

Comment Section Personalization: Algorithmic, Interface and Interaction Design

Yixue Wang

Northwestern University

yixue.wang@u.northwestern.edu

Abstract

Comment sections allow users to share their personal experiences, discuss and form different opinions, and build communities out of organic conversations. However, many comment sections present chronological ranking to all users. In this paper, I discuss personalization approaches in comment sections based on different objectives for newsrooms and researchers to consider. I propose algorithmic and interface designs when personalizing the presentation of comments based on different objectives including relevance, diversity, and education/background information. I further explain how transparency, user control, and comment type diversity could help users most benefit from the personalized interacting experience.

1 Introduction

Comment sections provide a public digital space for users to exchange ideas, share personal experiences, and form opinions, which are all key elements of deliberative democracy (Kim et al., 1999). However, many comments ranked highly by comment sections tend to be early comments due to greater visibility and resulting in greater capacity for a high community rating (Hsu et al., 2009), making other good quality and relevant comments less visible, and providing the same reading experience for all users. While comment sections can utilize different moderation strategies to promote high-quality comments (Wang and Diakopoulos, 2021b) and reduce the likelihood of uncivil conversations (Cheng et al., 2017), they lack the ability to promote diverse, and/or minority opinions and offer background information on the topics (Janssen and Kies, 2005) according to users' needs.

Personalization might help address this issue. News personalization has been defined as “a form of user-to-system interactivity that uses a set of

technological features to adapt the content, delivery, and arrangement of a communication to individual users' explicitly registered and/or implicitly determined preferences” (Thurman and Schifferes, 2012). 70 percent of 200 publishers personalize the content they deliver to their visitors (Weiss, 2019). Newsrooms have implemented different personalization approaches, including automatic content tagging and ad-targeting, documenting readers' locations, and reading behaviors (e.g., keywords and phrases in the articles), in order to customize the delivery of news and encourage users' engagement^{1 2}.

Though many newsrooms have incorporated different personalization approaches, the personalization of comments is still under-examined. How will comment personalization help the audience better understand the topic and promote deliberative conversations online in the future? And how can researchers, developers, and journalists design comment sections to customize readers' reading experience while maintaining the comment section as a common ground for all users? This short paper seeks to propose different design and algorithmic approaches to support different personalization objectives.

2 Objectives and Design of Personalized Comment Sections

People read news comments for various reasons: to learn about the opinions of others, to be entertained or amused by others' comments, to see how their opinion of the story or topic differs from others' views, to get more information on a story, to get

¹<https://www.nytimes.com/2017/03/18/public-editor/a-community-of-one-the-times-gets-tailored.html>

²<https://www.niemanlab.org/2016/05/the-washington-post-tests-personalized-pop-up-newsletters-to-promote-its-big-stories/>

additional reporting/updates on a story, or to gauge the pulse of the community. And people comment on news for various reasons: to express an emotion or opinion, to add information, to correct inaccuracies or misinformation, to take part in the debate, to discuss with others, etc. (Stroud et al., 2016). How can newsrooms better personalize the comment sections according to these reading and commenting needs? This section will introduce several personalization objectives including relevance, diversity, education/background information, and how algorithms could support them.

2.1 Relevance

Relevance is the key driver of news consumption (Schröder, 2019). People are more likely to like and understand those who are similar to them and their experiences, i.e., language and demographics (McPherson et al., 2001). Therefore, it is important to keep the personalized comments relevant to readers. Relevance could be achieved via different approaches, such as by localization based on self-reported geographic information (i.e., geographic relevance), by collaborative filtering based on previous like history in comments and articles (i.e., topic relevance), or by ranking content and language similarity based on word embeddings' cosine similarities (Kenter and De Rijke, 2015) (i.e., writing language relevance).

With this objective, newsrooms need to collect metrics around users' historical commenting behaviors (e.g., likes and comment content) and users' location information. Then comment sections could rank the comments from high to low relevance based on users' historical comments. This design would be similar to what (Wang and Diakopoulos, 2021a) proposed in their ranking algorithm, in which the algorithm automatically ranks the comments based on language relevance between users' example input query and the sample comments in the system. One potential problem with merely focusing on this objective is that users might fear being trapped in filter bubbles where most comments they interact with are from people who are very similar to them and share similar opinions (Monzer et al., 2020), which leads to the next objective I want to discuss: diversity.

2.2 Diversity

People not only look for similar personal experiences and opinions, but also compare their own opinions to others' views to gauge the community's

overall trends (Stroud et al., 2016). Therefore a comment section only focusing on users' relevance might make the user lose the full picture of public interest (Plattner, 2018). Offering a variety of comments could also help users better understand others' views, opinions, and eventually promote online deliberation, and enable "a diverse and in-depth news diet" that readers value (Bodó et al., 2019).

To personalize diversity across comments, newsrooms need to again collect metrics around users' historical commenting behaviors (e.g., likes and comment content), location, etc. Comment sections could be grouped into different groups based on whether or not the content is similar to users' previous comment content (e.g., "comments that you might find familiar" and "comments that you might find not familiar"), or whether the content is from a close location (i.e. rural and urban could be treated as different groups). These comments could be grouped into different tabs for users to interact with, similar to the three-column comment section structure (i.e., "Supporting Legalization", "Questions about Legalization", and "Opposing Legalization") that Peacock et al. (Peacock et al., 2019) proposed. Comments could also be tagged as "similar comments to yours" and "different comments compared to yours" along with the comment content.

2.3 Education/Background Information

Comments not only open a common ground for users to share their expertise, personal stories and opinions for every user to learn from and compare with the stories, but they also hold journalists accountable (Greenwald and Fillion, 2017). To provide such a common ground for all users, comment sections should work as a platform for users to either contribute their knowledge in the comment section to interact with journalists' reporting and/or learn background information while reading comments. When users are experts in a specific topic they are browsing, and/or they find a topic less familiar and they need more information, how can comment sections personalize their reading and commenting experience?

I propose that comment sections could collect users' expertise areas and topics unfamiliar to them, through implicitly inferring users' interests based on users' reading history and users' explicit feedback (e.g., self-report ratings in a survey about

Objective	Algorithm Approach	Interface Design
Relevance	Localization, collaborative filtering, and word embedding similarity	Ranking
Diversity	Word embedding similarity	Tab/tagging
Education/Background Information	Text similarity and keyword extraction/matching	Prompt/links to resources

Table 1: Summary of different objectives, their corresponding algorithmic design and interface design

users’ knowledge in different topics) (Thurman and Schifferes, 2012). Comment sections would then match these topics with the current article users interact with via text similarity (e.g. cosine similarity) and/or keyword matching. Comment sections would prompt users to comment in the comment section when an article potentially matches their expertise. If users find a specific topic in a comment unfamiliar and not directly related to the main topic in the article they interact with, and want to explore this unfamiliar topic in depth, the comment section could also aggregate a combination of external Wikipedia links and internal news article links to provide background information.

Note that these three objectives could be pursued by the newsrooms at the same time, which could eventually be helpful to avoid users’ concern of filter bubbles and losing the big picture of public interest (Monzer et al., 2020). I summarize the three objectives along with their algorithmic design and data collection methods in Table 1.

3 User Interaction with Personalized Comment Section

I discuss how comment sections could be personalized in different designs based on different objectives in Section 2 and summarize how algorithms could support each objective in Table 1. In this section, I streamline an ideal interaction between users and a personalized comment section in Section 3.1 based on the three objectives in Section 2 and I further discuss how transparency, user control, and diversity of content types could help users have a better interactive experience with a personalized comment section.

3.1 An example interaction between a personalized comment section and users

Imagine you are about to interact with a comment section. You open the personalized comment section, and then it shows the default ranking of all comments (either in chronological order or

by popularity) to provide the same reading experience to all users. On top of the comment section, you have the ability to turn the personalization on or off through a drop-down menu. In this drop-down menu, you could select how you want to personalize the comments (i.e., personalized by relevance/diversity, more details in Section 2.1 and 2.2).

Once you select your personalization objective, the comment section will then automatically show the personalized curation and notify you that the comments are personalized based on geographics, previous commenting history, or pre-selected topic interests. You can choose to comment directly in the comment section and/or reply to others’ comments in a sub-thread. The system presents the opportunity to interact easily with not just “personal stories” but other content types, such as “opinions” and “questions” from the community, by filtering and selecting content based on their tags. A pop-up window notifies you that this topic is within your area of expertise, and it encourages to share your expertise with other users (see Section 2.3).

When interacting with the comment section, you discover some relevant experiences and opinions, you understand what others are talking about, and you contribute back to the community. And if you are not satisfied with the personalization, there is always a way to go back to the default interface.

This is an ideal interaction experience with a personalized comment section. To better support this interaction with a personalized comment section, I propose two interaction objectives for researchers and newsrooms to consider in the design process: transparency/user control and comment type diversity.

3.1.1 Transparency and User Control

The lack of transparency about the personalization process may lead to a lack of trust in receiving personalized news (Monzer et al., 2020). To gain users’ trust, a personalized comment section should notify

users whenever the comments are being personalized. Power users (i.e., highly self-motivated learners who have the expertise and interest in adopting new technologies and interface features) prefer having user controls that allow them to determine when to start/stop personalization (Sundar and Marathe, 2010). Therefore, a personalized comment section should allow users to turn personalization on and off by selecting personalization objectives from a dropdown menu. Users should also have the ability to independently change different personalization objectives.

3.1.2 Comment Type Diversity

Apart from diversifying the content based on the similarity between comments and users' previous posts, one way to further diversify and personalize the experience would be to provide a mix of different types of comment content (e.g., personal stories, opinions, threads containing questions, expertise, etc.), which may be detected through clustering algorithms or classification algorithms based on crowd-sourcing tags. Comments could be tagged with multiple types (e.g., personal story and opinions). In a personalized system, users should be able to interact with various types of content (Stray, 2021).

Access to diverse content could further benefit users' personalization experience by allowing them to filter what they want to see based on different tags (i.e., "personal story", "opinions", etc.) attached to different comments. And it may encourage users to learn the topic and the community more deeply if they want to focus on a specific perspective to investigate the topic (e.g., to follow commenters who have specific domain knowledge, to participate in the community debate, and to understand if community members have questions/doubts on a topic). In order to encourage users to read and interact with various kinds of content, a personalized comment section could even extend this interaction by notifying users when they only consume one type of content (e.g., personal story) while ignoring other potential types (e.g., opinions).

To summarize, an ideal comment section should personalize comment content based on different objectives, including relevance, diversity, and education/background information, and also provide a transparent and diversified interaction experience for users. By implementing these design objectives and approaches, comment sections could achieve a

personalized yet representative reading experience for all users.

References

- Balázs Bodó, Natali Helberger, Sarah Eskens, and Judith Möller. 2019. Interested in diversity: The role of user attitudes, algorithmic feedback loops, and policy in news personalization. *Digital Journalism*, 7(2):206–229.
- Justin Cheng, Michael Bernstein, Cristian Danescu-Niculescu-Mizil, and Jure Leskovec. 2017. Anyone can become a troll: Causes of trolling behavior in online discussions. In *Proceedings of the 2017 ACM conference on computer supported cooperative work and social computing*, pages 1217–1230.
- Glenn Greenwald and Rubina Madan Fillion. 2017. Comment sections are essential for news sites. we're making changes to improve ours. Accessed: 2021-02-28.
- Chiao-Fang Hsu, Elham Khabiri, and James Caverlee. 2009. Ranking comments on the social web. In *2009 International Conference on Computational Science and Engineering*, volume 4, pages 90–97. IEEE.
- Davy Janssen and Raphaël Kies. 2005. Online forums and deliberative democracy. *Acta política*, 40(3):317–335.
- Tom Kenter and Maarten De Rijke. 2015. Short text similarity with word embeddings. In *Proceedings of the 24th ACM international on conference on information and knowledge management*, pages 1411–1420.
- Joochan Kim, Robert O Wyatt, and Elihu Katz. 1999. News, talk, opinion, participation: The part played by conversation in deliberative democracy. *Political communication*, 16(4):361–385.
- Miller McPherson, Lynn Smith-Lovin, and James M Cook. 2001. Birds of a feather: Homophily in social networks. *Annual review of sociology*, 27(1):415–444.
- Cristina Monzer, Judith Moeller, Natali Helberger, and Sarah Eskens. 2020. User perspectives on the news personalisation process: Agency, trust and utility as building blocks. *Digital Journalism*, 8(9):1142–1162.
- Cynthia Peacock, Joshua M Scacco, and Natalie Jomini Stroud. 2019. The deliberative influence of comment section structure. *Journalism*, 20(6):752–771.
- Titus Plattner. 2018. *Five risks of news personalization*. Accessed: 2020-10-07.

- Kim Christian Schröder. 2019. What do news readers really want to read about? how relevance works for news audiences. *Reuters Institute for the Study of Journalism*.
- Jonathan Stray. 2021. [Beyond engagement: Aligning algorithmic recommendations with prosocial goals](#). Accessed: 2021-01-30.
- Natalie Jomini Stroud, Emily Van Duyn, and Cynthia Peacock. 2016. News commenters and news comment readers. *Engaging News Project*, pages 1–21.
- S Shyam Sundar and Sampada S Marathe. 2010. Personalization versus customization: The importance of agency, privacy, and power usage. *Human Communication Research*, 36(3):298–322.
- Neil Thurman and Steve Schifferes. 2012. The future of personalization at news websites: Lessons from a longitudinal study. *Journalism Studies*, 13(5-6):775–790.
- Yixue Wang and Nicholas Diakopoulos. 2021a. Journalistic source discovery: Supporting the identification of news sources in user generated content. In *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems*, CHI '21. Association for Computing Machinery.
- Yixue Wang and Nicholas Diakopoulos. 2021b. The role of new york times picks in comment quality and engagement. In *Proceedings of the 54th Hawaii International Conference on System Sciences*, page 2924.
- Mark Weiss. 2019. [Digiday research: 70 percent of publishers say they personalize content](#). Accessed: 2020-10-07.