Flath et al. 2017). While providing insights into the formation of online collaboration and innovation, these studies also hint at a potential method to aid collective exploration: instead of relying on the users' independent choices, the evolution of the remixing network can be guided in a way that optimizes collective goals such as design space coverage. That is, human-based remixing algorithms can instruct crowd members to start from positions in the design space that are more likely to increase coverage.

Firms search for innovation in two ways, making use of search depth and search scope (Katila and Ahuja 2002). Similarly, we suggest that participants in online remixing communities also conduct their searching efforts in two ways. We characterize these as remixing depth, an effort to create more generations of new ideas, and remixing breadth, an effort to create more children in the same generation. These two independent methods lead to the formation of two poles of remixing algorithms for collective exploration: a widening remixing algorithm and a deepening remixing algorithm. These two pure remixing algorithms can be combined to form hybrid remixing algorithms. We propose two hybrid remixing algorithms based on two search algorithms in computer science: breadth-first search and depth-first search (Korf 1985). How would these human-based remixing algorithms shape and affect collective exploration for innovation?

To answer this research question and examine the effect of these four remixing algorithms on collective exploration, we conducted an exploratory study that compares the distances of ideas generated by the four human-based remixing algorithms. We use distances to calculate two measures: seed-generated novelty, the dissimilarity between the seed design and the remixed design, and diversity, the overall dispersion of the remixed ideas. In this paper, we first discuss related work and introduce the four human-based remixing algorithms. Then we present the exploratory study, an experiment that compares ideas generated from panel members of Amazon Mechanical Turk (MTurkers). Following that, we discuss the implications of our study and suggest potential follow-on studies.

Theoretical Development

Remixing, a Collective Innovation Search Method

Innovation is often perceived as a search process (Dahlander et al. 2016; March 1991; Petruzzelli and Savino 2014). Today, many organizations are gradually shifting their innovation model toward a paradigm—open innovation—that suggests organizations should use and connect internal ideas and external ideas (Chesbrough 2006). Based on this paradigm, organizations have conducted external innovation search with people from outside the organization, including competitors (Lim et al. 2010), suppliers (Schiele 2010), universities (Harryson et al. 2008), and customers (Gassmann et al. 2006). In these studies, customers are often identified as the main source of external innovation; they are also recognized as the key source of innovation in crowdsourcing (Bogers and West 2012; Poetz and Schreier 2012).

The development of Internet technologies has facilitated innovation search in online communities (Afuah and Tucci 2012; Füller et al. 2008). Crowdsourcing, also known as broadcast search, has been widely used to conduct external search for innovation (Lakhani et al. 2007). This search method outsources the innovation task to a large group of people in the form of an open call (Howe 2006; Jeppesen and Lakhani 2010). Initially, the implementation of this search method is intended to generate individual ideas from crowd members in an online community (e.g., InnoCentive, MyStarbucksIdea, Threadless). Later, innovation communities like Thingiverse started to include a modification process as this leads to more creative results (Kyriakou et al. 2017). This innovation search method with a modification process is called remixing, which allows crowd members to build on each other's ideas: for example, in the Climate CoLab (Malone et al. 2017) and Scratch communities (Han and Nickerson 2015; Resnick et al. 2009).

A major strength of crowdsourcing is that it allows organizations to obtain access to a large number of diverse ideas (Chiu et al. 2014; Majchrzak and Malhotra 2013). In addition to diversity, crowdsourcing brings organizations the benefits of distant search at a low cost (Afuah and Tucci 2012). However, studies also suggest that sometimes crowdsourcing is not efficient because a great many of the ideas generated are superficial or redundant (Bjelland and Wood 2008). Unlike the initial implementation of crowdsourcing that asks crowd members to work individually on ideas, remixing allows them to build on existing work, which has the potential to deepen ideas which in turn can lead to innovation at the collective level (Wisdom and Goldstone 2011). Is it possible to optimize this search process to improve efficiency while promoting the novelty and diversity of the ideas generated?

Visualizing Collective Exploration by Remix Networks

When users remix ideas in online communities, they are involved in a collective exploration activity that searches for the best position in a space of possibilities (Nickerson and Yu 2012). When we trace the remixing path among ideas, we can form remix networks that allow us to visualize the collective exploration. In these remix networks, nodes are the original ideas and its remixes and edges represent the remixing path. To understand how users generate ideas and collectively explore a design space through remixing, researchers have examined different online communities and summarized remixing patterns from these remix networks (Nickerson 2015; Flath et al. 2017). For example, remixing patterns can be classified based on inheritance characteristics, that is, whether a remix is modified based on one idea or recombined based on multiple ideas (Nickerson 2015). Data from an online 3D remixing community exhibits 8 distinct remixing patterns (Flath et al. 2017). However, previous studies have not addressed how different remixing patterns affect the ideas generated. We aim to examine how different remixing algorithms affect the generation of new artifacts.

In this paper, we experimentally drive the creation of remix networks and then analyze how crowd members collectively search for innovation in different conditions. In addition, we leverage computational methods to create an exploration map: We first measure the distances among all ideas generated in a remix network, and then converge the high dimensional distance matrix into a two-dimensional space using multidimensional scaling (Kruskal and Wish 1978). The exploration map not only shows the topological evolution of a remix network but, because it shows distances, also reveals the novelty and diversity of ideas relative to the seed ideas.

Human-based Remixing Algorithms

A classic understanding of innovation search is that the search space can be divided into local and distant (Fleming 2001; March 1991): search can occur in the territory near existing, well understood ideas, or far away from the comfort zone of a company. Innovation search can be characterized along two dimensions. Prior study calls these two dimensions search depth, "how deeply a firm reuses its existing knowledge", and search scope, "how widely a firm explores new knowledge" (Katila and Ahuja 2002). Alternatively, some researchers call the two dimensions search depth, "the extent to which firms draw intensively from different search channels or sources of innovative ideas", and search breadth, "the number of external sources or search channels" (Laursen and Salter 2006). Following a similar train of thought, we suggest that remixing, as a form of collective innovation search, also can be characterized by two distinct dimensions: remixing depth and remixing breadth. We define remixing depth as the effort to create more generations of new ideas and remixing breadth as the effort to create more children in the same generation. The former aims to conduct in-depth search, drilling sequentially down one path, while the latter aims to generate variations by starting many branching paths. Based on these two search methods, we propose two poles of exploration patterns we call human-based remixing algorithms: a deepening remixing algorithm that focuses on the remixing depth, and a widening remixing algorithm that focuses on the remixing breadth (Figure 1).

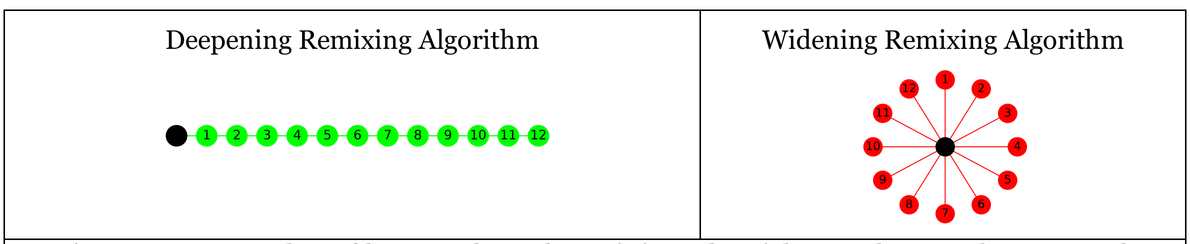


Figure 1. Two poles of human-based remixing algorithms. The number on each node refers to the sequence of creation.

As shown in Figure 1, in the deepening remixing algorithm, participants modify a design idea previously created in the remix network and form a chain of remixing (green nodes) that originates from the seed idea (black node). In this algorithm, the goal is to motivate crowd members to explore farther away from the seed idea as participants always have access to the latest design in the remix network. In this case, compared to the seed idea, the novelty of the ideas generated may increase. In contrast, in the widening remixing algorithm, all remixes (red nodes) are created based on the seed idea (black node). This algorithm encourages participants to explore in diverse directions in the design space and generate variations based on the seed idea. This thinking leads to the following two hypotheses:

H1: Ideas generated in the deepening remixing algorithm are more novel with respect to the seed idea than ideas generated in the widening remixing algorithm.

H2: Ideas generated in the widening remixing algorithm are more diverse than ideas generated in the deepening remixing algorithm.

In most online remixing communities, a user's movement in the exploration map appears random. For example, researchers have found that in the Scratch community, instead of moving forward to expand the solution space, sometimes participants move back towards the seed idea (Han and Nickerson 2015). This type of remix can be problematic because it (1) increases the possibility of redundancy and (2) may bring the exploration back to the starting seed, which slows collective exploration. Although we cannot fully avoid this type of remix since we cannot control individual behavior, it is possible to minimize the influence of such remixes and thereby improve the efficiency of the collective exploration. We propose two hybrid remixing algorithms originating from two simple search algorithms in computer science—depth-first search and breadth-first search (cf. Korf 1985)—to address this problem.

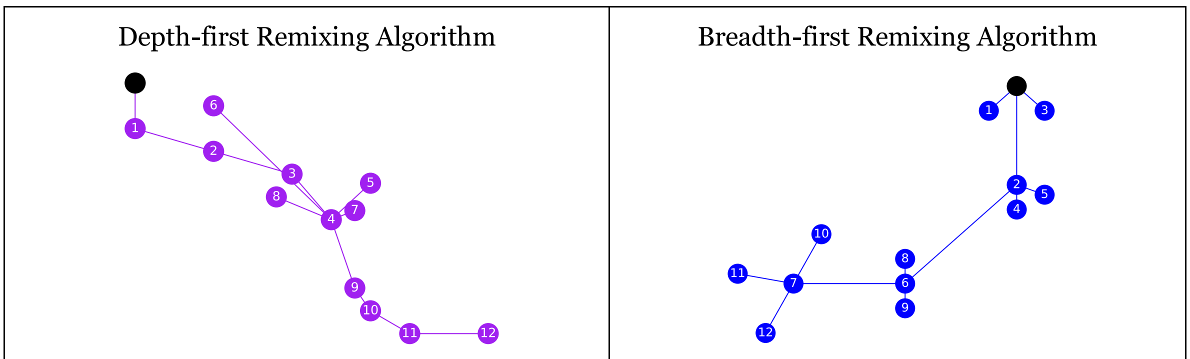


Figure 2. Two Hybrid Remixing Algorithms

As shown in Figure 2, the first hybrid remixing algorithm is created based on depth-first search. We call this remixing algorithm the depth-first remixing algorithm. Similar to depth-first search, this remixing algorithm prioritizes the search depth: it starts at the seed design and searches as far as possible on each remixing branch before backtracking. In this algorithm, we use the distance from the seed idea to determine the order in which remixing branches should be explored. In other words, each participant is presented with and asked to modify the design farthest from the seed idea at the time they start to remix. In the provided example, idea 1 was a remix generated based on the seed idea (black node). Idea 2, 3, 4, and 5 were generated based on the previous idea one after another. However, at the time before idea 6 was generated, the farthest idea from the seed was idea 4 rather than idea 5. Therefore, idea 4 was chosen as the parent idea for idea 6. Similarly, idea 7, 8, and 9 were all generated based on idea 4. After that, idea 10, 11, and 12 were generated one by one based on the previous idea.

The second hybrid remixing algorithm is called the breadth-first remixing algorithm and follows the process of breadth-first search. This remixing algorithm prioritizes the search scope: it starts at a seed design and searches the design space by generations. At each generation, three remixes are generated based on the same parent idea. Similar to the depth-first remixing algorithm, we also use the distance from the seed idea to determine which idea in a generation should be further remixed. In the provided example, in the first generation, idea 1, 2, and 3 were generated based on the seed idea. Since idea 2 was the farthest from the seed idea, it was selected as the parent idea for the next generation; and idea 4, 5, and 6 were generated based on idea 2. Following the same fashion, idea 7, 8, and 9 were generated based on idea 6, and idea 10, 11, and 12 were generated based on idea 7. Note that when the original computer science search algorithms were conceived, the assumption was that the data to be found already existed: here, the remixing algorithms are used to ask people to generate new data; these data in turn guide the next steps.

Since both hybrid remixing algorithms are essentially combinations of the deepening remixing algorithm and the widening remixing algorithm, they may have advantages with respect to both novelty and diversity. Therefore, we propose the following hypotheses:

H3: Ideas generated in hybrid remixing algorithms are more novel with respect to the seed idea than ideas generated in the widening remixing algorithm.

H4: Ideas generated in hybrid remixing algorithms are more diverse than ideas generated in the deepening remixing algorithm.

Exploratory Study: An Amazon Mechanical Turk Experiment

Experiment Design

In this exploratory study, we conducted an experiment to examine how the four human-based remixing algorithms aid collective exploration. In particular, we want to test our hypotheses and examine the seed-generated novelty and diversity of the ideas generated in each algorithm. We asked random crowd members on Amazon Mechanical Turk (Mason and Suri 2012) to generate ideas for new YouTube channels. We first generated one idea and established it as the seed idea for all four conditions. Then we randomly assigned MTurkers to create remixes in one of the four remixing algorithms. In the deepening condition, the remixes are generated one after another; each MTurker in this condition remixes the previous remix. In the widening condition, all MTurkers remix the seed idea. In the depth-first condition, the remixes are generated one after another. For each MTurker, we first compute the distances between the seed idea and all remixes that have been generated in this condition. The distance is calculated based on the frequency-inverse document frequency (Tf-idf) matrix (Pedregosa et al. 2011) using the text of the ideas. Then we ask the MTurker to remix the idea that is the most distant from the seed idea. In the breadth-first condition, the remixes are sequenced by generation. In each generation, we ask three MTurkers to remix the idea in the previous generation that is most distant from the seed idea. In this experiment, we controlled for the crowd effort of each condition: We recruited 121 participants: 1 for the seed idea and 30 in each condition. All participants were U.S. residents with an approval rate higher than 95%. Each participant was only allowed to create one idea.

To visualize how each remixing algorithm shapes the collective exploration, we used a previously studied technique and generated an exploration map for each condition according to the ideas' text distance (Han and Nickerson 2015; Siangliulue et al. 2016). For each condition, we first calculated the Euclidean distance between any pair of ideas based on the term frequency-inverse document frequency (Tf-idf) matrix. Then we used multidimensional scaling (MDS) to reduce the dimensions of the distance matrix and mapped all ideas to a two-dimensional space.

Exploration Maps

Figure 3 shows the exploration map of each condition. In each map, the seed idea is colored in black and remixes are colored in multicolor. We also qualitatively analyzed all ideas and found that overall each exploration map represents the semantic distances among different ideas quite well. We can see from the exploration maps that different remixing algorithms do influence the participants' exploration behavior. We observed that ideas in the deepening remixing algorithm first gradually move away from the seed idea, and at one point some ideas jump back towards the seed idea. Then the following ideas start to move away from the seed idea again. This pattern corresponds to observations made in Scratch that we mentioned in the previous section (Han and Nickerson 2015). Ideas in the widening remixing algorithm spread around the seed idea in a wheel shape. Compared to the deepening remixing algorithm, remixes generated in the widening remixing algorithm are more compact. The depth-first exploration map is quite interesting. The idea of providing the remix farthest from the seed helps the collective exploration: Although problematic remixes cannot be avoided, participants are pushed away from the seed idea by the more adventurous ideas. Similar to the depth-first exploration algorithm, idea selection also plays an important role in the breadth-first exploration algorithm: participants move away from the seed idea and explore the design space in clusters.

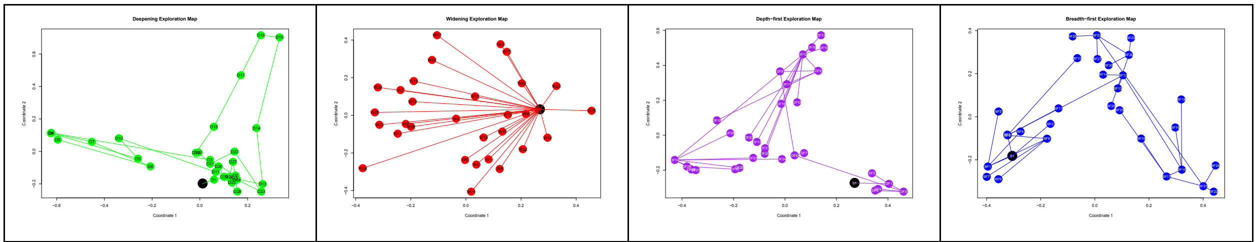
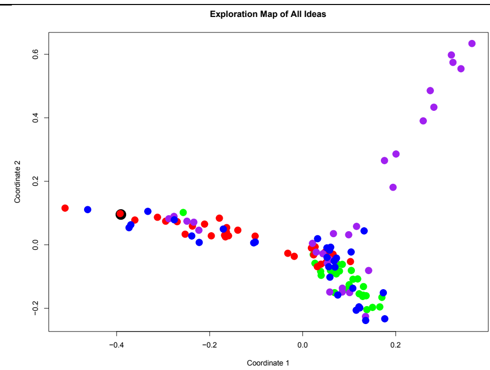


Figure 3. Exploration maps.
Deepening: Green; Widening: Red; Depth-first: Purple; Breath-first: Blue.

Analysis



Seed idea: Black Node.

Deepening: Green Nodes.

Widening: Red Nodes.

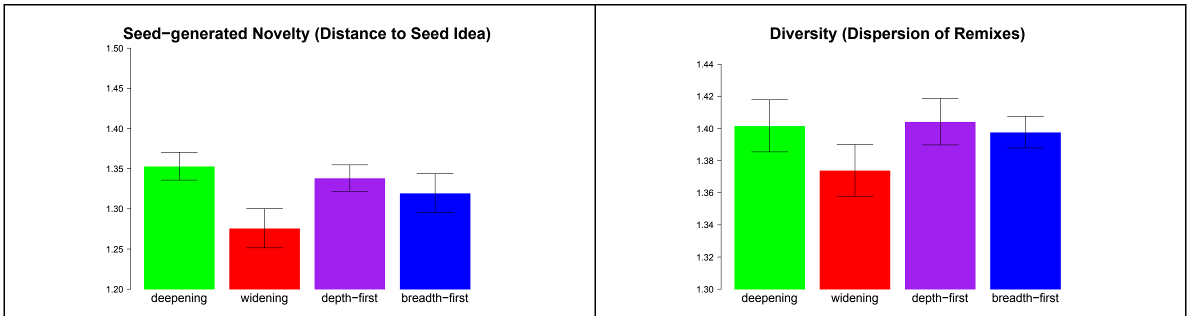
Depth-first: Purple Nodes.

Breath-first: Blue Nodes.

Figure 4. Exploration Map of All Ideas

To compare the remixing algorithms, we generated an exploration map containing all ideas generated in the four conditions. In this map, since the number of nodes is large, we eliminated the remixing path for better visualization. As seen in Figure 4, the participants in the depth-first algorithm (purple nodes) collectively explore the farthest while participants in the widening remixing algorithms mostly stay close to the seed idea. To test our hypotheses, we used distances to calculate two measures for each remixing algorithm: seed-generated novelty, the average distance between the remixes and the seed idea, and diversity, the overall dispersion of all the remixes. We measured these two variables based on the same distance matrix used to generate the above exploration map. Each remix's seed-generated novelty is calculated as the distance between the seed idea and the remix. We measured diversity as the multivariate dispersion of all remixes generated in each remixing algorithm (Anderson et al. 2006). Then we followed Anderson's (2006) PERMDISP2 procedure and conducted a permutation test for group diversity comparison (Oksanen et al. 2007).

Results



Figures 5a and 5b. Seed-generated novelty and diversity of ideas from different remixing algorithms. Error bars represent 95% confidence intervals.

As shown in Figure 5a, our t-test shows that ideas in the deepening condition are significantly more novel with respect to the seed than ideas in the widening condition (p-value<0.001); ideas in the depth-first condition (p-value<0.001) and breadth-first condition (p-value<0.05) are also significantly more novel with respect to the seed than ideas in the widening condition. Therefore, H1 and H3 are both supported, which indicates that the deepening remixing algorithm forms an in-depth collective exploration and generates more novel results. In hybrid remixing algorithms, since crowd members can build on each other's work, the ideas they generate are also more novel than ideas in the widening group.

Table 1 shows the permutation test result and Figure 5b shows the multidimensional dispersion in each condition: each remix's distance to the centroid of all remixes in the multidimensional distance space. The results suggest that ideas in the widening condition are significantly less diverse than ideas in the deepening condition (p-value<0.05); ideas in the depth-first condition (p-value>0.05) and breadth-first condition (p-value>0.05) are not significantly more diverse than ideas in the deepening condition. Therefore, H2 and H4 are not supported, which means the widening remixing algorithm does not have an advantage with respect to diversity.

Table 1. Permutation Test Pairwise Comparison

	Deepening	Widening	Depth-first	Breadth-first
Deepening		**0.019***	**0.820**	**0.669**
Widening	**0.016***		0.008**	0.010*
Depth-first	**0.807**	0.006**		0.438
Breadth-first	**0.672**	0.010*	0.446	

Observed p-value below diagonal, permuted p-value above diagonal. ***p<0.001; **p<0.01; *p<0.05

Discussion

In this exploratory study, we find that human-based remixing algorithms can aid collective exploration by guiding crowd members to efficiently explore the design space. The effect of deepening and hybrid algorithms are stronger because users can generate ideas that are both far away from the seed idea and other remixes. Our study has both theoretical and practical contributions. Exploration maps show the evolution ideas, which deepens our understanding of collective creativity. The preliminary results in our study also suggest that using remixing algorithms as interventions to improve the efficiency of crowd innovation search is promising. Organizations may consider redesigning the structure of online remixing community by controlling the visibility of ideas in order to guide crowd members and thereby shape the collective exploration.

Our exploratory study also suggests future research. We note is that there is fixation in collective exploration (Purcell and Gero 1996; Kohn and Smith 2011) regardless of the remixing algorithm. We conducted a preliminary analysis and examined each idea generated in the exploration map. We found that *music, human*, and *animal* are the three main topics of the ideas generated in our experiment. The topic of *human* seems to have a strong gravitational field. For example, in the depth-first remixing algorithm, the collective exploration starts from the topic music (e.g. mixing music) and gradually moves toward the topic human (e.g. show your talent). Then one remix jumps to the topic animal (e.g., llama) and the remixes after this jump fall into two camps: some carry on the exploration related to the animal topic (e.g., interviewing different animals), while others move back to the human topic (e.g. interviewing kids). Modifications of these remixing algorithms might be designed to address this issue in the future. For example, a mechanism could be added from classic taboo search (Cvijovic and Klinowsk 1995) in a way similar to the way Von Ahn and Dabbish (2004) use taboo search in a crowd setting. That is, the main clusters could be identified, and a taboo list generated representing those clusters to help participants avoid ideas that have already been explored so that the crowd can be steered away from the attractors.

Another further research direction is related to parallelism. Rooted in classic search algorithms, these human-based remixing algorithms also have time characteristics (Korf 1985): deepening and hybrid algorithms are more time-consuming since remixes are sequentially generated one after another, extending the duration of the process, in contrast to processes in which exploration can proceed in parallel. There is a trade-off between duration and performance in remixing algorithm selection. In particular, parallel methods do not allow as fine-grained a control of diversity. Future studies might explore algorithms that actively seek to measure and control density in particular parts of the design space, in order to discover more effective ways to generate diversity.

Concluding Thoughts

One of the main challenges in collective intelligence and crowdsourcing innovation is generating desired outcomes efficiently. By combining human intelligence with human-based computer algorithms, we may be able to create more effective systems that guide crowd members to explore the design space in a way that produces both novel and diverse ideas, catalyzing collective creativity.

Acknowledgements

This material is based upon work supported by the National Science Foundation under grants IIS-1422066, CCF-1442840, IIS-1717473, and IIS-1745463.

References

Afuah, A., and Tucci, C. L. 2012. “Crowdsourcing as a Solution to Distant Search,” *Academy of Management Review* (37:3), pp. 355-375.

Anderson, M. J. 2006. “Distance-based Tests for Homogeneity of Multivariate Dispersions,” *Biometrics* (62:1), pp. 245-253.

Anderson, M. J., Ellingsen, K. E., and McArdle, B. H. 2006. “Multivariate Dispersion as a Measure of Beta Diversity,” *Ecology Letters* (9:6), pp. 683-693.

Bjelland, O. M., and Wood, R. C. 2008. “An Inside View of IBM's' Innovation Jam',” *MIT Sloan Management Review* (50:1), pp. 32.

Bogers, M., and West, J. 2012. “Managing Distributed Innovation: Strategic Utilization of Open and User Innovation,” *Creativity and Innovation Management* (21:1), pp. 61-75.

Chesbrough, H. W. 2006. *Open Innovation: The New Imperative for Creating and Profiting from Technology*, Harvard Business Press.

Chiu, C. M., Liang, T. P., and Turban, E. 2014. “What Can Crowdsourcing Do for Decision Support?,” *Decision Support Systems* (65), pp. 40-49.

Cvijovic, D., and Klinowski, J. 1995. “Taboo Search: An Approach to the Multiple Minima Problem,” *Science* (267:5198), pp. 664-666.

Dahlander, L., O'Mahony, S., and Gann, D. M. 2016. “One foot in, one foot out: how does individuals' external search breadth affect innovation outcomes?,” *Strategic Management Journal* (37:2), pp. 280-302.

Flath, C. M., Friesike, S., Wirth, M., and Thiesse, F. 2017. “Copy, Transform, Combine: Exploring the Remix as a Form of Innovation,” *Journal of Information Technology* (32:4), pp. 306-325.

Fleming, L. 2001. “Recombinant Uncertainty in Technological Search,” *Management Science* (47:1), pp. 117-132.

Füller, J., Matzler, K., and Hoppe, M. 2008. “Brand Community Members as a Source of Innovation,” *Journal of Product Innovation Management* (25:6), pp. 608-619.

Gassmann, O., Sandmeier, P., and Wecht, C. H. 2005. “Extreme Customer Innovation in the Front-end: Learning from a New Software Paradigm,” *International Journal of Technology Management* (33;1), pp. 46-66.

Han, Y., and Nickerson, J. V. 2015. “Exploring Design Space Through Remixing,” *Collective Intelligence*.

Harryson, S., Kliknaite, S., and Dudkowski, R. 2008. “Flexibility in Innovation through External Learning: Exploring Two Models for Enhanced Industry University Collaboration,” *International Journal of Technology Management* (41:1-2), pp. 109-137.

Hill, B. M., and Monroy-Hernández, A. 2013. “The Remixing Dilemma: The Trade-off between Generativity and Originality,” *American Behavioral Scientist* (57:5), pp. 643-663.

Howe, J. 2006. “The Rise of Crowdsourcing,” *Wired Magazine* (14:6), pp. 1-4.

Jeppesen, L. B., and Lakhani, K. R. 2010. “Marginality and Problem-solving Effectiveness in Broadcast Search,” *Organization Science* (21:5), pp. 1016-1033.

Katila, R., and Ahuja, G. 2002. “Something Old, Something New: A Longitudinal Study of Search Behavior and New Product Introduction,” *Academy of Management Journal* (45:6), pp. 1183-1194.

Kohn, N. W., and Smith, S. M. 2011. "Collaborative Fixation: Effects of Others' Ideas on Brainstorming," *Applied Cognitive Psychology* (25:3), pp. 359-371.
Korf, R. E. 1985. "Depth-first Iterative-deepening: An Optimal Admissible Tree Search," *Artificial Intelligence* (27:1), pp. 97-109.
Kruskal, J. B., and Wish, M. 1978. *Multidimensional Scaling*, Sage.
Kyriakou, H., Nickerson, J. V., and Sabnis, G. 2017. "Knowledge Reuse for Customization: Metamodels in an Open Design Community for 3D Printing," *MIS Quarterly* (41:1), pp. 315-332.
Lakhani, K. R., Jeppesen, L. B., Lohse, P. A., and Panetta, J. A. 2007. *The Value of Openness in Scientific Problem Solving*, Division of Research, Harvard Business School.
Laursen, K., and Salter, A. 2006. "Open for Innovation: The Role of Openness in Explaining Innovation Performance among UK Manufacturing Firms," *Strategic Management Journal* (27:2), pp. 131-150.
Lim, K., Chesbrough, H., and Ruan, Y. 2010. "Open Innovation and Patterns of R&D Competition," *International Journal of Technology Management*, (52:3/4), pp. 295-321.
Majchrzak, A., and Malhotra, A. 2013. "Towards an Information Systems Perspective and Research Agenda on Crowdsourcing for Innovation," *The Journal of Strategic Information Systems* (22:4), pp. 257-268.
Majchrzak, A., and Malhotra, A. 2016. "Effect of knowledge-sharing trajectories on innovative outcomes in temporary online crowds," *Information Systems Research* (27:4), pp. 685-703.
Malone, T. W., Nickerson, J. V., Laubacher, R., Fisher, L, de Boer, P., Han, Y., and Towne, W. B. 2017. "Putting the Pieces Back Together Again: Contest Webs for Large-Scale Problem Solving," in *CSCW*, pp. 1661-1674.
March, J. G. 1991. "Exploration and Exploitation in Organizational Learning," *Organization Science* (2:1), pp. 71-87.
Mason, W., and Suri, S. 2012. "Conducting Behavioral Research on Amazon's Mechanical Turk," *Behavior Research Methods* (44:1), pp. 1-23.
Nickerson, J.V., and Yu, L. 2012. "Going Meta: Design Space and Evaluation Space in Software Design," in *Software Designers in Action*, M. Petre, and A.van der Hoek, (eds.), CRC Press, pp. 323-344
Nickerson, J. V. 2015. "Collective Design: Remixing and Visibility," in *Design Computing and Cognition'14,* J. Gero and S. Hanna (eds.), Springer, Cham, pp. 263-276.
Oksanen, J., Kindt, R., Legendre, P., O'Hara, B., Stevens, M. H. H., Oksanen, M. J., and Suggests, M. A. S. S. 2007. "The Vegan Package," *Community Ecology Package* (10), pp. 631-637.
Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., and Vanderplas, J. 2011. "Scikit-learn: Machine learning in Python," *Journal of Machine Learning Research* (12), pp. 2825-2830.
Petruzzelli, A. M., and Savino, T. 2014. "Search, Recombination, and Innovation: Lessons from Haute Cuisine," *Long Range Planning* (47:4), pp. 224-238.
Poetz, M. K., and Schreier, M. 2012. "The Value of Crowdsourcing: Can Users Really Compete with Professionals in Generating New Product Ideas?," *Journal of Product Innovation Management* (29:2), pp. 245-256.
Purcell, A. T., and Gero, J. S. 1996. "Design and Other Types of Fixation," *Design Studies* (17:4), pp. 363-383.
Resnick, M., Maloney, J., Monroy-Hernández, A., Rusk, N., Eastmond, E., Brennan, K., and Kafai, Y. 2009. "Scratch: Programming for All," *Communications of the ACM* (52:11), pp. 60-67.
Schiele, H. 2010. "Early Supplier Integration: The Dual Role of Purchasing in New Product Development," *Randd Management* (40:2), pp. 138-153.
Siangliulue, P., Chan, J., Dow, S. P., and Gajos, K. Z. 2016. "IdeaHound: Improving Large-scale Collaborative Ideation with Crowd-powered Real-time Semantic Modeling," in *Proceedings of the 29th Annual Symposium on User Interface Software and Technology*, Tokyo, Japan, pp. 609-624.
Stanko, M. A. 2016. "Toward a Theory of Remixing in Online Innovation Communities," *Information Systems Research* (27:4), pp. 773-791.
Von Ahn, L., and Dabbish, L. 2004. "Labeling Images with a Computer Game," in *Proceedings of the SIGCHI conference on Human factors in computing systems*, Vienna, Austria, pp. 319-326.
Wisdom, T. N., and Goldstone, R. L. 2011. "Innovation, Imitation, and Problem Solving in a Networked Group," *Nonlinear Dynamics-Psychology and Life Sciences* (15:2), pp. 229.