

and decayed every 10 rounds. Convergence is determined after a maximal number of 3000 rounds. The reported results are averaged over 5-fold validation.

We can see in Fig.5 that MCP performs consistently best in all NDCGs. It boosts the performance for about 5% than the best of MNAR models, i.e. MF-MNAR. This result demonstrates the competency of our model. Furthermore, it is worth-noting that the persona specific strength parameter learnt for MCP model $\tau_1 = 2$ for non-hardcore users and $\tau_0 = 0.4$ for hardcore users. The interpretation for this value is that, for the same rating that falls in the minority opinion with high divergent $|X_{ij} - E_{ij}|$, a hardcore user is more likely to display the rating than a non-hardcore user. This result is consistent with the empirical findings.

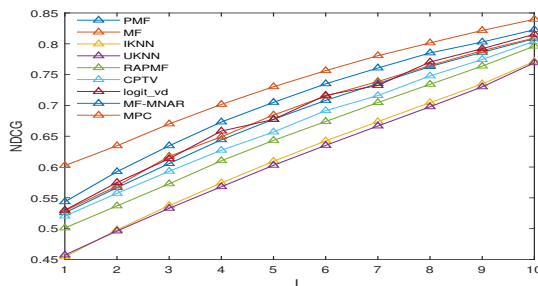


Figure 5: Comparable NDCG performance at top L items

8 CONCLUSION

In this paper we bring a social science perspective to the empirical study of missing not at random ratings in recommender systems. We verify the spiral of silence theory in large-scale real recommendation systems. We study the factors which contribute to the formation of the spiral of silence, i.e. the existence of hardcore users and the characteristics of a hardcore person. Our findings not only reveal that ratings in recommender systems are not missing at random, but also capture the mechanism of missing ratings. To demonstrate the impact of our empirical findings, we use the findings to guide the developments of a MNAR recommendation model. We experimentally show that such a model outperforms state-of-the-art models with and without MNAR assumptions.

In the future, we will also use the findings to model the evolution of public opinions and peer groups.

ACKNOWLEDGMENTS

Chen Lin is supported by the Natural Science Foundation of China (no.61472335). Yanghua Xiao is supported by National Key R&D Program of China under No.2017YFC0803700 and No.2017YFC1201200, by the National NSFC (No.61732004, No.61472085, No.U1509213, No.U1636207), by Shanghai Municipal Science and Technology project (No.16511102102, No.16JC1420401), Shanghai STCSMs R&D Program under Grant (No.16JC1420400). Chen Lin is the corresponding author.

REFERENCES

- [1] B. M. Marlin and R. S. Zemel. Collaborative prediction and ranking with non-random missing data. In *RecSys 2009*, pages 5–12.

- [2] H. Yang, G. Ling, Y. Su and MR. Lyu. Boosting response aware model-based collaborative filtering. *IEEE Transactions on Knowledge and Data Engineering*, 27(8):2064–2077, 2015.
- [3] YD. Kim and S. Choi. Bayesian binomial mixture model for collaborative prediction with non-random missing data. In *RecSys 2014*, pages 201–208.
- [4] JM. Hernández-Lobato, N. Houlisby and Z. Ghahramani. Probabilistic matrix factorization with non-random missing data. In *ICML 2014*, pages 1512–1520.
- [5] P. Popalan, JM. Hofman and DM. Blei. Scalable Recommendation with Hierarchical Poisson Factorization. In *UAI 2015*, pages 326–335.
- [6] D. Liang, L. Charlin, J. McInerney and DM. Blei. Modeling user exposure in recommendation. In *WWW 2016*, pages 951–961.
- [7] E. Noelle-Neumann. The spiral of silence a theory of public opinion. *Journal of communication*, 24(2):43–51, 1974.
- [8] J. D. Kenner. Self-serving bibias in perceiving the opinions of others. *Communication Research*, 17:393–404, 1990.
- [9] J. Matthes. Observing the “spiral” in the spiral of silence. *International Journal of Public Opinion Research*, 27(2):155–176, 2014.
- [10] RE. Anderson. Consumer dissatisfaction: The effect of disconfirmed expectancy on perceived product performance. *Journal of marketing research*, 38–44, 1973.
- [11] W. De. Koster and D. Houtman. ‘STORMFRONT IS LIKE A SECOND HOME TO ME’ On virtual community formation by right-wing extremists. *Information, Communication & Society*, 11(8):1155–1176, 2008.
- [12] E. Nekmat and WJ. Gonzenbach. Multiple opinion climates in online forums: Role of website source reference and within-forum opinion congruency. *Journalism & mass communication quarterly*, 90(4):736–756, 2013.
- [13] P. Porten-Cheé and C. Eliders. Spiral of silence online: How online communication affects opinion climate perception and opinion expression regarding the climate change debate. *Studies in communication sciences*, 15(1):143–150, 2015.
- [14] A. Schulz and P. Roessler. The spiral of silence and the Internet: Selection of opinion content and the perception of the public opinion climate in computer-mediated communication environments. *International Journal of Public Opinion Research*, 24(3):346–367, 2012.
- [15] J. G. Carroll, F. H. Andrew, and J. Shanahan. Perceived support for one’s opinions and willingness to speak out: A meta-analysis of survey studies on the ‘spiral of silence’. *The Public Opinion Quarterly*, 61(3):452–463, 1997.
- [16] HK. Meyer and B. Speakman. Quieting the Commenters: The Spiral of Silence’s Persistent Effect on Online News Forums. *Quieting the Commenters: The Spiral of Silence’s Persistent Effect*, 51, 2016.
- [17] N. N. Dalvi, R. Kumar, and B. Pang. Para ‘normal’ activity: On the distribution of average ratings. In *ICWSM 2013*, pages 110–119.
- [18] J. B. Cohen and M. E. Goldberg. The dissonance model in post-decision product evaluation. *Journal of Marketing Research*, 315–321, 1970.
- [19] N. Jindal and B. Liu. Opinion spam and analysis. In *WSDM 2008*, pages 219–230.
- [20] D. Godes and J. C. Silva. Sequential and temporal dynamics of online opinion. *Marketing Science*, 31(3):448–473, 2012.
- [21] N. Hu, J. Zhang, and P. A. Pavlou. Overcoming the j-shaped distribution of product reviews. *Communications of the ACM*, 52(10):144–147, 2009.
- [22] Liu, Y, Cao, X. and Yu, Y. Are You Influenced by Others When Rating?: Improve Rating Prediction by Conformity Modeling. In *RecSys 2016*, pages 269–272.
- [23] P. Massa and P. Avesani. Trust-aware recommender systems. In *RecSys 2007*, pages 17–24.
- [24] YL. Zhang, Q. Guo, J. Ni and JG. Liu. Memory effect of the online rating for movies. *Physica A: Statistical Mechanics and its Applications*, 417:261–166, 2015.
- [25] J. McAuley, R. Pandey and J. Leskovec. Inferring networks of substitutable and complementary products. In *SIGKDD 2015*, pages 785–794.
- [26] J. Tang, H. Gao and H. Liu. mTrust: discerning multi-faceted trust in a connected world. In *WSDM 2012*, pages 93–102.
- [27] FM. Harper and JA. Konstan. The movielens datasets: History and context. *ACM Transactions on Interactive Intelligent Systems*, 5(4):19, 2016.
- [28] P. McJones. Eachmovie collaborative filtering data set. *DEC Systems Research Center*, 249:57, 1997.
- [29] P. DiMaggio, J. Evans and B. Bryson. Have American’s social attitudes become more polarized?. *American journal of Sociology*, 102(3):690–755, 1996.
- [30] RL. West and LH. Tumer. Introducing communication theory: Analysis and application. *McGraw-Hill Humanities/Social Sciences/Languages*, 2006.
- [31] DA. Scheufle and P. Moy. Twenty-five years of the spiral of silence: A conceptual review and empirical outlook. *International journal of public opinion research*, 12(1):3–28, 2000.
- [32] E. Noelle-Neumann. *The spiral of silence: Public opinion, our social skin*. University of Chicago Press., 1993.
- [33] Zhang, Xiaoying and Zhao, Junzhou and Lui, John C.S. Modeling the Assimilation-Contrast Effects in Online Product Rating Systems: Debiasing and Recommendations *Proceedings of the Eleventh ACM Conference on Recommender Systems*
- [34] R. Salakhutdinov and A. Mnih. Probabilistic matrix factorization. *NIPS*, 20:1257–1264, 2008.
- [35] Y. Koren, R. Bell, and C. Volinsky. Matrix factorization techniques for recommender systems. *Computer*, 42(8):30–37, 2009.