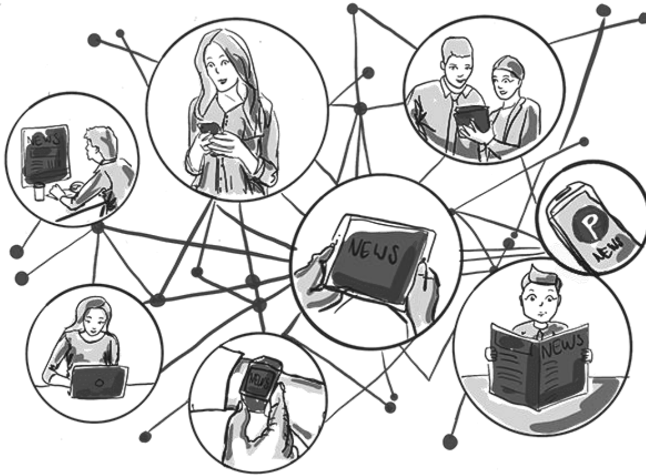


Being diverse is not enough:
Rethinking Diversity Evaluation to Meet Challenges of News Recommender Systems (NRS)

Célina Treuillier - Sylvain Castagnos - Evan Dufraisse - Armelle Brun

BOOM
MODELING & OPENING OPINION BUBBLES
Loria
BIRD
UNIVERSITÉ DE LORRAINE

1

Introduction

- Current challenges of NRS
- Diversity and NRS
- Evaluation of NRS

2

Experimental analysis

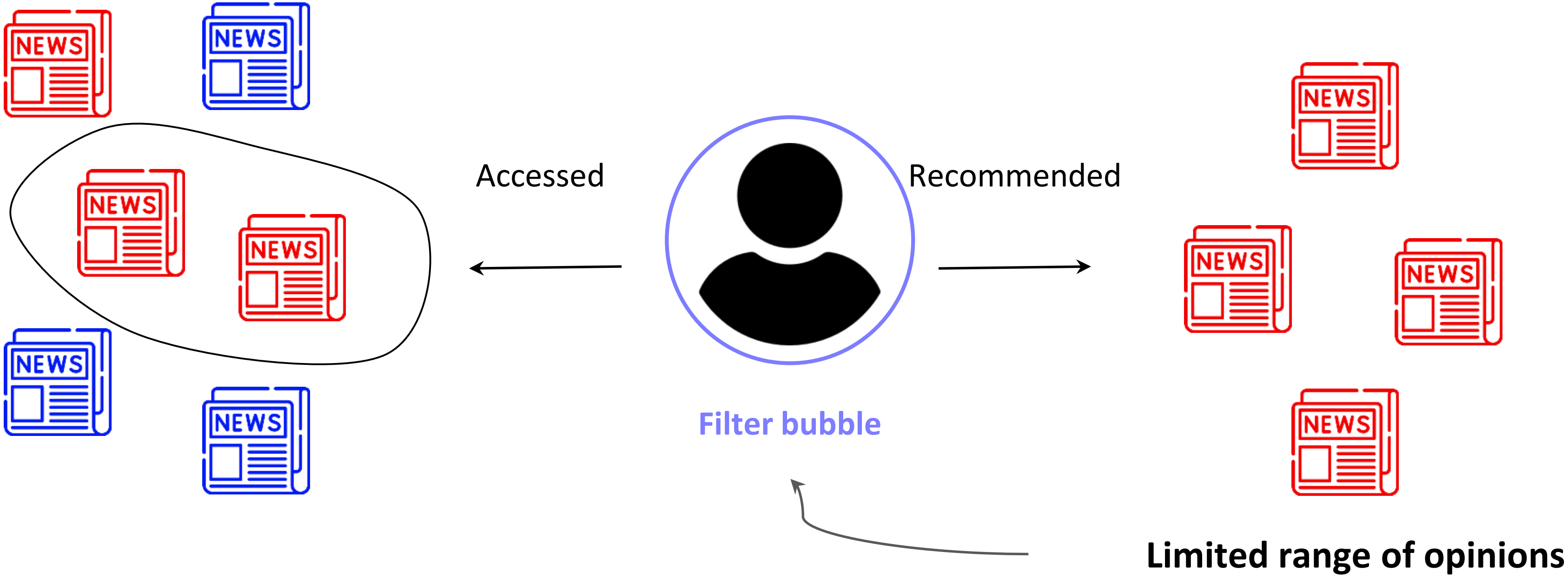
- Dataset selection
- Holistic analysis
- Temporal analysis

3

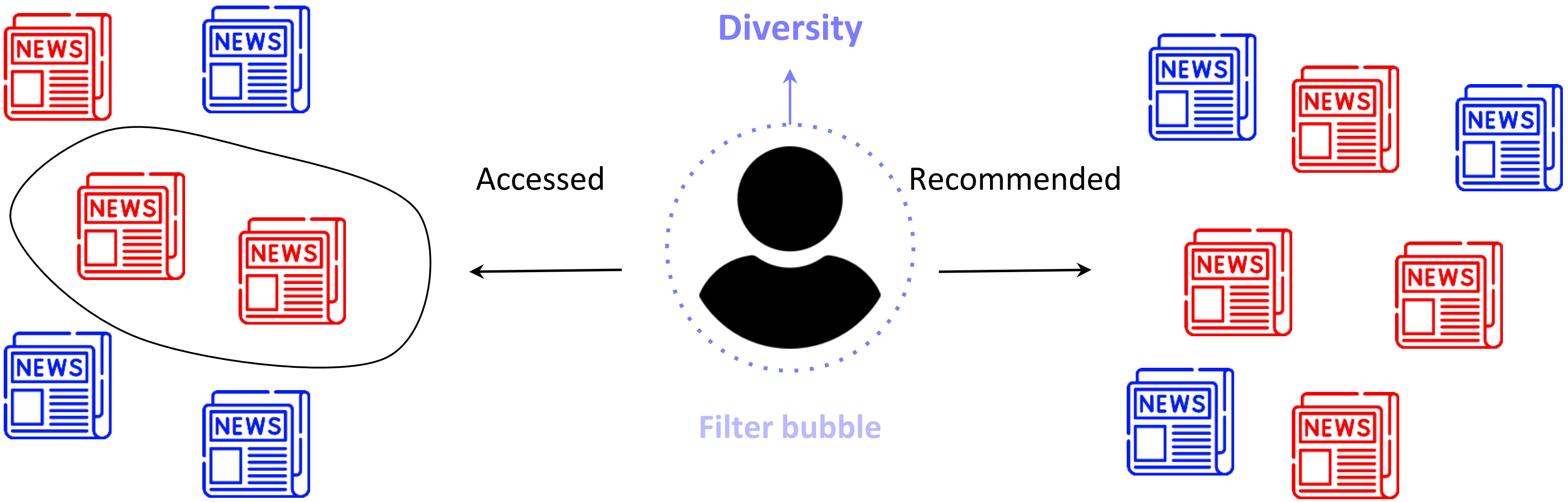
Take-home messages

Introduction

News Recommender Systems



News Recommender Systems



Diversity & NRS

- Provide a larger spectrum of opinions *(Heitz et al., 2022)*
- Ensure ethical and fair recommendations *(Lunardi et al., 2020)*
- Foster a healthy democratic debate *(Giunchiglia et al., 2021)*

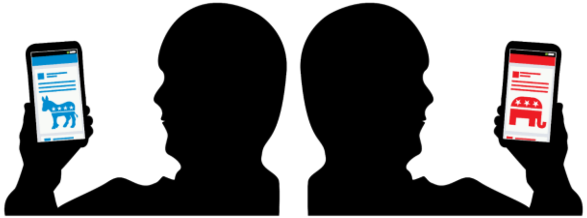
Role of diversity may be overestimated and must be finely controlled to help reduce polarization

⇒ **RQ1:** Does diversity of recommendations bring a systematic gain?

Evaluation of NRS

News characteristics: short lifespan, high turnover…

Democratic role of NRS *(Helberger, 2019)*



Adaptation of recommendation and evaluation

⇒ **RQ2:** Is it sufficient to measure the influence of an NRS afterward with single-number metrics, or does this influence occur with some variations over time?

Experimental analysis

NEWS MIND

Large-scale dataset for news recommendation research

Information about news

Users' interactions

5 weeks : October to November, 2019

Data selection ⇒1,475 users & 20,541 news

From URL
- News representation → ***LDA*** (Latent Dirichlet Allocation)
- Date of publication

User ID

History — *Week 1 to 4*

Recommended
Accessed — *Week 5*

MIND dataset

Average diversity *(Smyth & McClave, 2001) :*

$$Diversity(i_1, i_2, ..., i_n) = \frac{\sum_{k=1}^{n} \sum_{j=1}^{n} (1 - Similarity(i_k, i_j))}{\frac{n}{2} * (n-1)}$$

1. Holistic analysis

News of history

Recommended news

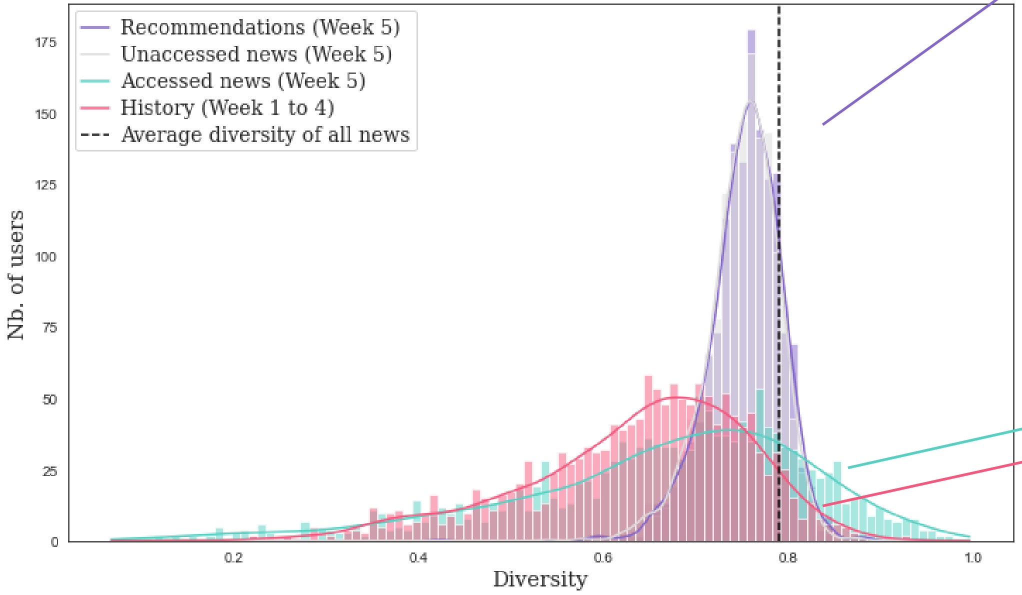
Accessed news

Unaccessed news

2. Temporal analysis

News accessed each week (1 to 5)

Analysis - *holistic*



Distribution of diversity among users

High average diversity *(0.75)* & small standard deviation *(0.04)*

Average diversity significantly lower

Great variability

Analysis - *holistic*

Conclusions:

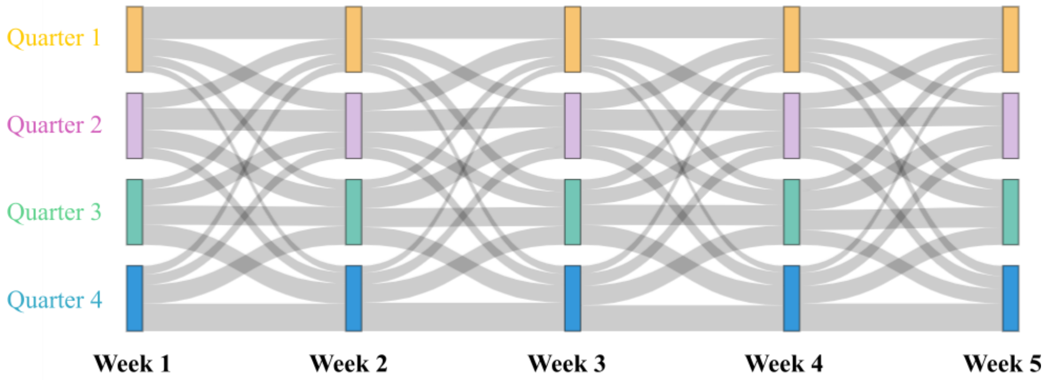
- Recommender system meets a predefined diversity level
- Impact of NRS on the news consumption
- Diversity input is not personalized

⇒ High diversity of recommendations does not systematically lead to a diverse news consumption **(RQ1)**

Analysis - *temporal*

Transition patterns remain stable over the weeks

⇒ Similar global impact of recommendations on users diversity of accessed news through weeks



Sankey diagram : variation flows over weeks

Analysis - *temporal*

Conclusions:

- Average diversity of accessed news differs over weeks
- Users observe diversity variations
- Impact is not the same for all users

3 types of users:

- Positively receptive users
- Negatively receptive users
- Resistant users

⇒ Single-number evaluations are insufficient

Need to take differences between users and temporal aspect into account **(RQ2)**

Overall conclusion

1. Diverse recommendations $\not\Rightarrow$ diverse consumption

   Need to adapt the evaluation

2. NRS does not impact all users equally

   Different classes of user behavior

Take-home messages

Take-home messages

1. Need of well-established methodologies to model users' diversity trajectories
2. Need of adapted diversity measures and personalized recommendation strategies
3. Lack of open datasets

Thank you - Thank you - Thank you - Thank you

Contact me!

celina.treuillier@loria.fr

References

Bernstein, A., de Vreese, C., Helberger, N., Schulz, W., Zweig, K., Baden, C., Beam, M. A., Hauer, M. P., Heitz, L., Katzenbach, C., Kille, B., Loosen, W., Moeller, J., Shani, G., Tintarev, N., Tolmeijer, S., van Atteveldt, W., Vrijenhoek, S., & Zueger, T. (2020). Diversity in News Recommendation. 19.

Ekstrand, M. D., Harper, F. M., Willemsen, M. C., & Konstan, J. A. (2014, October). User perception of differences in recommender algorithms. In Proceedings of the 8th ACM Conference on Recommender systems (pp. 161-168).

Giunchiglia, F., Otterbacher, J., Kleanthous, S., Batsuren, K., Bogin, V., Kuflik, T., & Tal, A. S. (2021). Towards Algorithmic Transparency: A Diversity Perspective.

Helberger, N. (2019). On the democratic role of news recommenders. Digital Journalism, 7(8), 993-1012.

Jones, N. (2010). User perceived qualities and acceptance of recommender systems (No. THESIS). EPFL.

Kozyreva, Anastasia, Stephan Lewandowsky, et Ralph Hertwig. 2020. « Citizens Versus the Internet: Confronting Digital Challenges With Cognitive Tools ». Psychological Science in the Public Interest 21(3):103-56. doi: 10.1177/1529100620946707.

References

Lunardi, G. M., Machado, G. M., Maran, V., & de Oliveira, J. P. M. (2020). A metric for Filter Bubble measurement in recommender algorithms considering the news domain. Applied Soft Computing, 97, 106771. https://doi.org/10.1016/j.asoc.2020.106771

Ricci, F., Rokach, L., & Shapira, B. (2011). Introduction to recommender systems handbook. In Recommender systems handbook (pp. 1-35). Springer, Boston, MA.

Wu, F., Qiao, Y., Chen, J. H., Wu, C., Qi, T., Lian, J., ... & Zhou, M. (2020, July). Mind: A large-scale dataset for news recommendation. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics (pp. 3597-3606).

Zhang, M., & Hurley, N. (2008, October). Avoiding monotony: improving the diversity of recommendation lists. In Proceedings of the 2008 ACM conference on Recommender systems (pp. 123-130).

Ziegler, Cai-Nicolas, Sean M. McNee, Joseph A. Konstan, et Georg Lausen. 2005. « Improving Recommendation Lists through Topic Diversification ». P. 22 in Proceedings of the 14th international conference on World Wide Web - WWW ’05. Chiba, Japan: ACM Press.