User Privacy in Recommender Systems

Duration: 2020 - early 2024

Peter Müllner

Advisors: Dominik Kowald, Elisabeth Lex

Motivation

- Recommender systems exploit user data to generate recommendations
- **Data Disclosure** through the recommendations! [7, 1]

Motivation

- Recommender systems exploit user data to generate recommendations
- **Data Disclosure** through the recommendations! [7, 1]
- Noise added to ensure **Differential Privacy** $\Rightarrow$ **Accuracy drop!**
- Our approach: Address the utilization of user data itself

Motivation

- Recommender systems exploit user data to generate recommendations
- **Data Disclosure** through the recommendations! [7, 1]
- Noise added to ensure **Differential Privacy** ⇒ **Accuracy drop!**
- Our approach: Address the utilization of user data itself

How do we address this?

Research Questions

RQ1: How can recommender systems use fewer data to generate meaningful recommendations?

Research Questions

RQ1: How can recommender systems use fewer data to generate meaningful recommendations?

RQ2: How can we improve the recommendation accuracy of differentially private recommender systems?

Research Questions

RQ1: How can recommender systems use fewer data to generate meaningful recommendations?

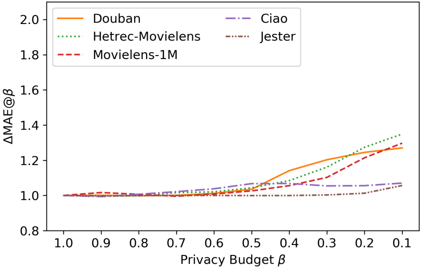
RQ2: How can we improve the recommendation accuracy of differentially private recommender systems?

RQ3: In which different ways does differential privacy impact personalized recommendations?

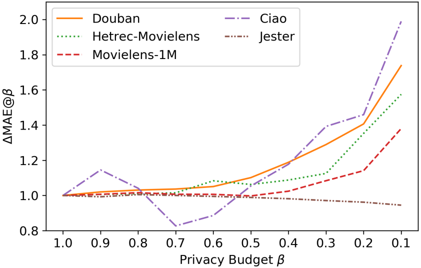
Preliminary Results

RQ1: How can recommender systems use fewer data to generate meaningful recommendations?

- Recommender system can only use a fraction β of each user's data
- **MetaMF** [2] with meta learning, **NoMetaMF** without meta learning
- Relative error change $\Delta\mathrm{MAE@}\beta = \mathrm{MAE@}\beta \ / \ \mathrm{MAE@1.0}$



(a) MetaMF

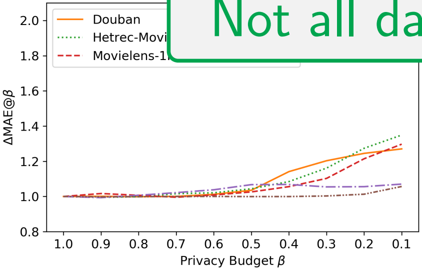


(b) MetaMF without meta learning

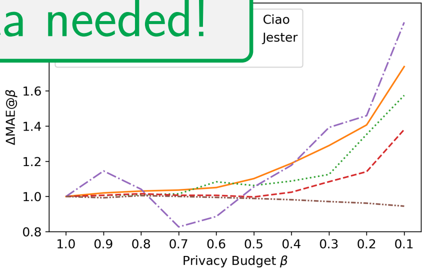
RQ1: How can recommender systems use fewer data to generate meaningful recommendations?

- Recommender system can only use a fraction β of each user's data
- **MetaMF** [2] with meta learning, **NoMetaMF** without meta learning
- Relative error change $\Delta\text{MAE@}\beta = \text{MAE@}\beta \ / \ \text{MAE@}1.0$

Not all data needed!



(a) MetaMF



(b) MetaMF without meta learning

RQ2: How can we improve the recommendation accuracy of differentially private recommender systems?

- *ReuseKNN*$_{DP}$ reuses the same neighbors for many recommendations
- **Many users** are not utilized and **do not need DP**
- Only **few users** are utilized and **need DP**

Method	ML 1M	Douban	LastFM	Goodreads
UserKNN$_{DP}$	80.39%	96.68%	99.89%	65.00%
ReuseKNN$_{DP}$	24.13%	34.40%	68.20%	29.12%

Table: Fraction of users that need DP.

RQ2: How can we improve the recommendation accuracy of differentially private recommender systems?

- *ReuseKNN$_{DP}$* reuses the same neighbors for many recommendations
- **Many users** are not utilized and **do not need DP**
- Only **fe**

Weaker accuracy drop, but privacy!

Met				eads
UserKNN$_{DP}$	80.39%	96.68%	99.89%	65.00%
ReuseKNN$_{DP}$	24.13%	34.40%	68.20%	29.12%

Table: Fraction of users that need DP.

RQ3: In which different ways does differential privacy impact personalized recommendations?

- DP leads to accuracy drop; any different impacts?
- Jaccard distance between recommendations with and without DP
- Hit Rate (HR) and **Average recommendation popularity** (ARP)

Method	No. Users	ΔHR	ΔARP
SlopeOne	65.01%	-4.79%	-12.92%
UserKNN	68.10%	-11.23%	-12.63%
NMF	68.84%	-23.09%	-30.34%

Table: Impact of DP on recommendations (MovieLens 1M).

RQ3: In which different ways does differential privacy impact personalized recommendations?

- DP leads to accuracy drop; any different impacts?
- Jaccard distance between recommendations with and without DP
- Hit Rate (HR … **opularity** (ARP)

With DP, weaker Popularity Bias!

Me…			…RP
SlopeOne	[illegible]	-4.79%	-12.92%
UserKNN	68.10%	-11.23%	-12.63%
NMF	68.84%	-23.09%	-30.34%

Table: Impact of DP on recommendations (MovieLens 1M).

Progress so far

Progress so far

Main Publications:

- *RQ1*: Robustness of Meta Matrix Factorization Against Strict Privacy Constraints (ECIR'21) [3] ✓
- *RQ2*: ReuseKNN: Neighborhood Reuse for Differentially-Private KNN-Based Recommendations [6] (under review)
- *RQ3*: The Impact of Differential Privacy on Personalized Recommendations (under review)

Additional Publications:

- User Privacy in Recommender Systems (ECIR'23) [4] ✓
- Position Paper on Simulating Privacy Dynamics in Recommender Systems (SimuRec@RecSys'21) [5] ✓

Future:

- [user-based countermeasures] (planned for end of 2023)

Topics for Discussion

Study DP on the user-level. Users are vastly different (preferences, privacy concerns, goals, ...) $\rightarrow$ **DP and Fairness?**

How do user characteristics correlate with the impact of DP?
How can we ensure fairness when applying DP?
How to study and evaluate fairness and DP? Simulation?

Thank you!

Source Code

github.com/pmuellner/

Contact

pmuellner[AT]know[MINUS]center[DOT]at

or

pmuellner.github.io

References I

J. A. Calandrino, A. Kilzer, A. Narayanan, E. W. Felten, and V. Shmatikov, *" you might also like:" privacy risks of collaborative filtering*, in 2011 IEEE symposium on security and privacy, IEEE, 2011, pp. 231–246.

Y. Lin, P. Ren, Z. Chen, Z. Ren, D. Yu, J. Ma, M. d. Rijke, and X. Cheng, *Meta matrix factorization for federated rating predictions*, in Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval, 2020, pp. 981–990.

P. Muellner, D. Kowald, and E. Lex, *Robustness of meta matrix factorization against strict privacy constraints*, in European Conference on Information Retrieval, Springer, 2021, pp. 107–119.

References II

- P. Müllner, *User privacy in recommender systems*, in Advances in Information Retrieval: 45th European Conference on Information Retrieval, ECIR 2023, Dublin, Ireland, April 2–6, 2023, Proceedings, Part III, Springer, 2023, pp. 456–461.
- P. Müllner, E. Lex, and D. Kowald, *Position paper on simulating privacy dynamics in recommender systems*, arXiv preprint arXiv:2109.06473, (2021).
- P. Müllner, E. Lex, M. Schedl, and D. Kowald, *Reuseknn: Neighborhood reuse for differentially-private knn-based recommendations*, 2022.
- N. Ramakrishnan, B. J. Keller, B. J. Mirza, A. Y. Grama, and G. Karypis, *Privacy risks to straddlers in recommender systems*, IEEE Internet Computing, 5 (2001), pp. 1–16.