Burst the Filter Bubble: Towards an Integrated Tool

Alireza Amrollahi

Australian Catholic University, alireza.amrollahi@acu.edu.au

Follow this and additional works at: https://aisel.aisnet.org/acis2019

Recommended Citation

https://aisel.aisnet.org/acis2019/2

This material is brought to you by the Australasian (ACIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in ACIS 2019 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact elibrary@aisnet.org.
Burst the Filter Bubble: Towards an Integrated Tool

Alireza Amrollahi
Peter Faber Business School
Australian Catholic University
Sydney, Australia
Email: alireza.amrollahi@acu.edu.au

Abstract
Formation of filter bubbles is known as a risk for democracy and can bring negative consequences like polarisation of the society, users' tendency to extremist viewpoints, and the proliferation of fake news. Previous studies, including prescriptive studies, focused on limited aspects of filter bubbles. The current study aims to propose a model for an integrated tool that assists users in avoiding filter bubbles in social networks. To this end, a systematic literature review has been adopted and 571 papers in six top-ranked scientific databases have been identified. After excluding irrelevant studies and an in-depth study of the remaining papers, a classification of research studies is proposed. This classification is then used to propose an overall architecture for an integrated tool that synthesises all previous studies and proposes new features for avoiding filter bubbles. The study explains the components and features of the proposed architecture and describes their focus on content and agents.

Keywords: Filter bubble, Social networks, Prescriptive study, Information bubble.
1 INTRODUCTION

The notion of filter bubble refers to the impact of our preferences and desires on the content and results we view on search engines, social media, and other online platforms. This concept has been central to social media and internet research since it was developed by Eli Pariser (Pariser 2011) and have been investigated by various scholars using various terms. Some of the terms used in the literature are information bubble (Liao and Fu 2013), online echo chamber (Möller et al. 2018), personal ecosystem of information (Helberger et al. 2015), partial information blindness (Haim et al. 2018), and information cocoons (Sunstein 2007).

Several undesirable impacts have been mentioned in the literature for filter bubbles. Especially, a potential risk to narrow the information sources for online users and "pushing users into the psychological comfort zone of self-confirmation and risking polarisation on a societal level" (Courtois et al. 2018, p. 2008) that can lead to polarisation of online debates (Seargeant and Tagg 2018) and even extremism (Costello et al. 2016; Liao and Fu 2013).

To avoid these negative impacts, several studies in the literature have recommended solutions to understand, avoid, and decrease the negative impacts of filter bubbles. These studies are focused on various topics including quantification of the bubble in social networks (Hannak et al. 2013), developing secondary apps (Wood et al. 2018), and approaches to stay anonymous in order to avoid filter bubbles (Ridgway 2017). However, none of the studies in the literature, offer a comprehensive and integrated tool to help users avoid the filter bubbles.

To tackle this shortcoming, in this study, we searched seven scientific databases with related phrases in an attempt to systematically review the prescriptive literature and suggest a conceptual model for an integrated tool. We investigated these studies based on their aim and approach to avoid filter bubble, used technology, and the effectiveness of the approach. The results of this research can help future research to find possible gaps in the literature and provide practitioners with a better understanding of the tools available to them for avoiding filter bubbles.

The remainder of this paper, in section two, provides a background to the concept of the filter bubble and posits the current study within the body of research. Section three introduces our methodology and approach for review and analysis of the literature. In section four the integrated framework is presented and results are discussed with possible implications for research and practice in section five. The paper is concluded with explaining the contributions to the body of research in section six.

2 RESEARCH BACKGROUND

2.1 Filter Bubble

Although the notion of limiting sources of information to one’s preferences has been largely studied in areas such as media (Jamieson and Cappella 2008) and psychology (Nickerson 1998), the application of this notion on online and social media came under the spotlight after the development of the term filter bubble.

There are two main research streams on the filter bubble. The first stream (inspired by the work of Pariser (2011)) is mainly focused on the impact of recommendation systems (LR et al. 2018; Nguyen et al. 2014; Sanz-Cruzado and Castells 2018). These recommendation systems consider the user’s demographic information, history, and search behaviour in suggesting new content by social media and search engines, creating a filter bubble for the information the user receives.

This stream of research has been increasingly challenged by the second wave of studies that focuses on the role of social media users rather than recommendation system technologies (Garrett 2017; Möller et al. 2018). This perspective is supported by empirical research including a study on Facebook content that found only 5-8% of the content provided to people with various political viewpoints is based on their profile (Bakshy et al. 2015).

Among previous studies, we found review research by Bozdag and van den Hoven (2015). The study considers two different perspectives about democracy (namely: liberal view of democracy and deliberative democracy) and introduces several software designs which have been introduced to combat filter bubble. The study then suggests design criteria against filter bubble based on the two models of democracy and concludes that the reviewed tools “do not define the filter bubble explicitly” and most of them “are performed for US politics” (Bozdag and van den Hoven 2015, p. 263). Except for
this comparison of tools, we did not find any research to go beyond algorithmic enhancements of recommendation systems and investigate possible social concepts leading to the formation of filter bubbles and how the system can deal with these factors.

2.2 Impacts of filter bubbles

The negative consequences associated with filter bubbles are extensively studied in the literature. Some of these impacts could be directly associated with the filter bubble. Some examples are a decline in user trust (Nagulendra and Vassileva 2016), limiting people’s access to information (Valdez et al. 2018), and social fragmentation (Möller et al. 2018).

A negative consequence that has been cited more specifically in the literature, is the polarization of political discussions in social media when people are stuck in a bubble that prevents them from receiving outsider information (Foth et al. 2016; Lahoti et al. 2018; Quaraishi et al. 2018; Thonet et al. 2017; Yang et al. 2017). Previous literature has not found a significant relationship between exposure to the opposite political view and a change in people’s political opinion (Bail et al. 2018). However, there are many studies which investigated the impact of filter bubbles on commitment to a populist cause (Postill 2018), avoidance of cross-referencing (Van den Bulck and Moe 2018), creating a risk to diversity of opinions and well-functioning democracy as a result (Bozdag and van den Hoven 2015; Dylko et al. 2018).

On the other hand, filter bubble can indirectly impact or result in proliferation of recent challenges in online media including fake news (Bhatt et al. 2018; Seargeant and Tagg 2018) as they “amplify any content, from genuine, factual news to emotionally charged, politically biased news” (Rehm 2017, p. 218). As social science studies have found that homogenous groups are more likely to become extreme in their thinking (Spohr 2017), the formation of these groups as a result of a filter bubble can lead to extremism. Finally, the negative impacts of filter bubbles have been studied in specific areas. For example, Taramigkou et al. (2013) investigated filter bubbles in music platforms and how the impacts platform users’ taste of music. Other researchers have studied the negative impacts of filter bubbles in areas such as online retail (Matt et al. 2014) and the source of information financial analysts receive (Shah et al. 2016).

Despite these breakthroughs in the literature, previous studies are less focused on the long-term of filter bubbles. For example, although there are several studies on the impact of social networks on extremism (Awan 2017; O’Callaghan et al. 2013; Spohr 2017), no empirical has investigated the impact of filter bubbles and human authority on the formation of extremist groups.

3 RESEARCH METHOD

The current paper aims to suggest an integrated tool that can deal with the problem of filter bubble in social networks. To do this a systematic literature review has been adopted and previous prescriptive studies on bursting filter bubbles have been reviewed. To conduct a systematic literature review the following steps were undertaken as suggested by Kitchenham and Charters (2007): (1) identifying resources; (2) study selection; (3) data extraction; (4) data synthesis; and (5) writing up the study as a report. To follow these steps, we searched six scientific databases: Science Direct, Scopus, ProQuest, ACM Digital Library, Association for Information Systems electronic library, and Springer Link. Table 1 shows the final set of papers in each scientific database.

<table>
<thead>
<tr>
<th>Database</th>
<th>The first set of papers</th>
<th>The final set of papers</th>
</tr>
</thead>
<tbody>
<tr>
<td>Association for Information Systems electronic library</td>
<td>99</td>
<td>2</td>
</tr>
<tr>
<td>Pro Quest</td>
<td>119</td>
<td>5</td>
</tr>
<tr>
<td>Science Direct</td>
<td>19</td>
<td>8</td>
</tr>
<tr>
<td>Scopus</td>
<td>147</td>
<td>32</td>
</tr>
<tr>
<td>Springer Link</td>
<td>146</td>
<td>4</td>
</tr>
<tr>
<td>ACM Digital Library</td>
<td>41</td>
<td>20</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>571</strong></td>
<td><strong>71</strong></td>
</tr>
</tbody>
</table>

Table 1 - Distribution of first/final set of papers in different databases

We searched for the following terms in title, keywords and abstracts depending on the services offered by the relevant search engines:
"filter bubble" or "information bubble" or "social recommendation systems", “social personalisation”, “news gatekeepers”, “information gatekeepers”, “personalised filtering”, “online echo chamber”

We found 571 papers after our initial search for the above keywords. This number was reduced to 147 papers when reading the titles and abstracts and excluded irrelevant papers. Finally, duplicated papers included in more than one database were removed and in another round, we referred to papers in full text to formulate the final pool of 71 papers. Table 2 illustrates the process through which we arrived at the final pool of research papers.

<table>
<thead>
<tr>
<th>Round</th>
<th>Number of Papers Excluded</th>
<th>Number of Papers Remaining</th>
</tr>
</thead>
<tbody>
<tr>
<td>The initial list of papers</td>
<td></td>
<td>571</td>
</tr>
<tr>
<td>Exclusion based on the title</td>
<td>107</td>
<td>464</td>
</tr>
<tr>
<td>Exclusion based on abstract</td>
<td>317</td>
<td>147</td>
</tr>
<tr>
<td>Removal of duplicate papers</td>
<td>9</td>
<td>138</td>
</tr>
<tr>
<td>Exclusion based on full text (Final list)</td>
<td>67</td>
<td>71</td>
</tr>
</tbody>
</table>

Table 2- Different Stages of Inclusion / Exclusion and Number of Papers in Each Round

For the current paper, we studied all the papers in our pool of 71 papers and focused on those which were prescriptive and suggested an approach for dealing with filter bubbles. The results of further analysis on the content of those papers are presented in the future sections. While the study focuses on the prescriptive approaches suggested in each article to prevent the formation of filter bubbles, there is less focus on the approach used to evaluate the rigour research data or data analysis. Instead, we have focused on the robustness of the proposed approach and if it can be applied in practice.

4 RESULTS

The final set of prescriptive papers formed the basis of the results described below. These final papers are studied in detail, and their advice on how to overcome the filter bubble phenomenon has been investigated. The results of this study helped us to differentiate two different perspectives in the reviewed studies on the filter bubble. Each of these perspectives is studied in further detail to understand and categorise their recommended approach.

There was a group of papers which proposed approaches to identify the filter bubble. This category includes identification of the bubble, confirming its existence, and quantification of the impact of the bubble. The second category of research studies, however, is directly focused on approaches to take users out of the filter bubble or as we call it in this paper, burst the bubble. Details about each category are explained below:

4.1 Alert about the bubble

Research studies under this category are focused on the identification and evaluation of filter bubbles. Filter identification studies are mainly focused on tools that check and alert users (in different ways) if they are trapped in a filter bubble. Nagulendra and Vassileva (2014) for example, proposed an interactive visualisation to enable the user to see the filtering. According to the authors, the tool has four goals which are: awareness, understanding, control of personalised filtering, and to increase the users’ trust in the system. The tool has implemented and tested on an independent platform.

The authors have extended this tool in their future work (Nagulendra and Vassileva 2016) to include both content and agent (users) in visualisation. This goal to detect the bubble has been continued in future studies using different approaches including diffusion of topics (content) (TK et al. 2015), network theory (agents) (Thonet et al. 2017), and machine learning (Lahoti et al. 2018).

Another stream of research on alerting filter bubble goes one step beyond bubble identification and considers evaluation and quantification of filter bubbles. Hannak et al. (2013) for example, developed a methodology for measuring personalisation in Web search results. The proposed methodology compares different search results on Google considering the attributes of the agent (user) who performs the search. Also, Matakos et al. (2017) developed an index to measure the tendency of opinion polarisation in network communities that can lead to the development of filter bubbles. This measure considers the opinion of agents (users) and the structure of the network.

Despite these studies and a few in-progress studies, no study has been used as a comprehensive framework for measuring the filter bubble. The proposed frameworks are mainly focused on agent and less focused on the content. Finally, the proposed frameworks are less applied in real-time on social networks to show the users existence or significance of a filter bubble.
4.2 Burst the bubble

This category of studies explains the suggested approach to disable or decrease the negative impact of recommendation systems creating filter bubble through exploring new ideas and diverse perspectives. The first stream of research in this category is formed around bypassing or changing algorithms. Ridgway (2017) for example, suggest staying anonymous while being online as a solution to avoid potential filter bubbles. Bozdag and van den Hoven (2015) on the other hand, reviewed possible design criteria proposed for bursting filter bubble.

Another stream of research under this category is focused on bursting the bubble by extending users’ awareness and encouraging them to explore different ideas. In this stream of research, again the focus is on either new content or new agents. The majority of studies in our final pool are focused on viewing new content. Among them, the work of Taramigkou et al. (2013) is the first study we identified that proposed a methodology helping users to explore new music genres outside their zone of interest. Webberley et al. (2016) also suggested an algorithm for avoiding filter bubble through focusing on re-tweeting behaviour of users rather the scope of the user’s social circle. Finally, several studies have suggested applications to enable users to view new content outside their preference (Linder et al. 2018; Wood et al. 2018).

Recent studies, however, have shifted the focus from content to the agent. Quraishi et al. (2018) proposed a graph partitioning method that exploits social interactions to represent different viewpoints in a social network. A qualitative evaluation of the proposed method is also presented based on implementing the method on a dataset retrieved from Twitter. Also, Sanz-Cruzado and Castells (2018) focused on contact recommendation and based on the concept of weak connection, proposed an index for diversity. Despite the diversity of research studies in this category, a lack of focus on the real-time application of the proposed methods in social networks is observable among all the studies.

5 DISCUSSION

The results of our review indicate that studying the filter bubble phenomenon is an interdisciplinary research area. Our final pool of research studies includes work from areas ranging from information systems, information technology, and management to political science, sociology, law, and journalism. In the analysis of the related literature, we could observe a focus on the content provided in social networks in few studies in various categories while many others were focused on the users posting the content (agent) and how they impact the formation of filter bubbles.

Based on the outcome of our review we propose an architecture for an integrated tool that can be implemented in social networks (regardless of the content type) and help to avoid the formation of filter bubbles. Based on the outcome of our literature review, we propose two major functions for this integrated tool: (i) alerting users about a potential filter bubble; and (ii) bursting the bubble.

Under the alert component, the integrated tool first focuses on the identification of a filter bubble. While many users of social networks are not aware of the filter bubble they were kept in as a result of a lack of transparency in the algorithms used by social networks (Bozdag and Timmermans 2011), the tool needs to alert users. As the cause of filter bubbles shift from the recommendation algorithms to features enabling users to put themselves in filter bubble (Amrollahi and McBride 2019), this is important to also inform the users about the consequences of their actions. For example, users should have the right to see how blocking or muting one specific user, may result in missing out a network of users and their perspective. Therefore, this bubble identification feature should be designed in connection with a bubble evaluation feature that assesses the significance of the bubble and also can predict future significance after certain events.

Potential improvements in the recommendation systems have been suggested as a possible solution for bursting the bubble. However, in the current study we have not considered this as a feature in the proposed tool. Recommendation systems are not included in the tool for two reasons: (i) the proposed tool is proposed independent to the social network, type of content they provide, and the recommendation algorithm they use; (ii) the proposed integrated tool is focused on not only the recommendation algorithm, but also on social network facilities that enable users to build a bubble around them. The architecture of the proposed integrated tool is illustrated in Figure 1.

As illustrated in Figure 1, the proposed tool, will focus on both agent and content. It means that in order to identify filter bubbles (as part of the alert component), both connections of and the content viewed by the user under study should be investigated. Also, the significance of the bubble should be evaluated by considering both the recommended content and recommended users. Finally, the
awareness of users should be increased by suggesting both novel (out of bubble) content and connections.

Figure 1 The proposed architecture for an integrated tool

6 CONCLUSION

Filter bubbles are problematic consequences of modern media and social networks as they create barriers to rational and diversified dialogue that is necessary for a democratic society. This research looks at filter bubble as not only a product of recommendation algorithms in social networks but also a social issue which considers an active role for users building a filter bubble around them.

In the current paper, we proposed an architecture for an integrated tool that can be incorporated into social networks to prevent the formation of filter bubbles around users. This tool is proposed based on a systematic review of the literature and classification of the studies under different categories considering their aim. The proposed components of the integrated tool cover both liberal (through alert component) and deliberative (through awareness component) models of democracy (Bozdag and van den Hoven 2015).

The results of our review show a lack of empirical studies on the effectiveness of the proposed tools. Even those studies which included their empirical results, did so by using test data on developed hypothetical platforms. Therefore, the application of the proposed methods in social networks should be considered in future studies. This improvement in future research will lead to a better evaluation of the effectiveness of the proposed methods which is another shortcoming we identified in the literature.

Based on the results of our literature review, the components of an integrated tool are proposed in the form of an architecture map. The proposed method, unlike what is proposed in the previous literature, considers different perspectives on filter bubbles. These perspectives include a concurrent focus on alerting users about the formation of bubble and bursting bubbles through increasing users’ awareness. The proposed integrated tool also has a dual focus on both content and agent which cannot be found in the previous literature.

Although the proposed framework in this study is expected to consider various approaches for busting the filter bubble in the literature, it has not been put to practice as a solution in any social network.
Therefore, we propose future study to consider this as a principle of implementation and evaluation of the system. In particular, the results of the current study can inform future studies especially upcoming design research. Moreover, there is currently no evidence of effectiveness of the proposed components of the system as a whole. Future studies should focus on the effectiveness of the proposed approach and check if people in such filter bubbles will break free when given an opportunity. Future studies can also focus on different parts of the proposed tool and develop algorithms for each part. The study also proposes a real-time (predictive) bubble evaluation which is not developed in the previous studies. Finally, the results of this study can benefit practitioners and managers of social networks to see various avenues for improving their service or possible pitfalls in their social network allowing the formation of filter bubbles.

7 REFERENCES


Copyright: © 2019 Amrollahi. This is an open-access article distributed under the terms of the Creative Commons Attribution-NonCommercial 3.0 Australia License, which permits non-commercial use, distribution, and reproduction in any medium, provided the original author and ACIS are credited.