HOW TO BURST THE BUBBLE IN SOCIAL NETWORKS?

Abstract

Filter bubble has considered as a serious risk for democracy and freedom of information on the internet and social media. This phenomenon can restrict users' access to information sources outside their comfort zone and increase the risk of polarisation of opinions on different topics. This in-progress paper explains our plan for conducting a prescriptive research aiming at decreasing the chance of filter bubbles formation on social networks. The paper explains a gap in the literature which is a prescriptive work considering both human and technology perspectives. To focus on this research gap, a design perspective has been selected covering two different bodies of theory as kernel theories. The paper explains the relevance of these theories, some of the primarily formed requirements derived from them and the future steps in this research. The explained future steps includes various phases of developing an Information Systems Design Theory and our strategy to evaluate the effectiveness of the developed theory.

Keywords: Filter Bubble, Information Bubble, Design Theory, Habermas Theory of Discourse, Attitude Polarisation.

1.0   Introduction

The concept of *filter bubble* (also known as *information bubble*) was first coined by the internet activist, Eli Pariser in a book with the same title (Pariser, 2011). This concept 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. Significant attention in both academia and industry has been attracted to this notion since its development. In particular, the 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" have been mentioned in the literature (Courtois, Slechten, and Coenen, 2018, p. 2008).

Earlier literature on filter bubble, are mainly focused on the role of recommendation systems and how understanding users' information and preferences may impact the results they view on search engines (Hannak et al., 2013; Ridgway, 2017; Tran and Yerbury, 2015). Also, our review of the literature indicates that previous research work is either descriptive (Bozdag, Gao, Houben, and Warnier, 2014; Courtois et al., 2018; Matt, Benlian, Hess, and Weiß, 2014) or focused on technical improvements of related recommendation algorithms (Apel, Yom-Tov, and Tennenholtz, 2018; Knijnenburg, Sivakumar, and Wilkinson, 2016; Nguyen, Hui, Harper, Terveen, and Konstan, 2014).

This study, however, undertakes a different perspective and focuses on the reciprocal role of human and technology in creating such a bubble. The study also, takes a design perspective to prescribe an Information Systems Design Theory (ISDT) decreasing the chance of forming a filter bubble for users of social networks. To do this, the study refers to various theoretical bodies of research as kernel theories; to explore the role of technology, the study focuses on Habermas Theory of Discourse (Habermas, 2005), and to discover human behaviour, Attitude Polarisation (Corner, Whitmarsh, and Xenias, 2012) has been selected as a theoretical basis. The remainder of this paper, in sections 2, reviews the related literature and tries to position the current study within the body of research. Section 3 explains the used research methods including the kernel theories and how we are going to incorporate those theoretical concepts in our design research. Section 4 explains our primarily outcomes including a number of design requirements and explains our plan for future research in forming and evaluating the design theory.

2.0   Filter bubble

Shortly after the development of the term filter bubble, the concept found its way to academic research. Initial research in this area was mainly focused on verifying the existence of the filter bubble (Hannak et al., 2013) and its impact (Forsblom, Nurmi, Åman, and Liikkanen, 2012). Several negative impacts are associated with filter bubbles in the literature. Taramigkou, Bothos, Christidis, Apostolou, and Mentzas (2013) mentioned that developing filter bubbles in music platforms makes it difficult for users to go out of their personalised world and change their taste and opinion. Also, proliferation of fake news has been considered as a possible consequence of filter bubbles (Bhatt, Joglekar, Bano, and Sastry, 2018; Seargeant and Tagg, 2018) as this make entrance of new information hard into the developed bubble. Other potential negative impacts include decline in user trust (Nagulendra and Vassileva, 2016), limiting people's access to information (Valdez, Kluge, and Ziefle, 2018), and social fragmentation (Möller, Trilling, Helberger, and van Es, 2018).

In particular, polarisation of political discussions in social media has been cited as a major consequence which may happen when people are trapped in a bubble that prevents them from receiving outsider information (Foth, Tomitsch, Forlano, Haeusler, and Satchell, 2016; Lahoti, Garimella, and Gionis, 2018; Quraishi, Fafalios,

and Herder, 2018; Thonet, Cabanac, Boughanem, and Pinel-Sauvagnat, 2017). Network studies (Kelly and Francios, 2018) illustrate how Twitter groups do not extend political discourse but isolate factions in self-confirming chatter. Although exposure to opposite political view is not approved to significantly impact (or change) people's political opinion (Bail et al., 2018), this could be a risk to diversity of opinions; and well-functioning democracy as a result (Bozdag and van den Hoven, 2015; Dylko et al., 2018).

The initial work on filter bubble mainly focuses on the impact of recommendation systems (LR, Tamhane, and Pervin, 2018; Nguyen et al., 2014; Sanz-Cruzado and Castells, 2018). This perspective, which is similar to the initial description of Pariser (2011) about the topic, considers the use of user's demographic information, history, and search behaviour in suggesting new content by social media and search engines, as the main factor which creates filter bubbles.

However, this perspective has been increasingly challenged by recent research studies in this area (Garrett, 2017). For example, a study on Facebook content found that only 5-8% of the content provided to people with various political viewpoints is based on their profile (E. Bakshy, S. Messing, and L. A. Adamic, 2015). Companies such as Facebook and Google have also claimed to improve their algorithms to avoid the impacts of filter bubble (E. Bakshy, S. Messing, and L. Adamic, 2015; Hao, 2018).

On the other hand, recent studies are more focused on the role of social media users (rather than technologies). For example, the study of Möller et al. (2018) shows a match between news recommendation systems and journalistic recommendations. The study concludes that future research should focus on factors other than recommendation algorithms to achieve diversity. Following this call, in the current study we shift the focus to human aspects of social networks and how the provided features in the social networks enable users to create a filter bubble around them.

3.0   Research Method

The current research will use a process model for design science research suggested by Peffers, Tuunanen, Rothenberger, and Chatterjee (2007). The research design process starts with *problem and motive identification* and continues with *defining objectives of a solution*. In this research a solution will be converted to an artefact during the *design and development* phase.

Considering the prescriptive nature of this study a developed Information Systems Design Theory (ISDT) is selected as the output. This ISDT will entail the structure and function of an information system and a process facilitating the implementation of the first component which can respectively represented by the principles of form and function and principles of implementation in ISDT according to Gregor and Jones (2007).

Requirements in ISDT are governed by core theories from natural or social science areas known as kernel theories (Walls, Widermeyer, and El Sawy, 2004). As explained before, the twofold nature of this study in considering both human and technology aspects of the filter bubble, will be reflected in the selected kernel theories. Habermas theory of discourse has been set as *product kernel theory* to explain the architecture of the required information systems and Attitude Polarisation is selected as *process kernel theory* to explain the behavioural aspects.

3.1 Habermas Theory of Discourse

Habermas identifies discourse as "a processes of argumentation and dialogue in which the claims implicit in the speech act are tested for their rational justifiability as true, correct or authentic" and can coordinate human actions (Bohman and Rehg, 2007, p. Section 3.1). In the current study, the ideal (free from filter bubble) social network is considered as an ideal form of discourse. Aier, Fischer, and Winter (2011) interpretation of discourse theory identifies four pragmatic presuppositions of an ideal discourse which are listed below:

- No-one capable of making a relevant contribution has been excluded.
- Participants have equal voice.
- Participants are internally free to speak their honest opinion without deception or self-deception.
- There is no source of coercion built into the process and procedures of discourse.

Considering the above presuppositions, in the study will help the research to form the requirements in a way that they decrease the gap between the developed system and an ideal discourse.

3.2 Attitude Polarisation

According to Corner et al. (2012, p. 6), attitude polarisation refers to the fact that "having assimilated information in a biased way, people with opposing attitudes may diverge in their opinions". The reason this has been selected as a kernel theory for the

current study is to consider the most significant impact of filtered information in the social network and how it can lead to polarisation of people on various aspects.

Previous studies has investigated the impact of attitude polarisation on social networks like Facebook (Howarth and Sharman, 2015) and Twitter (Pearce, Holmberg, Hellsten, and Nerlich, 2014) and how these social networks are divided on topics such as climate change and politics.

In this study, we will particularly consider the interpretation of Parsell (2008) about the concept of attitude polarisation in the context of internet and web 2.0:

- People seek out others with the same prejudices as themselves;
- The Web 2.0 provides the necessary resources to build communities with whomever we like;
- Being in a community of people with the same prejudices increases our own prejudices;
- Hence, the Web 2.0 is likely to lead to greater prejudice, social cleavage and community division.

The developed requirements for the ISDT are set to help avoid these "worrying argument".

4.0   Future work

4.1 Design theory

The developed ISDT in this study will be a macro-level design theory that satisfies a number of design requirements for avoiding filter bubble in social networks. These requirements will be derived from kernel theories and include the following:

- Informing users about the concept of filter bubble on social networks and its negative impacts.
- Informing users from potential and actual filter bubbles in the content they view.
- Facilitating exploration of different (and opposing) viewpoints on certain topics (i.e. hashtags).
- Reminding users about the possibility of forming an filter bubble when they block or mute users outside their comfort zone.

By referring to two main references of ISDT, Table 1 explains how requirements of an ISDT will be met in the current study.

Requirements of ISDT		How the requirement will be met in the current study
Elements of ISDT (Walls, Widmeyer, and El Sawy, 1992)	Anatomical skeleton of design theory (Gregor and Jones, 2007)	
Product kernel theory	Justificatory knowledge	Habermas' theory of discourse will be used to form the technology related requirements
Process kernel theory		Attitude Polarisation will be used to form the human related requirements
Meta-requirements	Purpose and scope	Defined in the introduction section of the paper as forming a number of design

Requirements of ISDT		How the requirement will be met in the current study
Elements of ISDT (Walls, Widmeyer, and El Sawy, 1992)	Anatomical skeleton of design theory (Gregor and Jones, 2007)	
		principles in social networks to avoid filter bubbles
	Constructs	A main focus on psychological and technical factors leading to form a filter bubble
Meta-description	Principle of form and function	A number of technology focused design principles for developing a social network free of filter bubbles
	Artefact mutability	A number of human focused design principles for developing a social network free of filter bubbles
Product hypotheses	Testable propositions	A number of propositions on the effectiveness of the developed principles
Process hypotheses		
Design method	Principles of implementation	An overall design and architecture of the designed system as well as guidelines on how to implement them in social network.
	Expository instantiation	

Table 1.   Components of design theory in the current study

4.2 Implementation and evaluation

Artefact evaluation has been considered as a crucial phase in design research (Hevner, March, Park, and Ram, 2004; Peffers et al., 2007). For this reason, the developed design principles will be implemented and evaluated in a social network. Both *ex ante* and *ex post* evaluations will be subject of attention in this study. Ex ante evaluation is the predictive evaluation of the design effort with respect to the future impacts. This will be done through surveying experts in areas such as social networks and ethics about the developed ISDT. On the other hand, ex post evaluation evaluates functional value of a developed artefact (Neff et al., 2014; Stefanou, 2001). Ex post evaluation in this study will be through implementation of the principles in real social network environment and formally evaluate their impact in avoiding filter bubbles.

5. Conclusions

Filter bubbles are problematic because they create barriers to rational discussion and the dialogue that is necessary for a democratic society. This research considers how the application of design science techniques in information systems might contribute to enable social media users to increase their awareness of filter bubbles and seek to avoid being trapped in them. However, this is a social issue because no-one has to be trapped in a filter bubble. Users can seek out different opinions, join different political

Twitter threads, and contribute to dialogs. And yet often users do not challenge their own social media tastes.

In exploring this area and considering both discourse and polarisation, we will need to query the motives which result in people remaining in filter bubbles. For some, perhaps, restriction to a filter bubble, is motivated by their perception of their identity within a group and their wish to remain in the perceived safety of a social bubble. When some move to other social groupings in, for example, Twitter, there may be a propensity, rather than entering to the social discourse envisioned by Habermas, to participating in the trolling, sarcasm and denigration that is so prevalent in social networks,

References

Aier, S., Fischer, C., & Winter, R. (2011). "Theoretical stability of information systems design theory evaluations based upon Habermas's discourse theory". *Paper presented at the 19th European Conference on Information Systems (ECIS 2011)*, Helsinki, Finland.

Apel, R., Yom-Tov, E., & Tennenholtz, M. (2018). *Characterizing Efficient Referrals in Social Networks*. arXiv preprint arXiv:1805.00252.

Bail, C. A., Argyle, L. P., Brown, T. W., Bumpus, J. P., Chen, H., Hunzaker, M. F., . . . Volfovsky, A. (2018). *Exposure to opposing views on social media can increase political polarization*. Proceedings of the National Academy of Sciences*, 115*(37), 9216-9221.

Bakshy, E., Messing, S., & Adamic, L. (2015). *Exposure to diverse information on facebook*. Facebook Research Blog.

Bakshy, E., Messing, S., & Adamic, L. A. (2015). *Exposure to ideologically diverse news and opinion on Facebook*. Science*, 348*(6239), 1130-1132.

Bhatt, S., Joglekar, S., Bano, S., & Sastry, N. (2018). *Illuminating an Ecosystem of Partisan Websites*. arXiv preprint arXiv:1803.03576.

Bohman, J., & Rehg, W. (2007). Jürgen habermas. *Stanford Encyclopedia of Philosophy.* Retrieved from http://stanford.library.usyd.edu.au/archives/spr2012/entries/habermas/

Bozdag, E., Gao, Q., Houben, G.-J., & Warnier, M. (2014). *Does offline political segregation affect the filter bubble? An empirical analysis of information diversity for Dutch and Turkish Twitter users*. Computers in human behavior*, 41*, 405-415.

Bozdag, E., & van den Hoven, J. (2015). *Breaking the filter bubble: democracy and design*. Ethics and Information Technology*, 17*(4), 249-265.

Corner, A., Whitmarsh, L., & Xenias, D. (2012). *Uncertainty, scepticism and attitudes towards climate change: biased assimilation and attitude polarisation*. Climatic change*, 114*(3-4), 463-478.

Courtois, C., Slechten, L., & Coenen, L. (2018). *Challenging Google Search filter bubbles in social and political information: Disconforming evidence from a digital methods case study*. Telematics and Informatics*, 35*(7), 2006-2015.

Dylko, I., Dolgov, I., Hoffman, W., Eckhart, N., Molina, M., & Aaziz, O. (2018). *Impact of customizability technology on political polarization*. Journal of Information Technology & Politics*, 15*(1), 19-33.

Forsblom, A., Nurmi, P., Åman, P., & Liikkanen, L. (2012). "Out of the bubble: Serendipitous even recommendations at an urban music festival". *Paper presented at the Proceedings of the 2012 ACM international conference on Intelligent User Interfaces*.

Foth, M., Tomitsch, M., Forlano, L., Haeusler, M. H., & Satchell, C. (2016). "Citizens breaking out of filter bubbles: urban screens as civic media". *Paper presented at the Proceedings of the 5th ACM International Symposium on Pervasive Displays*.

Garrett, R. K. (2017). *The "echo chamber" distraction: Disinformation campaigns are the problem, not audience fragmentation*. Journal of Applied Research in Memory and Cognition*, 6*(4), 370-376.

Gregor, S., & Jones, D. (2007). *The anatomy of a design theory*. Journal of the Association for Information systems*, 8*(5), 312.

Habermas, J. (2005). *Zwischen naturalismus und religion*: Suhrkamp Frankfurt am Main.

Hannak, A., Sapiezynski, P., Molavi Kakhki, A., Krishnamurthy, B., Lazer, D., Mislove, A., & Wilson, C. (2013). "Measuring personalization of web search". *Paper presented at the Proceedings of the 22nd international conference on World Wide Web*.

Hao, K. (2018). Google is finally admitting it has a filter-bubble problem. Retrieved from https://qz.com/1194566/google-is-finally-admitting-it-has-a-filter-bubble-problem/

Hevner, A. R., March, S. T., Park, J., & Ram, S. (2004). *Design science in information systems research*. MIS Quarterly*, 28*(1), 75-105.

Howarth, C. C., & Sharman, A. G. (2015). *Labeling opinions in the climate debate: A critical review*. Wiley Interdisciplinary Reviews: Climate Change*, 6*(2), 239-254.

Kelly, J and Francios, C. (2018) This is what filter buble actually look like. MIT Technology Review August 22, 2018. https://www.technologyreview.com/s/611807/this-is-what-filter-bubbles-actually-look-like/ Accessed 7/1/2019

Knijnenburg, B. P., Sivakumar, S., & Wilkinson, D. (2016). "Recommender systems for self-actualization". *Paper presented at the Proceedings of the 10th ACM Conference on Recommender Systems*.

Lahoti, P., Garimella, K., & Gionis, A. (2018). "Joint non-negative matrix factorization for learning ideological leaning on Twitter". *Paper presented at the Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining*.

LR, D., Tamhane, A., & Pervin, N. (2018). *A Clustering Based Social Matrix Factorization Technique for Personalized Recommender Systems*.

Matt, C., Benlian, A., Hess, T., & Weiß, C. (2014). *Escaping from the filter bubble? The effects of novelty and serendipity on users' evaluations of online recommendations*.

Möller, J., Trilling, D., Helberger, N., & van Es, B. (2018). *Do not blame it on the algorithm: An empirical assessment of multiple recommender systems and their impact on content diversity*. Information, Communication & Society*, 21*(7), 959-977.

Nagulendra, S., & Vassileva, J. (2016). *Providing awareness, explanation and control of personalized filtering in a social networking site*. Information Systems Frontiers*, 18*(1), 145-158.

Neff, A. A., Hamel, F., Herz, T. P., Uebernickel, F., Brenner, W., & vom Brocke, J. (2014). *Developing a maturity model for service systems in heavy equipment manufacturing enterprises*. Information & Management*, 51*(7), 895-911.

Nguyen, T. T., Hui, P.-M., Harper, F. M., Terveen, L., & Konstan, J. A. (2014). "Exploring the filter bubble: the effect of using recommender systems on content diversity". *Paper presented at the Proceedings of the 23rd international conference on World wide web*.

Pariser, E. (2011). *The filter bubble: What the Internet is hiding from you*: Penguin UK.

Parsell, M. (2008). *Pernicious virtual communities: Identity, polarisation and the Web 2.0*. Ethics and Information Technology*, 10*(1), 41-56.

Pearce, W., Holmberg, K., Hellsten, I., & Nerlich, B. (2014). *Climate change on Twitter: Topics, communities and conversations about the 2013 IPCC Working Group 1 report*. PloS one*, 9*(4), e94785.

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). *A design science research methodology for information systems research*. Journal of Management Information Systems*, 24*(3), 45-77.

Quraishi, M., Fafalios, P., & Herder, E. (2018). "Viewpoint Discovery and Understanding in Social Networks". *Paper presented at the Proceedings of the 10th ACM Conference on Web Science*.

Ridgway, R. (2017). *Against a personalisation of the self*. ephemera: theory & politics in organization*, 17*(2).

Sanz-Cruzado, J., & Castells, P. (2018). "Enhancing structural diversity in social networks by recommending weak ties". *Paper presented at the Proceedings of the 12th ACM Conference on Recommender Systems*.

Seargeant, P., & Tagg, C. (2018). *Social media and the future of open debate: A user-oriented approach to Facebook's filter bubble conundrum*. Discourse, Context & Media.

Stefanou, C. J. (2001). *A framework for the ex-ante evaluation of ERP software*. European Journal of Information Systems*, 10*(4), 204-215.

Taramigkou, M., Bothos, E., Christidis, K., Apostolou, D., & Mentzas, G. (2013). "Escape the bubble: Guided exploration of music preferences for serendipity and novelty". *Paper presented at the Proceedings of the 7th ACM conference on Recommender systems*.

Thonet, T., Cabanac, G., Boughanem, M., & Pinel-Sauvagnat, K. (2017). "Users Are Known by the Company They Keep: Topic Models for Viewpoint Discovery in Social Networks". *Paper presented at the Proceedings of the 2017 ACM on Conference on Information and Knowledge Management*.

Tran, T., & Yerbury, H. (2015). *New Perspectives on Personalised Search Results: Expertise and Institutionalisation*. Australian Academic & Research Libraries*, 46*(4), 277-290.

Valdez, A. C., Kluge, J., & Ziefle, M. (2018). *Elitism, trust, opinion leadership and politics in social protests in Germany*. Energy Research & Social Science.

Walls, J. G., Widermeyer, G. R., & El Sawy, O. A. (2004). *Assessing information system design theory in perspective: how useful was our 1992 initial rendition?* Journal of Information Technology Theory and Application (JITTA)*, 6*(2), 43-58.

Walls, J. G., Widmeyer, G. R., & El Sawy, O. A. (1992). *Building an information system design theory for vigilant EIS*. Information Systems Research, *3*(1), 36-59.