Algorithmically-Generated Communities: A Case Study

Austin L. Toombs
Michael A. Davidge
toombsa@purdue.edu
mdavidge@purdue.edu
Purdue University
Lafayette, Indiana

Jeongjin Park
Gregory Sirko
park1125@purdue.edu
gsirko@purdue.edu
Purdue University
Lafayette, Indiana

Michael LaPeter
the100.io
Portland, Oregon
mike@the100.io

ABSTRACT

In this poster, we present a case study of the100.io, an online platform that sorts gamers into groups that support gameplay and other activities. Through our analysis, we have identified that communities in which individuals are algorithmically sorted can thrive, even when the conditions of that sorting are somewhat arbitrary. We identify opportunities for CSCW researchers to further engage in the presence of shared identities, conversations about group atmosphere, and engagement with external tools as a means of indicating group health and success.

INTRODUCTION AND BACKGROUND

We frame this case among scholarship investigating how communities can be conceptualized as the users of sociotechnical systems [7], inclusive of communities that are situated in online, hybrid, and physical spaces. In this particular study, we focus on the100.io, an online community associated with the online multiplayer games Destiny and Destiny 2. How satisfied players are with these types of online communities can make or break a game [2], and member retention and commitment has a

Woodstock '18, June 03–05, 2018, Woodstock, NY

ACM ISBN 978-1-4503-9999-9/18/06…$15.00
https://doi.org/10.1145/1122445.1122456

Table 1: Descriptive Statistics about the Data Set

Attribute	Larger Data Set	Focused Data Set
Avg. Age	30	33
Avg. Size	52	68
Avg. Activity	7	23
% on PC	16.1%	9.8%
% on XBOX One	37.5%	49.2%
% on PS4	43.3%	40.1%
% on Other	3.1%	NA

[1]We refer to these as "recommended groups" because they are groups that the100 recommends that people join based on having a group size of less than 100 and at least a minimum level of activity.
Activity scores on the100 are calculated each day and depend on recent events. Our initial data set was generated based on the most active groups on June 3, 2019.

large impact on that satisfaction and sustainability [3]. Thus, we have begun to study how the100 as a platform supports the online community associated with these games, building on work related to online community maintenance practices [1, 5] and their related social processes [4].

One unique approach that the100 takes to maintaining its online communities is in ensuring that the groups do not become so large that their members no longer feel a sense of connection or a sense of belonging within them. Through the system, gamers can designate a platform, a timezone, and a number of other characteristics about how they like to play in order to be "matched" with a group of 100 similar gamers. This platform has been studied before (e.g., [6]) because of this unique matchmaking process. Our interests in this platform are in exploring how online communities can be bounded or limited along their own set of (*helpful*) arbitrary dimensions. Where online groups are not limited by geography, what other factors can help ensure the communities are manageable? When individuals are looking to join a group of "like-minded" people, how do they search and filter for those characteristics? Through this paper, we contribute to research on online communities and community maintenance practices by presenting the preliminary results of our case study about this platform. Our goal with this study is to begin a conversation about what we have been calling "algorithmically-generated communities" in order to see how else such a concept can be applied.

RESEARCH APPROACH

After exploring the groups the100 placed us in upon signing up, we used two approaches to generate our data set. First, we collected recommended groups from the "most active groups" list on the100.io/community[1]. Second, we crawled through the full list of groups available for the destiny games, selecting every 4th group that had more than 5 members and was not already included in our data set. This secondary strategy ensured that our data set included groups that were too large or not active enough to have been recommended groups. This resulted in a data set of 386 groups, which we filtered down to 61 (the "Focused Data Set") for in-depth coding. These data sets are summarized in Table 1. Through open coding, 3 members of the team generated 100 codes. After axial coding, our new codebook of 36 codes were reapplied to the Focused Data Set. Two coders coded each of the groups, and each pair of coders discussed their coding strategies afterward. The goal was not to reach consensus but to identify patterns among the groups. A code co-occurrence table helped to further identify such patterns.

FINDINGS

Our preliminary findings focus on patterns we identified among highly active groups, or groups that appeared to chat frequently or who appeared to have frequent play sessions planned. We focus in this preliminary analysis on patterns in our data that involve the 6 codes listed in Sidebar 1, which provide the most relevant cross-section of our results.

	Group Identity: Branding (33)	Group Identity: Making It Their Own (21)	External Tools: Communication and Extension (39)	Reference Material: Gameplay Improvement (23)	Group Descriptors: High Level of Activeness (36)	Group Atmosphere: Fun (25)
Group Identity: Branding (33)	x	18	30	20	29	16
Group Identity: Making It Their Own (21)	18	x	20	13	19	12
External Tools: Communication and Extension (39)	30	20	x	20	33	22
Reference Material: Gameplay Improvement (23)	20	13	20	x	20	12
Group Descriptors: High Level of Activeness (36)	29	19	33	20	x	19
Group Atmosphere: Fun (25)	16	12	22	12	19	x

Figure 1: Co-occurrence Matrix. This presents the number of times each of these codes of interest occurred with each other.

	Group Identity: Branding (33)	Group Identity: Making It Their Own (21)	External Tools: Communication and Extension (39)	Reference Material: Gameplay Improvement (23)	Group Descriptors: High Level of Activeness (36)	Group Atmosphere: Fun (25)
Group Identity: Branding (33)	-	85.71%	76.92%	86.96%	80.56%	64.00%
Group Identity: Making It Their Own (21)	54.55%	-	51.28%	56.52%	52.78%	48.00%
External Tools: Communication and Extension (39)	90.91%	95.24%	-	86.96%	91.67%	88.00%
Reference Material: Gameplay Improvement (23)	60.61%	61.90%	51.28%	-	55.56%	48.00%
Group Descriptors: High Level of Activeness (36)	87.88%	90.48%	84.62%	86.96%	-	76.00%
Group Atmosphere: Fun (25)	48.48%	57.14%	56.41%	52.17%	52.78%	-

Figure 2: Co-occurrence Percentage Matrix. This compares the number of times each of the codes occurred with each other as a proportion of the number of times each code was used overall.

Branding and Reference Materials: Highly active groups tended to incorporate their own branding, graphics, and even merchandising into their the100 pages. In fact, the codes Group Descriptors: High Level of Activeness and Group Identity: Branding co-occurred 29 times, accounting for 80.56% of the times that Group Descriptors: High Level of Activeness was coded and 87.88% of the times that Group Identity: Branding was coded. (To save space, we will not report these statistics going forward, which can be found in Figures 1 and 2.) Branding also co-occurred highly with the codes we applied when groups linked to external reference materials or employed other mechanisms for improving members' gameplay. Highly active groups tended to provide more reference and onboarding materials for new and current members, such as guides for specific in-game mechanics or instructions for how to create an event in the100 that others could join.

External Tools and Group Atmosphere: Groups coded as highly active also tended to use external tools to facilitate communication, such as Slack, Discord, or similar services: *"As a group, we have set up a discord channel [Discord link] that has mostly replaced the chat on this site. More fluid, with separated channels and the ability to retain posts longer. Check it out."* The external communication tools code also co-occurred highly with the codes for branding and the code applied to groups that customize their own pages and group names beyond what has been generated by the system ("Making it their own"). These co-occurrences align with some of our initial assumptions that prolific groups would make use of external platforms to further accommodate new members or more efficiently manage different avenues of interaction. Finally, groups that described themselves as having a "fun" atmosphere also tended to be more active and to use external tools. This suggests that the goal of having fun was attractive to the kind of potential members who actually intended to engage with others on this platform, and that members enjoyed engaging with each other enough to continue those engagements across platforms.

DISCUSSION AND FUTURE WORK

We interpret the above data as having implications for how these ecologies of tools interact with the group's successes or failures and how individuals develop identities as members of these online groups. During initial stages of group analysis, we noticed that activity scores, which are calculated based on how many events a group runs and how many individuals join it, did not directly reflect group activity. After reaching some minimum of regular activity on the100, some groups moved operations to an external platform, such as Discord or Slack. However this transition was not consistent, which surprised us. While some groups reached an activity score of zero on the100 despite being fairly active outside of the platform, other groups continued to use the platform in addition to their external tools. This might be because the100 provides a low barrier to entry for joining upcoming events (a press of a button) that would be cumbersome to replicate in Discord or Slack. Through interviews and observations, we will continue to explore how and why these groups use their external tools. Is

Group Identity: Branding
A group appropriates their group name (nickname or the100.io given name) for use in graphics or merchandising.

Group Identity: Making It Their Own
A group customizes given clan features (e.g.: their auto-generated name), either by shortening it or referring to themselves through a nickname of some kind.

External Tools:
Communication and Extension
A group employs the use of external, third-party tools on their homepage. These include things like Slack, Discord, YouTube, Bungie.net, and social media.

Reference Material:
Gameplay Improvement
A group links or lists informational tools for a certain game with the intent of helping players improve.

Group Descriptors:
High Level of Activeness
A group operates at a high level of activeness (e.g.: having multiple games played, an active chat, or future game events planned out.)

Group Atmosphere: Fun
A group explicitly states their desire for having fun and demonstrates activities of camaraderie through jokes, chat, or other activities.

Sidebar 1: Highlights from our codebook.

[2]A fictitious group name

it because they have become too large for the100 to be a viable tool for them? Additionally, why do these groups choose not to use the group-building tools developed for Destiny and Destiny 2 by the developer (Bungie)? We hypothesize this may have something to do with how the interpersonal dynamics in these groups are supported (or not) by these various systems.

Beyond the use of these external tools, we found the strong identities that groups developed through the100 to be surprising, given that the individuals were largely "sorted" into those groups by the platform's algorithms. The fifth author, who created the100, has been specifically surprised with how some of the groups latch on so strongly to the name of their group, which is generated by a simple algorithm. Groups will take a name like "Delta Company 209"[2] and create graphics and clothing with that group name on them. Some groups do take advantage of the ability to create a group with a name they have chosen, but we were unable to see any differences among how individuals interacted in algorithmically-named groups versus self-named groups in this preliminary analysis. This leads us to question what roles community members might be taking on to be able to create these graphics and related identity materials. How are these activities facilitated? Are there a few dominant members leading the charge, or are these decisions made democratically? At what point might the group size have an affect on these activities? In our future work, we hope to explore these questions.

CONCLUSION

What we have seen in this early study is that communities in which individuals are algorithmically sorted can thrive, even when the conditions of that sorting algorithm are somewhat arbitrary. We found that the presence of shared identity development, conversations about community atmosphere, and the willingness of a few group members to set up external tools were among the most important factors in determining how successful a group could be. These findings surprised us, because we assumed that some other measure of potential community homophily (such as stage of life, nationality, or some other expressed personal identity) would have a strong enough impact to be noticeable in our data set. During the poster session at CSCW 2020, we hope to discuss these insights with other academics and practitioners in order to see how else this concept of an algorithmically-generated community can be applied. As we progress in our study, we continue to find opportunities for additional application, such as how such groups could help those moving to a city, those who want to shrink the sizes of their social media audiences, or those who wish to shift the online echo chambers in which they currently participate.

ACKNOWLEDGMENTS

We would like to thank the other members of the Community-Computer Interaction Lab (www.c-cilab.com) at Purdue University for their support.

REFERENCES

[1] Brian Butler, Lee Sproull, Sara Kiesler, and Robert Kraut. 2008. Community Effort in Online Groups: Who Does the Work and Why? In *Leadership at a distance Research in technologically supported work*, Suzanne P Weisband (Ed.). Psychology Press, 171–194.

[2] Yiwei Cao, Anna Glukhova, Ralf Klamma, and Dominik Renzel. 2009. A Community Success Model for Gaming Communities. *Journal of Multimedia* 4 (04 2009). https://doi.org/10.4304/jmm.4.2.87-93

[3] Rosta Farzan, Laura A. Dabbish, Robert E. Kraut, and Tom Postmes. 2011. Increasing Commitment to Online Communities by Designing for Social Presence. In *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work (CSCW '11)*. ACM, New York, NY, USA, 321–330. https://doi.org/10.1145/1958824.1958874

[4] Derek L Hansen, Mark S Ackerman, Paul J Resnick, and Sean Munson. 2007. Virtual community maintenance with a collaborative repository. *Proceedings of the Association for Information Science and Technology* 44, 1 (Jan. 2007), 1–20.

[5] Austin L Toombs, Shaowen Bardzell, and Jeffrey Bardzell. 2015. The Proper Care and Feeding of Hackerspaces: Care Ethics and Cultures of Making. In *CHI '15: Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM Request Permissions, 629–638.

[6] Günter Wallner and Simone Kriglstein. 2018. Introducing PlaNet: A Tool for Visualizing Player Communities. In *Proceedings of the 13th International Conference on the Foundations of Digital Games (FDG '18)*. ACM, New York, NY, USA, Article 42, 4 pages. https://doi.org/10.1145/3235765.3235808

[7] Brian Whitworth and Aldo de Moor. 2003. Legitimate by design: towards trusted socio-technical systems. *Behaviour IT* 22 (2003), 31–51.