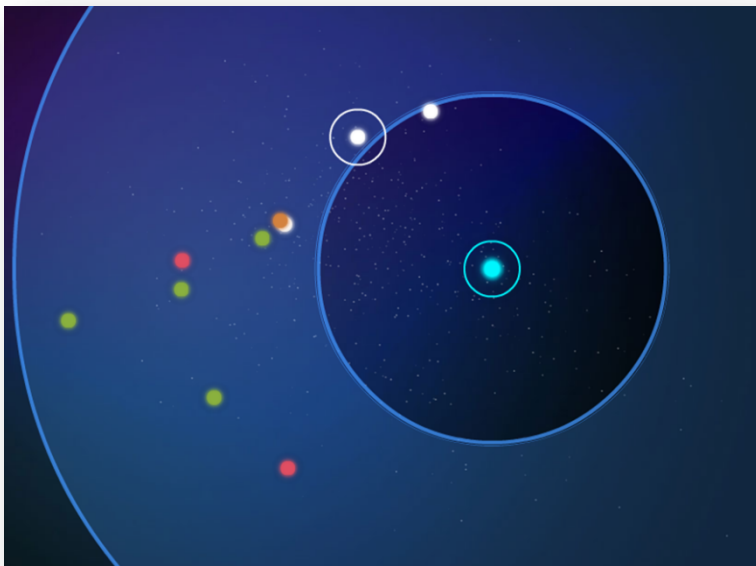# The Diversity Donut: Enabling Participant Control Over the Diversity of Recommended Responses

Figure 1. The Diversity Donut allows participants to indicate the level of diversity they would like to see in a recommended set of responses. Each point in the space represents a participant's response. By changing the radii of the inner and outer rings of the Diversity Donut, a participant changes the degree of diversity in the responses that will be presented.

David Wong[1]
david_wong@berkeley.edu

Siamak Faridani[2]
faridani@berkeley.edu

Ephrat Bitton[2]
ebitton@berkeley.edu

Bjoern Hartmann[1]
bjoern@eecs.berkeley.edu

Ken Goldberg[1,2]
goldberg@berkeley.edu

[1]EECS Department
University of California, Berkeley
[2]IEOR Department
University of California, Berkeley

Abstract
Most online discussion interfaces organize textual responses using linear lists. Such lists do not scale to the number of responses and cannot convey the diversity of the participants who have contributed. The Opinion Space system is designed to address these issues. In this paper, we augment Opinion Space with two features. The first is a new user interface tool and recommendation system: the Diversity Donut (Figure 1). While the Diversity Donut did not establish a statistical advantage over other recommendation methods, participant self-reported data suggested that participants found the Diversity Donut to yield the most diverse set of comments. The second contribution is a new dimensionality reduction technique in Opinion Space: Canonical Correlation Analysis (CCA). Our analysis suggests that CCA is a better algorithm for opinion visualization than Principal Component Analysis (PCA).

Author Keywords
Diverse Opinion Browsing, Recommendation Algorithms, Canonical Correlation Analysis, Opinion Visualization

Figure 3. Each participant of Opinion Space enters their opinions on five statements on a slider scale. Participants also answer a discussion question with a textual response.

Figure 2. A screenshot of the Opinion Space system.

Introduction

Since 2000, the volume of interaction in social media has grown significantly. However, existing systems for online discussion such as linear lists do not scale well. The first problem with lists is that the amount of data presented to a participant can be overwhelming. For example, news stories and blog posts can generate hundreds or thousands of responses. Lists bias responses at the top and make navigating through responses unmanageable. Furthermore, the linear interface impedes consideration of diverse responses.

Opinion Space is a social media technology that addresses these challenges using techniques from deliberative polling, collaborative filtering, and multidimensional visualization [7]. Opinion Space collects opinions on baseline statements as scalar values on a continuous scale (Figure 3) and applies dimensionality reduction to project the data onto a two-dimensional plane for visualization and navigation. This technique effectively places all participants onto a level playing field. Points far apart correspond to participants with very different opinions while participants with similar opinions are proximal (the converses do not necessarily hold). Participants in Opinion Space contribute textual responses to discussion questions and are encouraged to earn points through reading and rating the responses of others (Figure 4). In this paper, we describe the following two innovations.
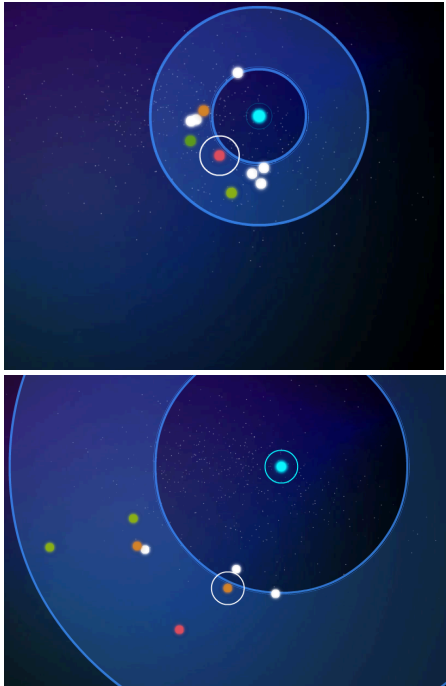
The Diversity Donut
The Diversity Donut is an interactive recommender tool. A participant controls the level of response diversity displayed by adjusting the inner and outer rings of a "donut" that centers about the participant's point in the space (Figure 5).

Canonical Correlation Analysis (CCA)
The existing Opinion Space platform uses Principle Component Analysis (PCA) to project participants onto a two-dimensional plane. This method only utilizes the participant's slider ratings and does not consider the textual response in the calculation of a participant's position in the space. CCA, on the other hand, uses both the participant's slider data and textual response to calculate the participant's position. This enables us to capture the sentiment of the participant's textual response as part of the projection. A participant's response is stemmed, featurized, and fed into CCA along with their slider data to determine the participant's position. We designed a formal
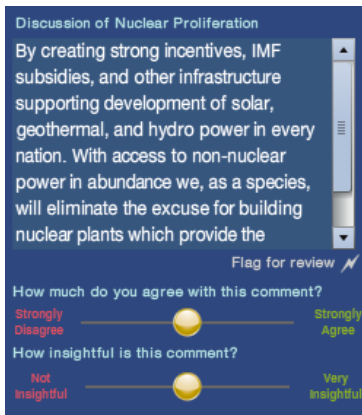
Figure 4. A participant rates the responses of others on two scales: how much they agree with the response and how insightful they find the response.

Figure 5. The Diversity Donut allows a participant to indicate the level of diversity in the responses they would like to see. A user can shrink the donut to see responses from like-minded participants (above), or expand the donut to see responses from participants who differ in opinion (below).

quantitative framework to compare the quality of the projections yielded by PCA and CCA. Our results suggest that CCA is a better projection method than PCA for Opinion Space.

Related Work

Aggregation of Opinions and Deliberative Polling
Fishkin proposed a way to aggregate opinions called "Deliberative Polling" [1]. This method first polls individuals on issues and then allows participants to discuss the issues together. Following the discussion, participants are polled again. Fishkin argues that deliberative polling yields more informed conclusions.

Diversity in Opinions
Munson and Resnick suggest that participants in online environments can be clustered into two distinct subgroups: "Diversity Seeking" and "Challenge Averse" [2]. Generally speaking, challenge averse participants are primarily interested in opinions that are similar to their own, while diversity seeking participants are interested in considering opinions that challenge their own. Munson and Resnick found that even with several UI interfaces that highlighted responses that were different from the participant's opinion, challenge averse participants did not become more accepting of opinions that differed from their own.

Sunstein describes the problem of "cyberpolarization" in online environments [3]. He argues that participants who frequently read responses from other like-minded participants have the tendency to be pushed to extremes of the political spectrum.

Canonical Correlation Analysis
CCA finds a linear transformation for two sets of data that maximizes the correlation between the transformed sets of data. Jordan and Bach provided a probabilistic interpretation of CCA [4]. Hardoon et al. extended Jordon and Bach's work, created a kernel method, and demonstrated that a kernel CCA method provides better results than a purely CCA method [5].
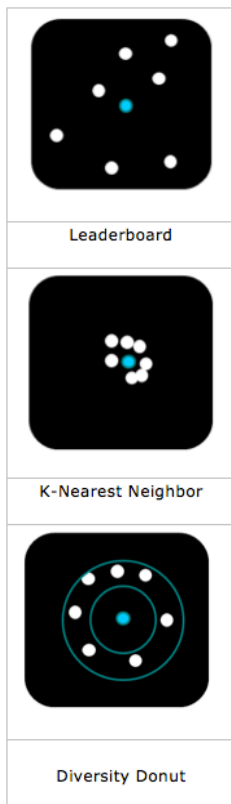
Evaluation of Projection Quality

We evaluate the quality of a dimensionality reduction method by calculating the average Pearson correlation between the agreement rating (the first slider in Figure 4) and the distance between the rater and author of a response for all ratings in the system. This framework allows us to quantitatively compare dimensionality reduction techniques. A higher quality dimensionality reduction technique yields a higher negative correlation.

The Diversity Donut

The Diversity Donut empowers participants to directly indicate the level of diversity in the responses they view. Centered on the participant's point in Opinion Space, the Diversity Donut can be adjusted to present the participant with responses that are from like-minded participants and from participants with differing opinions (Figure 5). The area between the two rings of the donut defines the spatial region that is queried for responses. After the participant defines the region, the underlying recommendation system retrieves responses from participants whose points fall within that region.

We implemented a basic recommendation engine for the Diversity Donut that intended to recommend participants a diverse set of responses. This means that

Leaderboard

K-Nearest Neighbor

Diversity Donut

Table 1. Each participant in the study evaluated responses recommended by the three recommendation methods above.

Dimensionality Reduction Method	Pearson's Correlation Between Distance and Agree
CCA	$\rho=-0.352$
PCA	$\rho=-0.134$
RND	$\rho=0.0021$

Table 2. Canonical Correlation Analysis provides the highest negative correlation between spatial distances and agreement rating values of responses. We conclude that this dimensionality reduction method is best suited for the visualizing opinions in Opinion Space.

the algorithm may recommend a response that a user disagrees with, but may still find insightful. This approach differs from traditional collaborative filtering and recommendation systems (e.g. Amazon.com and Netflix), which recommend participants a set of items that they most likely would agree with. Arguably, the application of a traditional recommender algorithm in Opinion Space would encourage cyberpolarization: if our goal were to recommend responses that participants are most likely to rate highly, then participants would only see responses that reinforce their own opinions.

The recommendation algorithm first clusters all participants in CCA space using k-means clustering. For every participant cluster, the system aggregates each participant's agreement ratings and calculates the average agreement rating for all responses in the system. After a participant defines a spatial search region with the donut, the system returns the responses that have the lowest average agreement in the cluster and are within the search region.

Participant Study
To evaluate the Diversity Donut, we selected a set of 118 thoughtful responses that reflect a range of opinions from our Opinion Space 2.0 dataset. We conducted a controlled user study with 13 participants. Five participants were Berkeley students and eight were recruited from Mechanical Turk. An experimenter was available online for any questions that the Mechanical Turk participants may have had. Before starting the experiment, participants completed a tutorial that used both images and explanatory text to explain the Opinion Space interface.

Study Design
Participants were first asked to fill out a prescreening survey to determine whether they were challenge averse or diversity seeking. Next, they were asked to complete their opinion profile (Figure 3). They rated five statements using a sliding scale that ranged from 'Strongly Disagree' to 'Strongly Agree' and entered a textual response to a discussion question. Participants were then recommended three sets of 10 or more responses and were asked to read and rate each response on how much they agreed with the response and how insightful they found it. Each set of responses was shown using a different recommendation algorithm: the leaderboard model, K-Nearest-Neighbors (KNN) algorithm, or the Diversity Donut (Table 1). The leaderboard model is the existing global ranking reputation system in Opinion Space. This metric takes into account the spatial relationship between participants in the space and ranks responses using all of the ratings in the system [6]. The KNN model is a benchmark recommendation system that makes personalized recommendations based on the preferences of those in the immediate neighborhood of the participant. After each recommendation method, the participant filled out a survey asking questions regarding the quality of the responses and the participant's satisfaction. At the end of the experiment, the participant filled out an exit survey.
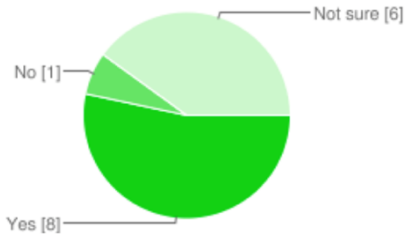
Evaluation
All participant activity on the space was recorded, including responses viewed, agreement and insight rating values of other responses, dwell time per response, and the radii of the Diversity Donut's inner and outer rings. The surveys asked the participants questions regarding their experience with the system,

Figure 6. "Do you prefer the diversity donut to a purely automated approach?"

Figure 7. "How satisfied were you with the diversity of the opinions expressed in the recommended set of responses?"

Figure 8. "Did you see a good range of opinions in the recommended responses?"

such as "How satisfied were you with the recommended set of responses?" The surveys also included 7-point Likert scales that asked about the participant's satisfaction with various aspects of the experiment.

Hypotheses

Hypothesis 1: Dimensionality reduction techniques
CCA will outperform PCA as a dimensionality reduction technique according to the method of evaluation described in "Evaluation of Projection Quality".

Hypothesis 2A: Insightfulness of responses
Participants will find the responses recommended by the Diversity Donut to be more insightful than those recommended by KNN.

Hypothesis 2B: Agreement with responses
Participants will agree more with the responses recommended by the KNN model than with those recommended by the Diversity Donut.

Hypothesis 3: Self-reported data
Participants will report that the responses recommended by the global ranking method and the Diversity Donut will be more diverse than those recommended by KNN and that there will be higher satisfaction in using the Diversity Donut.

Results

Projection Quality
We found that CCA had the greatest negative correlation between spatial distance and agreement rating of responses. We concluded that Hypothesis 1 is supported and CCA is the most effective dimensionality reduction method for Opinion Space. Table 2

summarizes the results for our evaluation of different dimensionality reduction methods.

Diversity Donut
A single factor ANOVA analysis with repeated measures was performed on the insightfulness ratings ($p=0.2167$) and agreement ratings ($p=0.2047$) across all three recommender methods. The results from this analysis were inconclusive in determining which recommendation method yields the highest agreement or highest insightfulness ratings. There are three possible explanations for this result: 1) 13 participants is not sufficiently large enough to establish a statistical significance, 2) there is no significant difference between the three algorithms in terms of the mean values of insightfulness and agreement ratings, or 3) the dataset of responses used was not sufficiently diverse in content and opinion. All of these possibilities require further statistical analysis. Hypothesis 2A and 2B cannot be supported based on this preliminary data. Figures 6, 7, and 8 show the self-reported data from the surveys. Figures 7 and 8 illustrate the mean value and one standard deviation in each direction for the Likert scale data. These data suggest that participants generally preferred the Diversity Donut to an entirely automated approach. Additionally participants were more satisfied with the diversity of responses that were presented to them by the Diversity Donut, suggesting support for Hypothesis 3.
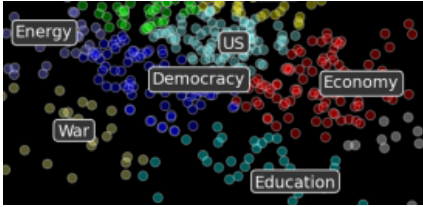
Conclusion

In this preliminary report, we compare CCA with PCA and investigate the effectiveness of the Diversity Donut. Our analysis of projection quality suggests that CCA is a more effective dimensionality reduction method for Opinion Space. Data on the Diversity Donut
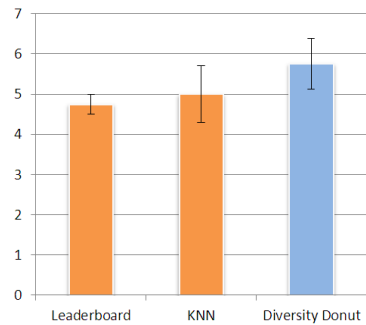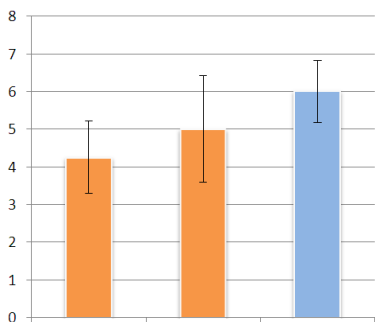
Figure 9. Region labeling of the topics of responses in CCA space.

is inconclusive, however participant self-reported data suggests that participants found the Diversity Donut to be an effective tool for recommending diverse responses.

Future work

The dataset of responses may not have been sufficiently diverse in content and opinion. We plan to further develop the system and evaluate it again using a dataset of responses that has a wider range of diversity in content. We are also planning to run another control experiment with a larger group of in-lab participants.

The inner and outer rings of the Diversity Donut allow a participant to select a symmetric search region, but the CCA space is not symmetric around the participant. This suggests that a "Diversity Lasso" that is not centered on the participant's point may be more useful. Since CCA provides us with strong topic modeling and region labeling capability (Figure 9), we can use this data and augment the space with topic labels, in which participants can see the main topic for each region in space. Participants can then use the lasso to select a region that they are interested in. We would like to analyze how more general tools such as the lasso perform in comparison to symmetric tools like the Diversity Donut.

References

1. J.S. Fishkin and R.C. Luskin. Experimenting with a democratic ideal: Deliberative polling and public opinion. Acta Politica, 40(3):284-298, 2005.
2. Munson, S.A. and Resnick, P. "Presenting diverse political opinions: how and how much", Proceedings of the 28th international conference on Human factors in computing systems. 1457-1466, ACM, 2010.
3. Sunstein, C. (2002). "The Law of Group Polarization," The Journal of Political Philosophy 10(2): 175-195.
4. F. Bach, M. I. Jordan. A probabilistic interpretation of canonical correlation analysis. Technical Report 688, Department of Statistics, University of California, Berkeley, 2005.
5. Hardoon, D.R. and Szedmak, S. and Shawe-Taylor, J. Canonical correlation analysis: an overview with application to learning methods, Neural Computation, 16,2639--2664,2004,MIT Press.
6. E. Bitton. A spatial model for collaborative filtering of responses in an online discussion forum. In Proceedings of the third ACM conference on Recommender systems, pages 393–396. ACM, 2009.
7. S. Faridani, E. Bitton, K. Ryokai, and K. Goldberg. Opinion space: a scalable tool for browsing online responses. In Proceedings of the 28th international conference on Human factors in computing systems, pages 1175–1184. ACM, 2010.