Robustness of Meta Matrix Factorization Against Strict Privacy Constraints

Peter Müllner, Dominik Kowald, Elisabeth Lex

ECIR 2021 - Reproducibility Track

Motivation

- Many large-scale recommender systems
- No “one size fits all” recommendations
- Need for personalization
- Inclusion of more user data, e.g., interaction data, personal data, ...

Privacy Issues

- Utilization of enormous amounts of user data
- Sensitive attributes could get leaked
  - Data breach
  - Infer private attributes
- Also state of the art could leak data, e.g., federated learning [1]

State-of-the-art Solution by Lin et al.

- Meta Matrix Factorization for Federated Rating Prediction [2]
- Private and federated matrix factorization framework
  1. Users share updates with global model (FL)
  2. Exploits collaborative information among users (MF)
  3. Generates private rating prediction model for each user (ML)
- Learns the parameters of the rating prediction model via meta learning

One Limitation of MetaMF

Although MetaMF incorporates [...] it still needs a certain amount of data for each user to achieve satisfactory performance. As a result, it does not perform well when there is not enough personalized data for each user. - Lin et al.

One Limitation of MetaMF

Although MetaMF incorporates [...] it still needs a certain amount of data for each user to achieve satisfactory performance. As a result, it does not perform well when there is not enough personalized data for each user. - Lin et al.

How much data is needed?

The Problem

- There can be no data disclosure if there is no data
- **Privacy budget** refers to the fraction of data a user shares with the recommender system
- Large privacy budget $\Rightarrow$ high utility $\Rightarrow$ low privacy
- Small privacy budget $\Rightarrow$ low utility $\Rightarrow$ high privacy

The Problem

- There can be no data disclosure if there is no data
- **Privacy budget** refers to the fraction of data a user shares with the recommender system
- Large privacy budget $\Rightarrow$ high utility $\Rightarrow$ low privacy
- Small privacy budget $\Rightarrow$ low utility $\Rightarrow$ high privacy

Limitation of MetaMF! (How much data is needed?)

The Problem

- There can be no data disclosure if there is no data
- **Privacy budget** refers to the fraction of data a user shares with the recommender system
- Large privacy budget $\Rightarrow$ high utility $\Rightarrow$ low privacy
- Small privacy budget $\Rightarrow$ low utility $\Rightarrow$ high privacy

Limitation of MetaMF! (How much data is needed?)
We test MetaMF's robustness against small privacy budgets

Reproducibility Study

Reproduction of Meta Matrix Factorization by Lin et al. [2]

Privacy-focused Study

MetaMF's robustness against strict privacy budgets

Reproducibility Study

Reproduction of Meta Matrix Factorization by Lin et al. [2]

Privacy-focused Study

MetaMF's robustness against strict privacy budgets

Approach

- RQ1a: Performance of MetaMF on different datasets?
  - Details given in the paper
  - Code and data publicly available
- RQ1b: Ability to personalize and exploit collaborative information?
  - Details not clear in paper
  - Code not publicly available
  - Additional information provided by Lin et al.

RQ1a: Performance on different datasets

- Trainset and testset (80/20) provided by Lin et al.
- Plus same parameters, e.g., learning rate, batch size, ...
- Our/Lin et al.'s results

Dataset	MAE	MSE
Douban	0.588/0.584	0.554/0.549
Hetrec-MovieLens	0.577/0.571	0.587/0.578
MovieLens 1M	0.687/0.687	0.765/0.760
Ciao	0.774/0.774	1.125/1.043
Jester	0.856/-	1.105/-

RQ1a: Performance on different datasets

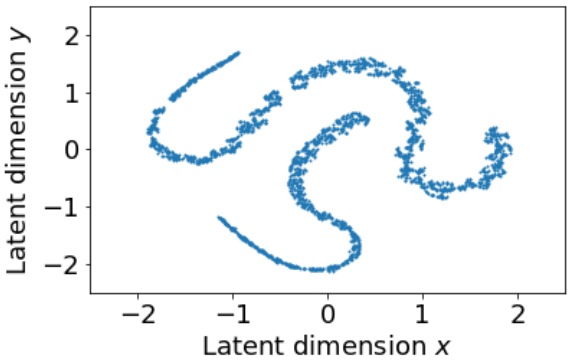
- Trainset and testset (80/20) provided by Lin et al.
- Plus same parameters, e.g., learning rate, batch size, ...
- Our/Lin et al.'s results

Reproducible

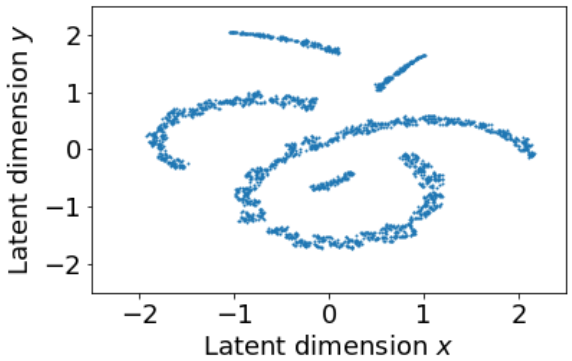
Dataset		
Douban		/0.549
Hetrec-MovieLens	0.577/0.571	0.587/0.578
MovieLens 1M	0.687/0.687	0.765/0.760
Ciao	0.774/0.774	1.125/1.043
Jester	0.856/-	1.105/-

RQ1b: Personalization and collaboration

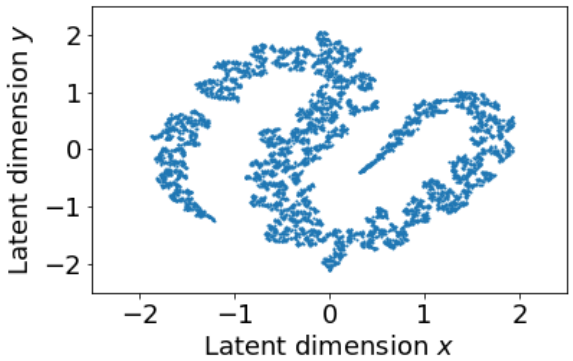
- Weights of the rating prediction model
- Unique embeddings correspond to personalization
- Clusters correspond to collaboration



(a) Douban



(b) Hetrec-MovieLens

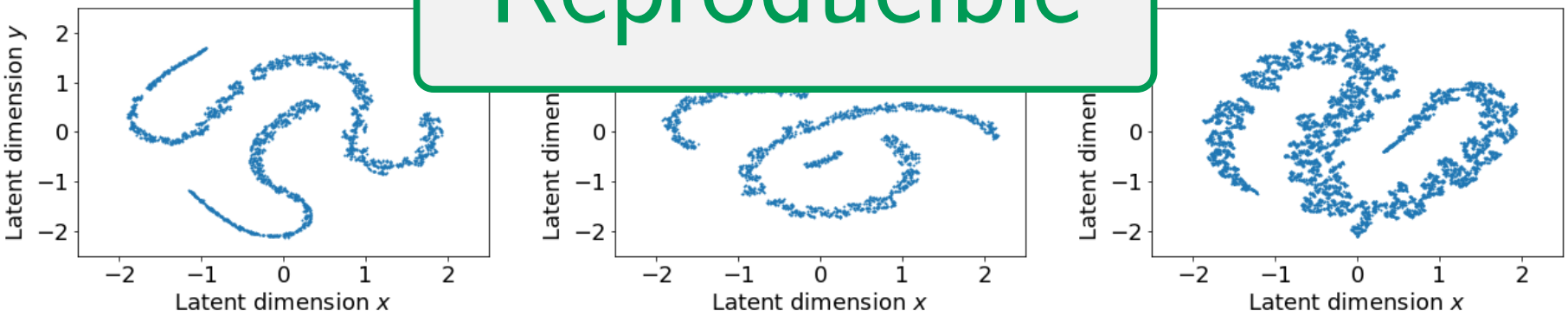


(c) MovieLens 1M

RQ1b: Personalization and collaboration

- Weights of the rating prediction model
- Unique embeddings correspond to personalization
- Clusters correspond to collaboration

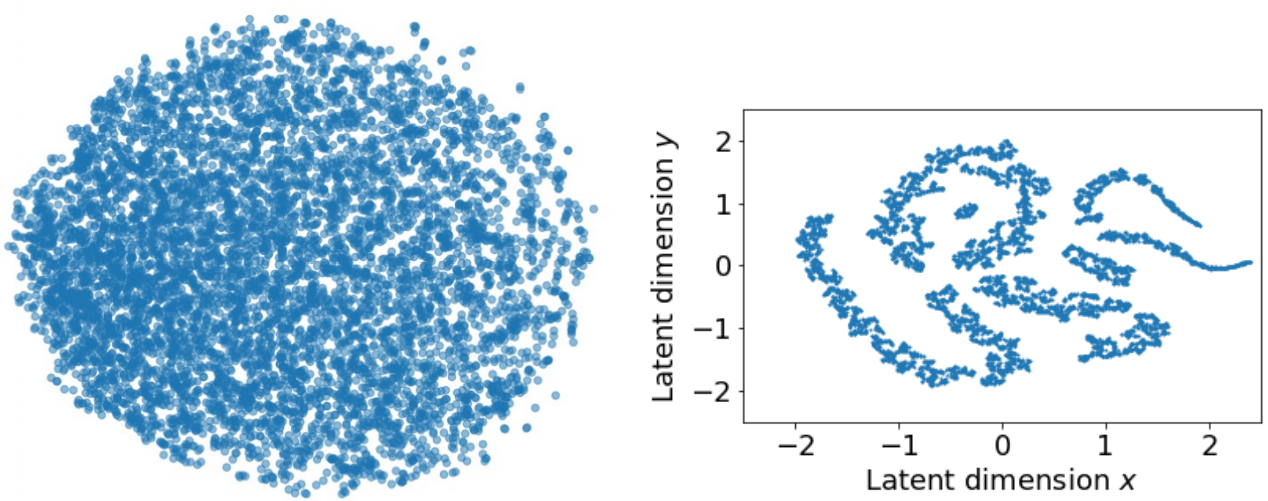
Reproducible



(a) Douban (b) Hetrec-MovieLens (c) MovieLens 1M

RQ1b: Personalization and collaboration

- **Ciao: Lin et al.'s results different to our results**
- Too few ratings per user ⇒ no clusters?

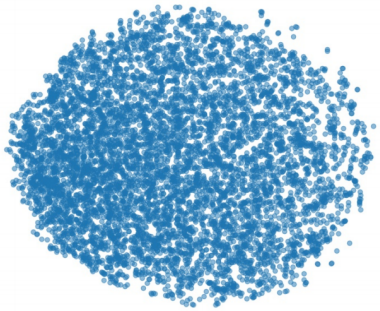


(a) Lin et al.

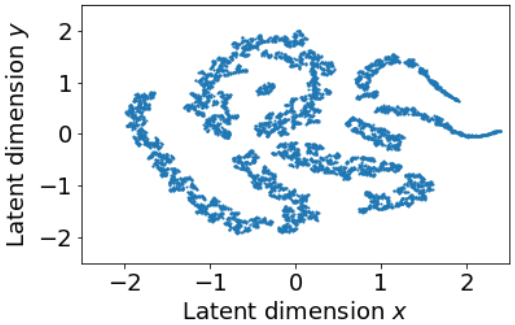
(b) Our reproduction

RQ1b: Personalization and collaboration

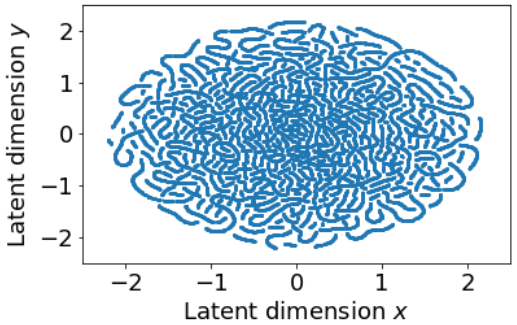
- **Ciao: Lin et al.'s results different to our results**
- Too few ratings per user ⇒ no clusters?
- Ciao and Jester with small amount of data per user



(a) Ciao by Lin et al.



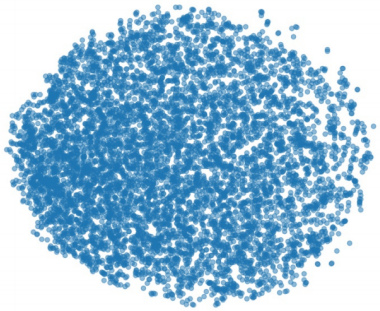
(b) Ciao



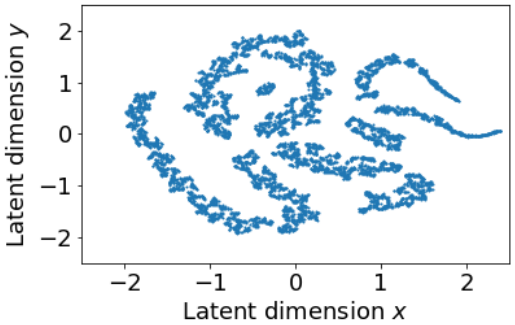
(c) Jester

RQ1b: Personalization and collaboration

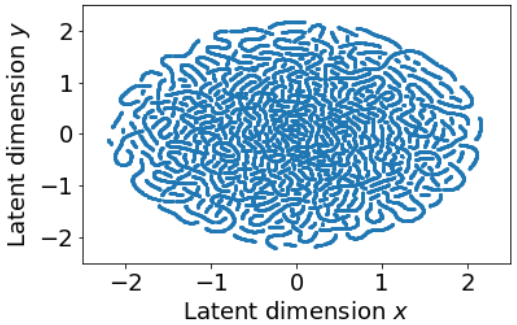
- **Ciao: Lin et al.'s results different to our results**
- Too few ratings per user ⇒ no clusters? Wrong!
- Ciao and Jester with small amount of data per user



(a) Ciao by Lin et al.



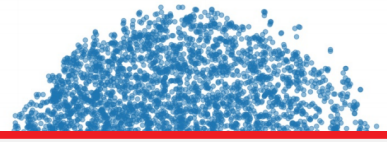
(b) Ciao



(c) Jester

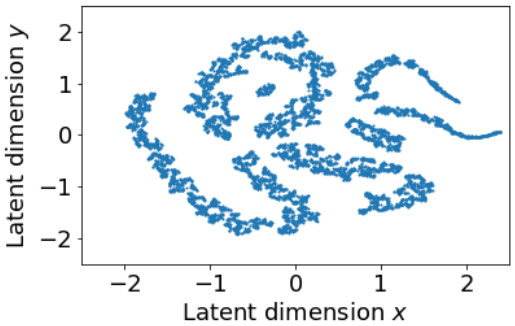
RQ1b: Personalization and collaboration

- **Ciao: Lin et al.'s results different to our results**
- Too few ratings per user ⇒ no clusters? Wrong!
- Ciao and Jester with small amount of data per user

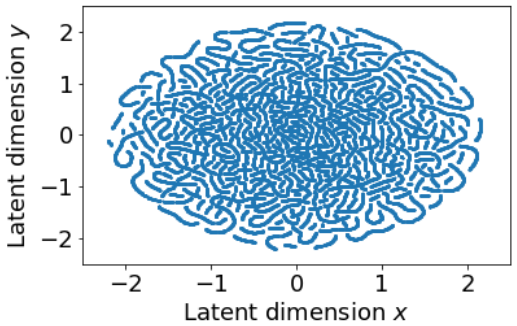


Not Reproducible

(a) Ciao by Lin et al.



(b) Ciao



(c) Jester

Takeaways

- Error measurements are reproducible (margin of error $\leq 2\%$)
- Weight embeddings are (mostly) reproducible
- Code only partially provided
- Some details missing in paper
- Authors provided additional information

Code

github.com/pmuellner/RobustnessOfMetaMF

Reproducibility Study

Reproduction of Meta Matrix Factorization by Lin et al. [2]

Privacy-focused Study

MetaMF's robustness against strict privacy budgets

Reproducibility Study

Reproduction of Meta Matrix Factorization by Lin et al. [2]

Privacy-focused Study

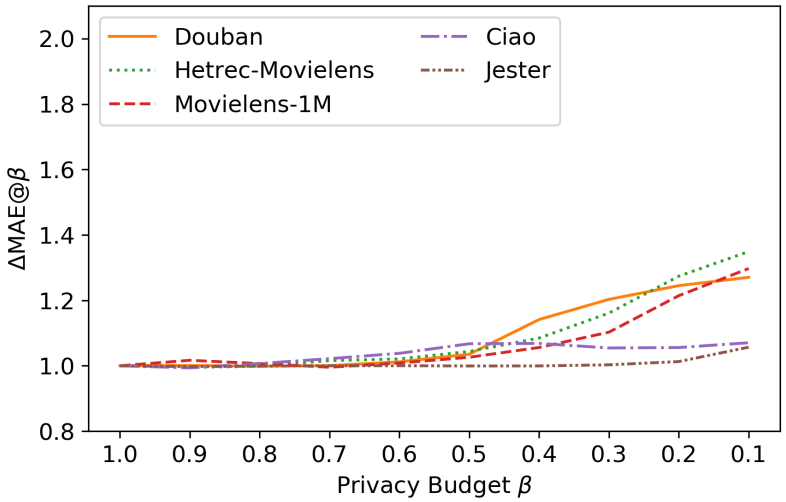
MetaMF's robustness against strict privacy budgets

Approach

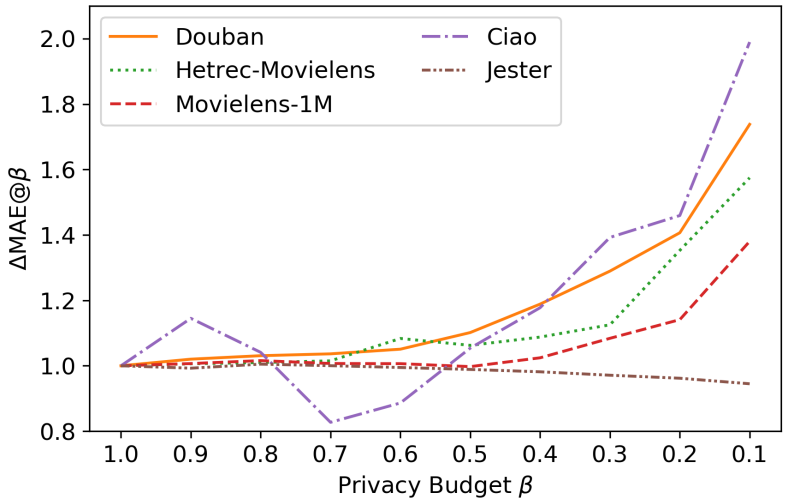
- RQ2a: Benefit of meta learning for small privacy budgets?
  - **NoMetaMF**, i.e., MetaMF with no meta learning
  - Comparison between MetaMF and NoMetaMF
- RQ2b: Effect of small privacy budgets on different user groups?
  - Identify three user groups
  - Evaluate MetaMF on these user groups

RQ2a: Benefit of meta learning

- **MetaMF** with meta learning, **NoMetaMF** without meta learning
- Privacy budget β is the fraction of a user's shared data
- Relative performance loss $\Delta\mathrm{MAE@}\beta = \mathrm{MAE@}\beta \;/\; \mathrm{MAE@1.0}$



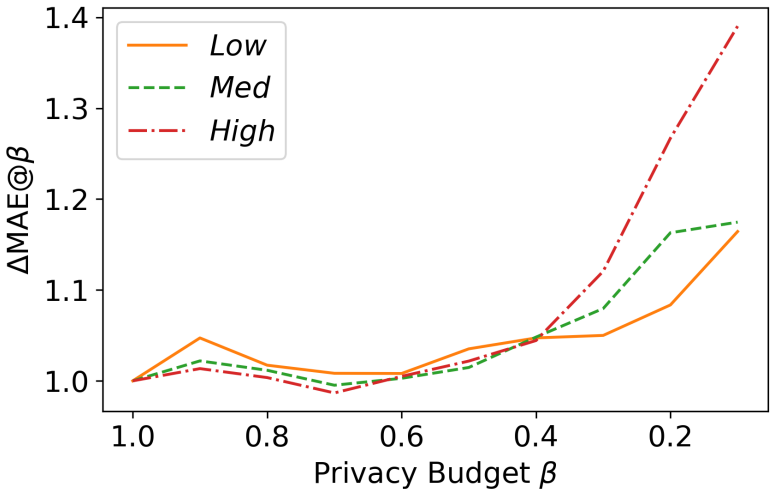
(a) MetaMF



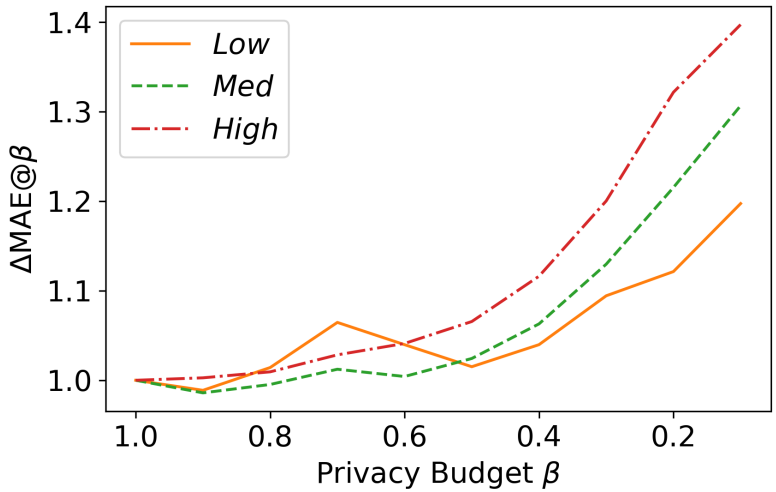
(b) NoMetaMF

RQ2b: Effect on different user groups

- Identify user groups based on number of ratings
- Train MetaMF on all users, evaluate on *Low*, *Med*, and *High*



(a) MovieLens 1M



(b) Hetrec-MovieLens

Takeaways

- Sharing $\geq 50\%$ of data is no big problem
- Sharing $< 50\%$ of data requires meta learning
- *Low* can afford privacy
- *High* has to sacrifice a lot of utility for privacy
- Different privacy-utility tradeoffs!

Conclusions & Future Work

Reproducibility Study

Reproduction of Meta Matrix Factorization by Lin et al. [2]

Privacy-focused Study

MetaMF's robustness against strict privacy budgets

Reproducibility Study

Reproduction of Meta Matrix Factorization by Lin et al. [2]

- Not all information in paper ✗
- Error measurements ✓
- Weight embeddings ✓

Privacy-focused Study

MetaMF's robustness against strict privacy budgets

Reproducibility Study

Reproduction of Meta Matrix Factorization by Lin et al. [2]

- Not all information in paper ✗
- Error measurements ✓
- Weight embeddings ✓

Mostly Reproducible

Privacy-focused Study

MetaMF's robustness against strict privacy budgets

Reproducibility Study

Reproduction of Meta Matrix Factorization by Lin et al. [2]

- Not all information in paper ✗
- Error measurements ✓
- Weight embeddings ✓

Mostly Reproducible

Privacy-focused Study

MetaMF's robustness against strict privacy budgets

- Robust if users share $\geq 50\%$
- Small privacy budgets
  $\Rightarrow$ meta learning
- Tradeoffs differ per user group

Future Work

- Data with different privacy requirements
- What data should be shared based on data characteristics
- Not sharing data could decrease utility of other users

Thank you for your attention!

Source Code
github.com/pmuellner/RobustnessOfMetaMF

User Groups Data
zenodo.org/record/4031011

Contacts
{pmuellner, dkowald}@know-center.at
elisabeth.lex@tugraz.at

Peter Müllner
Know-Center GmbH

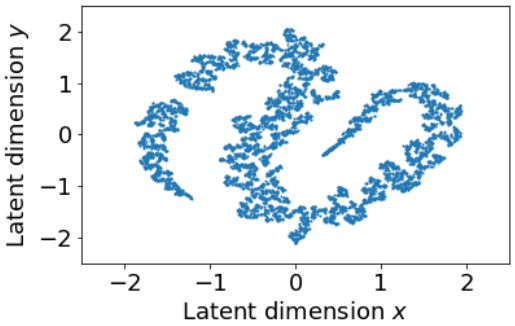
Dominik Kowald
Know-Center GmbH

Elisabeth Lex
Graz University of Technology

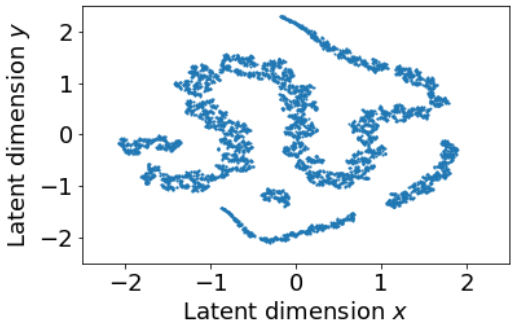
Dataset statistics

Dataset	$\|U\|$	$\|I\|$	$\|R\|$	User Group Size
Douban	2,509	39,576	893,575	125
Hetrec-MovieLens	2,113	10,109	855,598	106
MovieLens 1M	6,040	3,706	1,000,209	302
Ciao	7,373	105,096	282,619	369
Jester	73,321	100	4,136,360	2,671

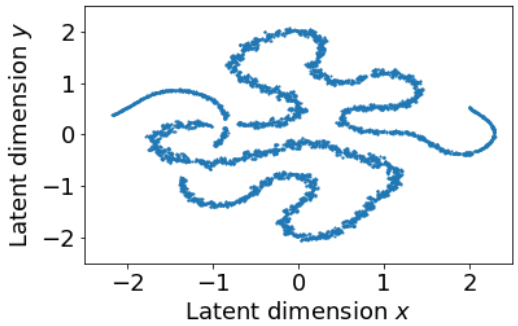
Weight embeddings of MetamF and NoMetaMF



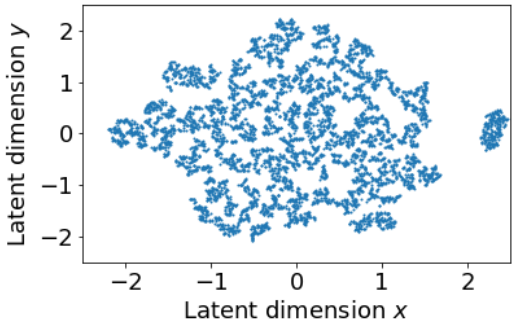
(a) $\beta = 1.0$



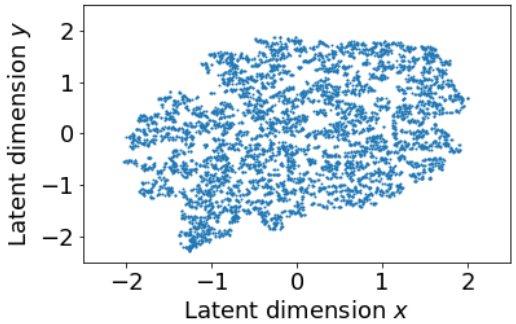
(b) $\beta = 0.5$



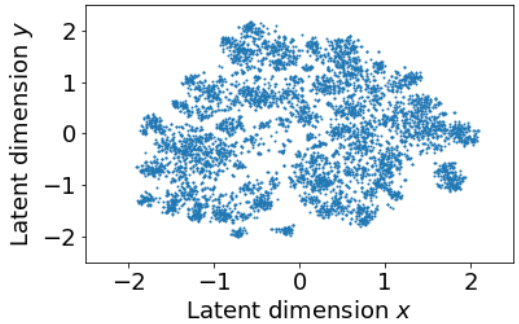
(c) $\beta = 0.1$



(d) $\beta = 1.0$



(e) $\beta = 0.5$

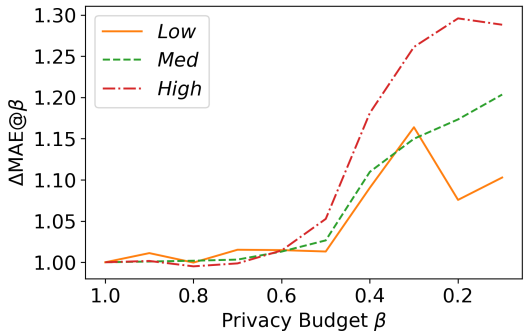


(f) $\beta = 0.1$

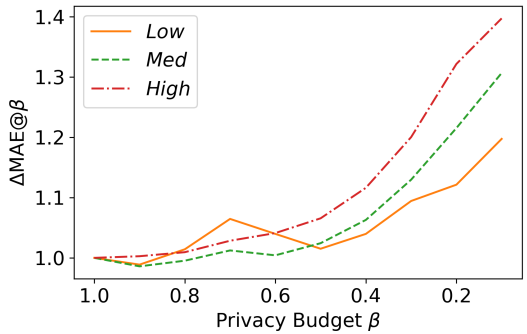
MAE measurements for our user groups

Dataset	*Low*	*Med*	*High*
Douban*	0.638	0.582	0.571
Hetrec-MovieLens****	0.790	0.603	0.581
MovieLens 1M****	0.770	0.706	0.673
Ciao	0.773	0.771	0.766
Jester****	1.135	0.855	0.811

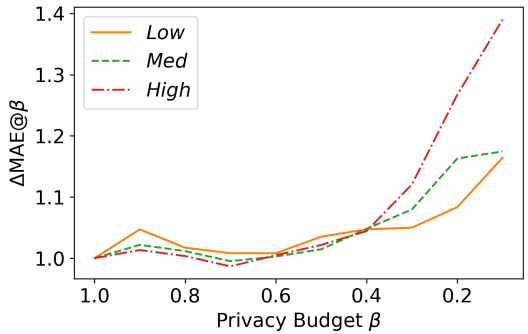
ΔMAE@β measurements for our user groups



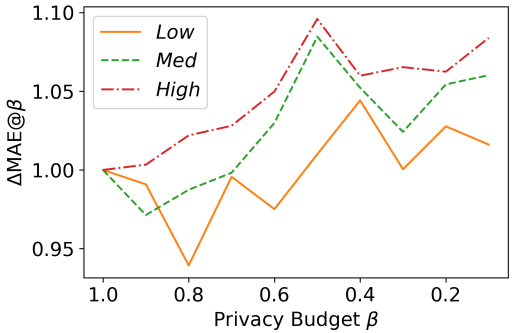
(a) Douban



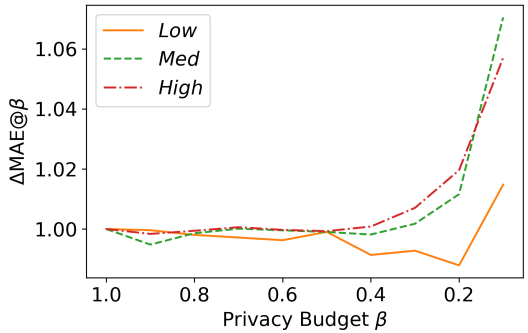
(b) Hetrec-MovieLens



(c) MovieLens 1M



(d) Ciao



(e) Jester

References I

E. Bagdasaryan, A. Veit, Y. Hua, D. Estrin, and V. Shmatikov, *How to backdoor federated learning*, in International Conference on Artificial Intelligence and Statistics, PMLR, 2020, pp. 2938–2948.

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.