PRIVACY TUTORIAL

Bart Knijnenburg and Shlomo Berkovsky
RecSys-2017, Como

Motivation

- Effective personalization requires large amounts of user data
  - Accurate, detailed, and up-to-date user profiles
    - More reliable recommendations
  - Correlation between quality of user data and quality of recommendations
- Trade-off between personalization and data privacy (privacy-personalization paradox)
  - More personal data is available
    - Better recommendations are generated
  - More personal data is available
    - Less privacy users are remained with

Motivation

- From a promotional brochure

- $1M question: Can we have both at the same time?!

Data collection

- Recommender systems collect user data
- Fair Information Practices (FIPS) for data collection
  - Purposes of collection should be specified at the time of collection
  - Data should
    - be collected within limits, by lawful and fair means and with consent
    - not be used or disclosed for other purposes except with consent or by law
    - be protected against unauthorized access, destruction, use, or disclosure
  - Users should
    - be able to know what is being collected, for what purposes, and who controls the data
    - be allowed to inspect the data, and have them erased, rectified, completed or amended
  - Collector should be accountable for complying with the above measures
    - Any violation of FIPS is a privacy breach

Privacy risks in recommenders

Adversary	Direct access to existing data	Inference of new data
Recommender system	Unsolicited data collection	Exposure of sensitive information
	Sharing data with third parties	Targeted advertising
	Unsolicited access by employees	Discrimination
Other users	Leaks through shared device or service	Inference from the recommender output
External entities	Lawful data disclosure	Exposure of sensitive information
	Hacking	
	Re-identification of anonymized data	

Adversary = recommender system

- Direct access
  - Unsolicited collection: any breach of FIPS
  - Sharing: sell, recommendations as a service, Netflix prize competition, ..
  - Access by staff
- Inference
  - Much attention in the DM/ML community
    - Kosinski: FB likes → ethnicity, sexual orientation, religious/political views, personality, ..
  - RecSys scenarios
    - Cross-system recommendations
    - Targeted advertisement
    - Price discrimination

Other adversaries

- Adversary = other system users
  - Direct: shared account or shared device
  - Inference
    - Collaborative recommendations based on user-similarity
    - Fake profiles will uncover data of similar real profiles
- Adversary = external entities
  - Direct
    - Cybersecurity attacks
    - Court and law enforcement orders
  - Inference: de-anonymization of Netflix dataset using IMDb ratings
    - "with background knowledge of 8 ratings (of which 2 may be wrong) and dates (within a 14-day error), 99% of records can be uniquely re-identified"

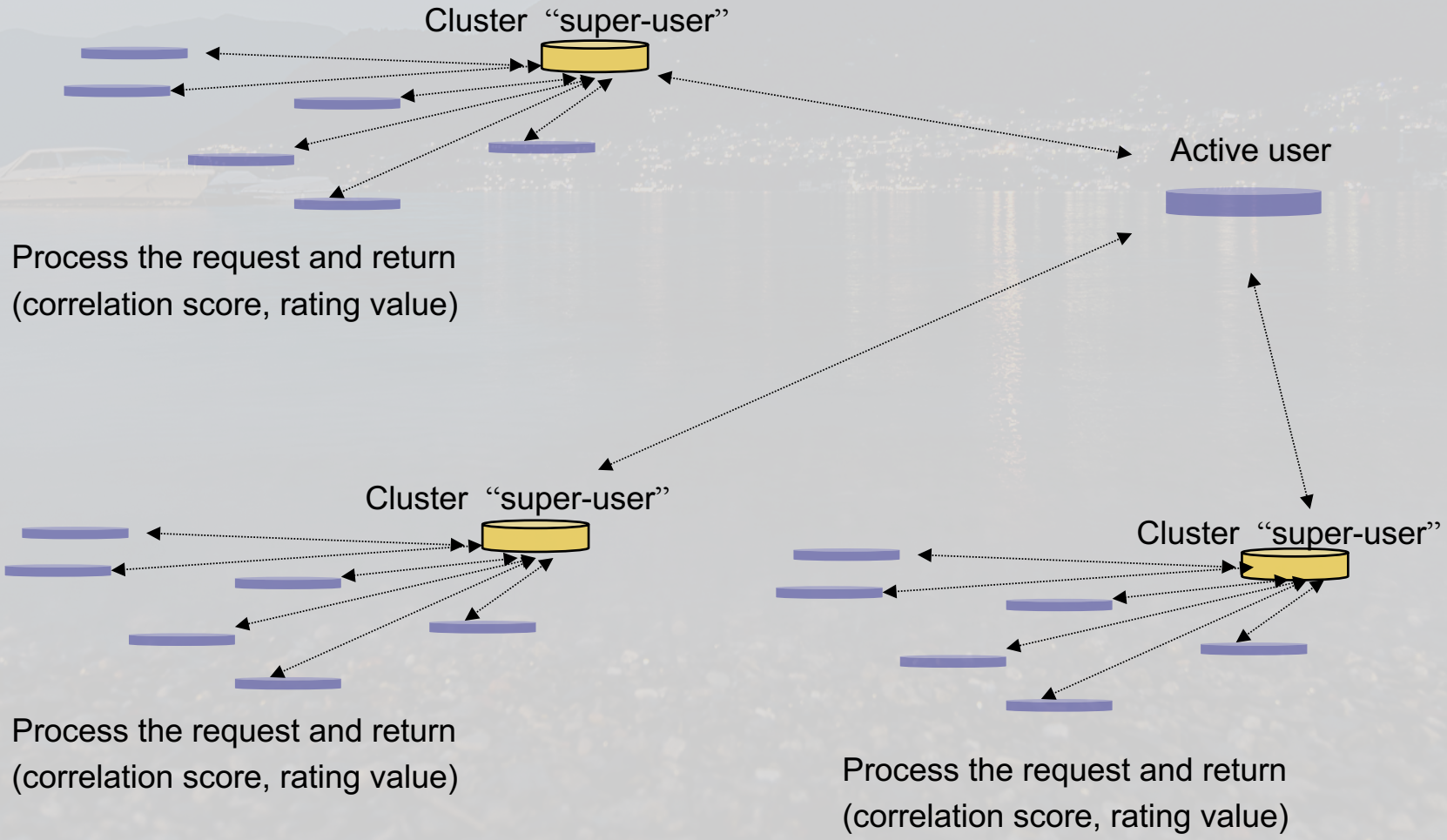
Taxonomy of solutions

- Architectural
  - Software architectures, platforms, and standards that minimize the personal data leakage threat
- Algorithmic
  - Data modification approaches and formal guarantees that ensure that leaked personal data discloses only modified or encrypted information
- User-centric
  - User behaviors, attitudes, and decisions related to the disclosure of their personal data

Architectural solutions

- Trusted software
  - System guarantees for data storage, linkability, and disclosure
  - Technical audit and certification by a trusted third party
- Distributed and decentralised profile storage
  - User profile storage by the (mobile) end-user device
  - Disclosure using Semantic Web and Social Web technologies
- Distributed recommendation generation
  - Peer-to-peer decentralised recommendation process
  - Virtual user communities masking individual users

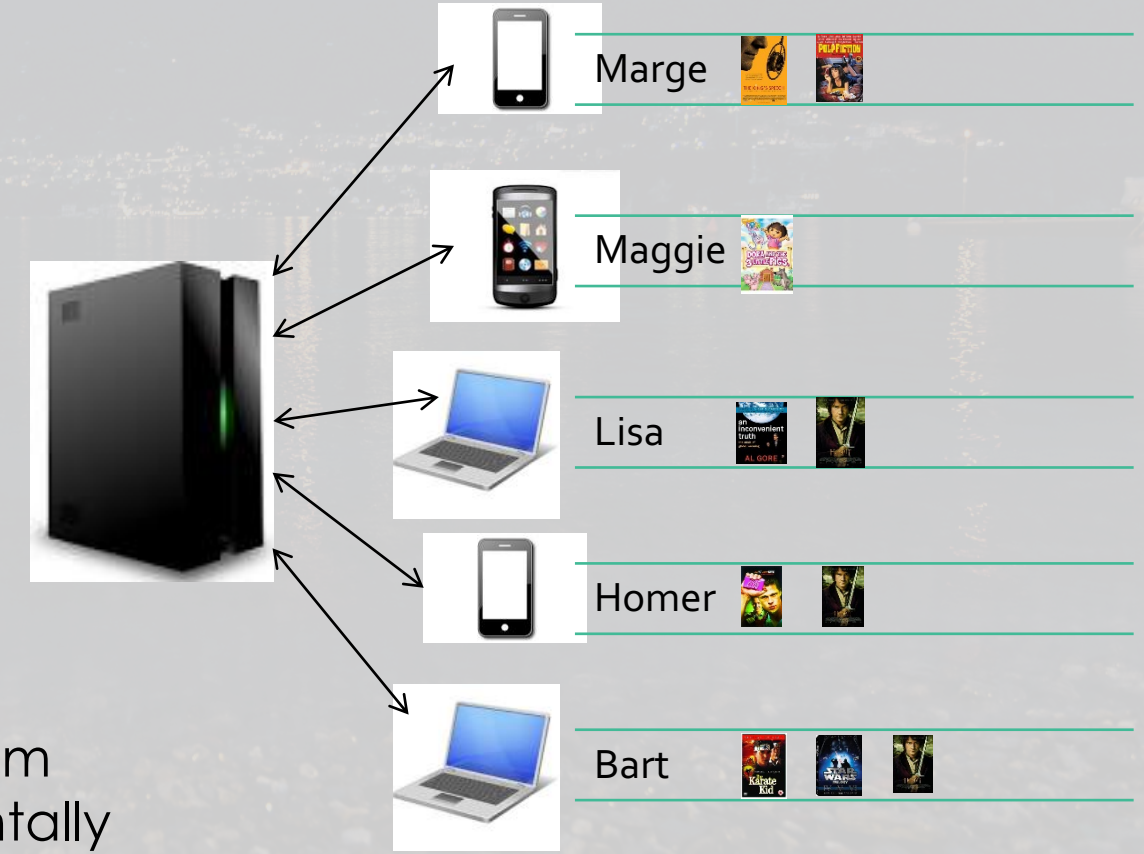
Hierarchical topology for CF



MF without user data retention

Q: Item-features matrix

	f_1	f_2	f_3
	3	1	-3
	2	-2	-1
	-4	-2	1
	2	-1	4



No user vectors or profiles are stored
1. User inputs are processed as a stream
2. Item vectors are updated incrementally
3. User vectors and profiles are discarded

Algorithmic solutions

- Groups of approaches
  - Anonymization
    - Hide users by pseudonyms and anonymous identity
  - Data obfuscation
    - Modify user data by adding noise to the real data
  - Differential privacy
    - Modify user data by adding noise to the real data
    - Provides formal privacy guarantees
  - Cryptography
    - Secure computation protocols using encrypted data

Anonymization

- Inherently misses the point of personalization
  - Addressed through
    - personas or stereotypical user profiling
    - role-based data access permissions
- Can be neutralised using external data records
  - De-anonymization of Netflix dataset
  - Massachusetts medical records disclosure
    - 54K people → (birth date, gender, postcode) → single person

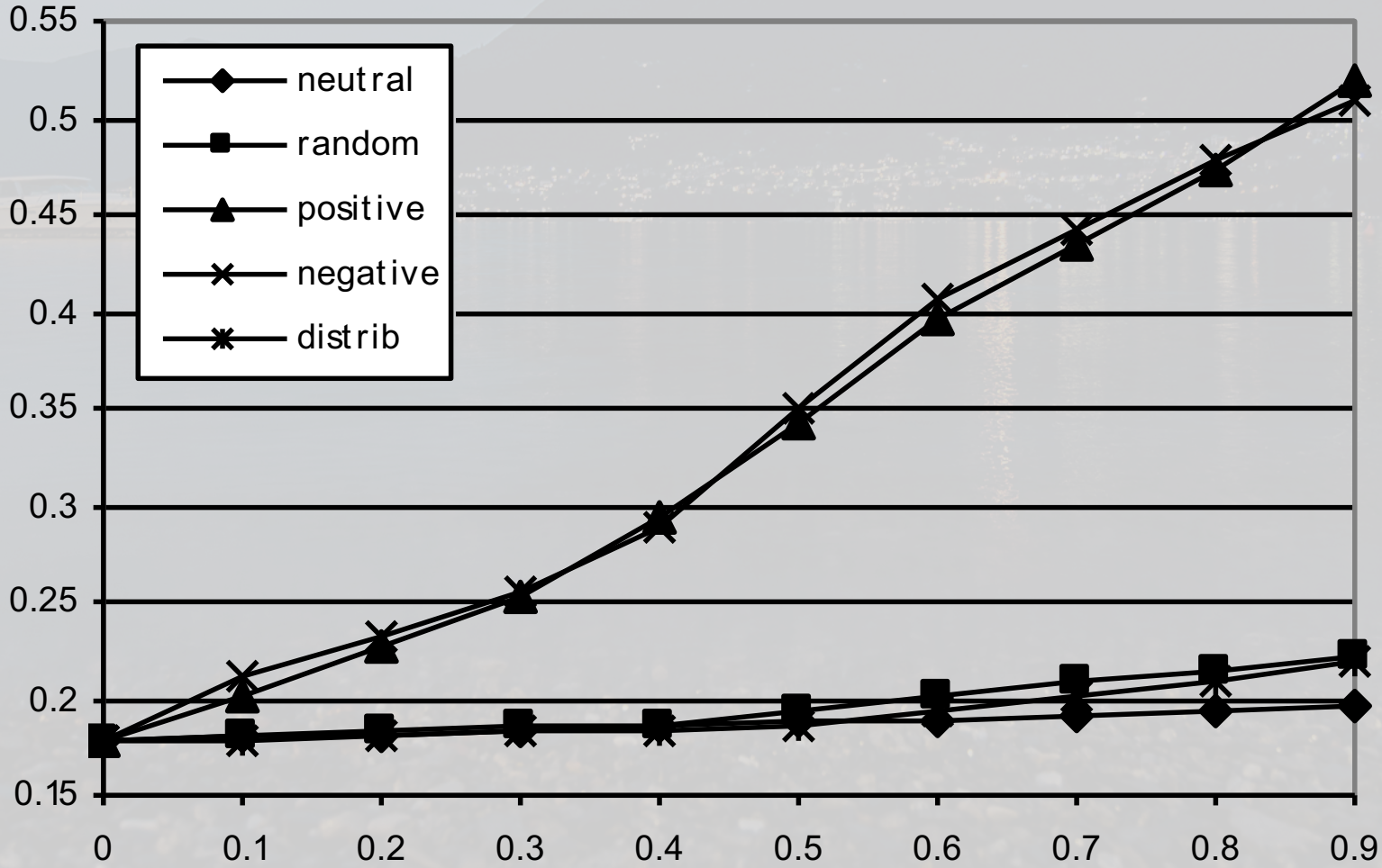
Obfuscation – intuition

- Users or system modify user profile data
  - Noise is added to real ratings
- Profiles may be revealed by attacker
  - For example, multiple recommendations requests
  - But only modified profiles will be revealed
- Will modifying profiles may affect recommendations?
  - Privacy-personalization trade-off

Obfuscation in CF – experimental setting

- 5 obfuscation policies
  - Positive: substitute by the highest rating in the dataset
  - Negative: substitute by the lowest rating in the dataset
  - Neutral: substitute by the neutral rating in the dataset
  - Random – substitute by random values in the range of dataset ratings
  - Distribution – substitute by values reflecting the distribution of ratings
- What is obfuscated?
  - All the ratings or extreme (positive/negative) ratings
- What is predicted?
  - All the ratings or extreme (positive/negative) ratings

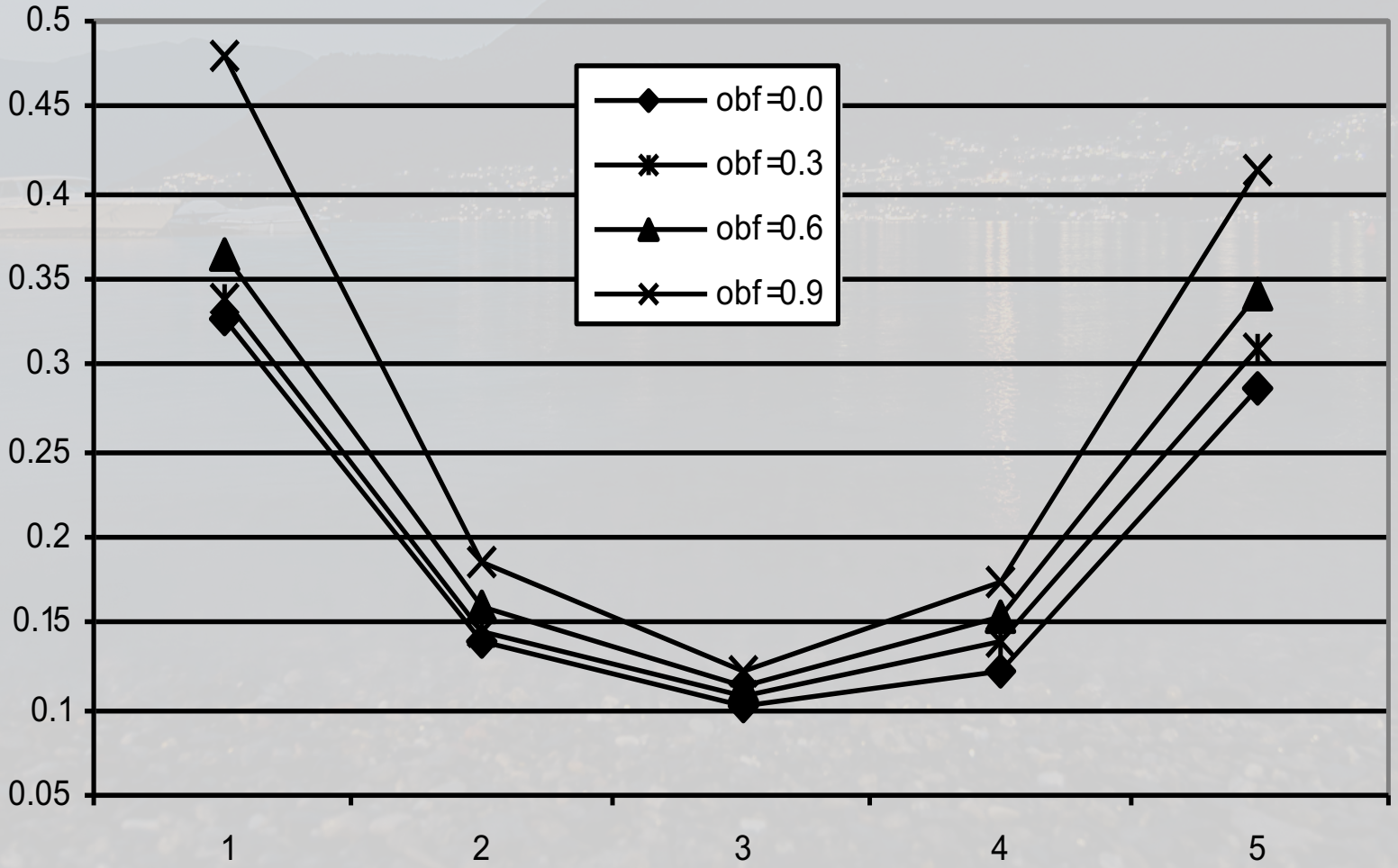
All obfuscated – all predicted



All obfuscated – all predicted

- For random, neutral, and distribution policies
  - Modified ratings are similar to the original
    - Observation: close to normal distribution of ratings
  - Error increases with the obfuscation rate
- For positive and negative policies
  - Modified ratings are different from the real
  - Error increases with the obfuscation rate
- error(positive,negative) > error(random,neutral,distribution)
- Hard to optimize privacy and accuracy at the same time

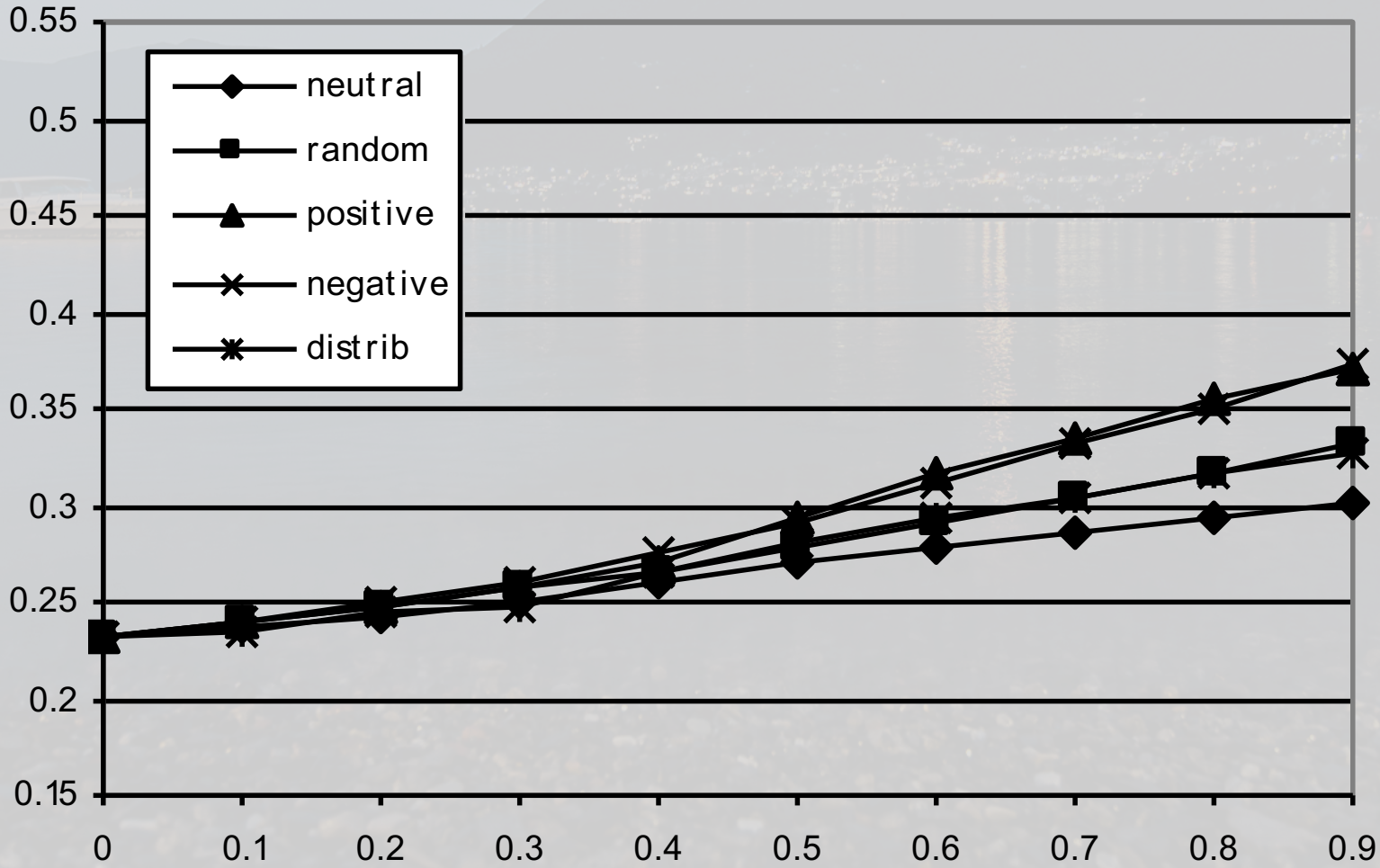
All obfuscated – extreme predicted



All obfuscated – extreme predicted

- Distribution policy applied
- Obfuscation effect
  - Is manageable on moderate ratings
    - These are easier to predict
  - Is stronger on extreme ratings
    - These are harder to predict
- error(extreme) > error(moderate)
- error_increase(extreme) > error_increase(moderate)

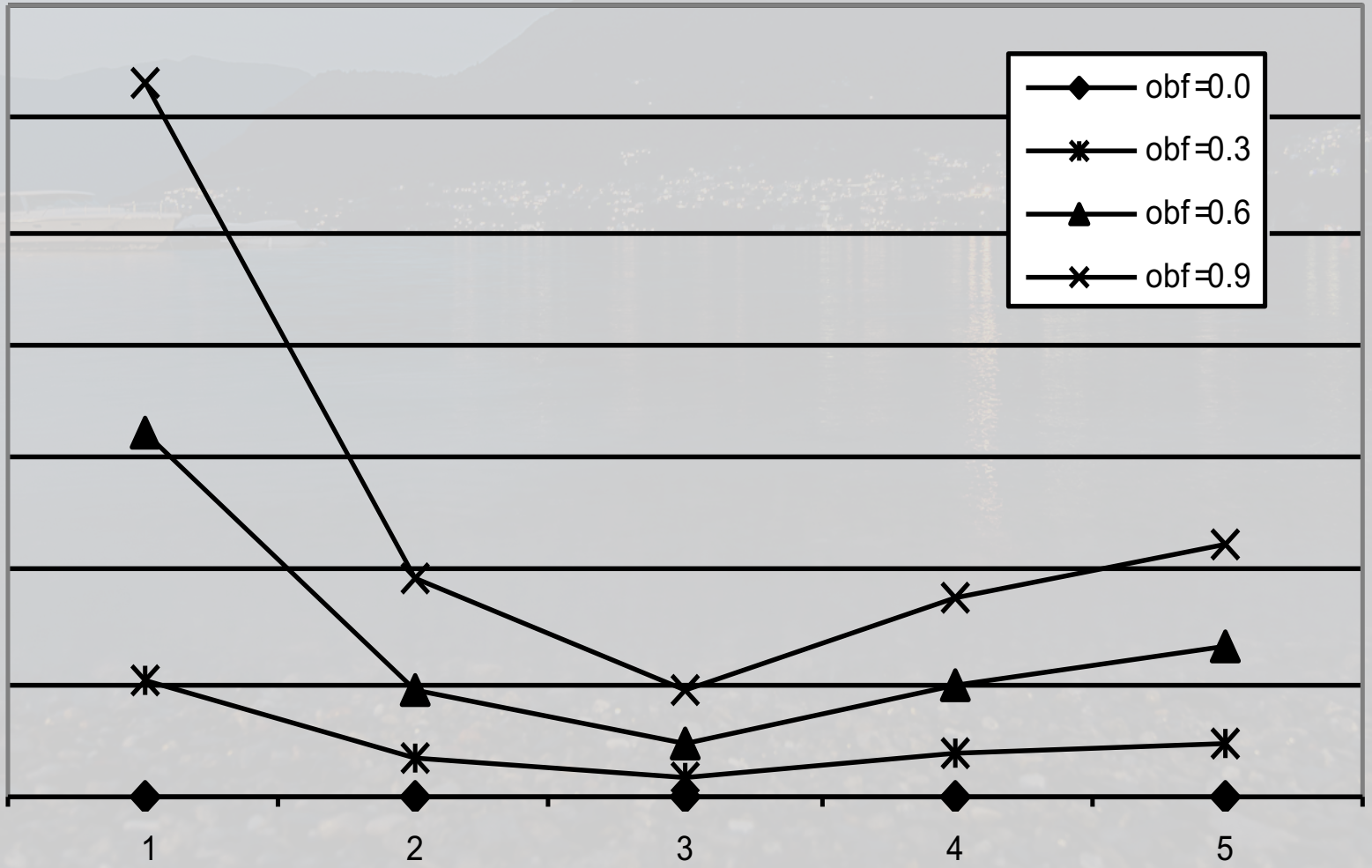
Extreme obfuscated – all predicted



Extreme obfuscated – all predicted

- For positive and negative policies
  - Error increases with the obfuscation rate
  - Weaker than for all predictions
- For random, neutral and distribution policies
  - Error increases with the obfuscation rate
  - Stronger than for all predictions
- error(positive,negative) is closer to error(random,neutral,distribution)

Extreme obfuscated – extreme predicted

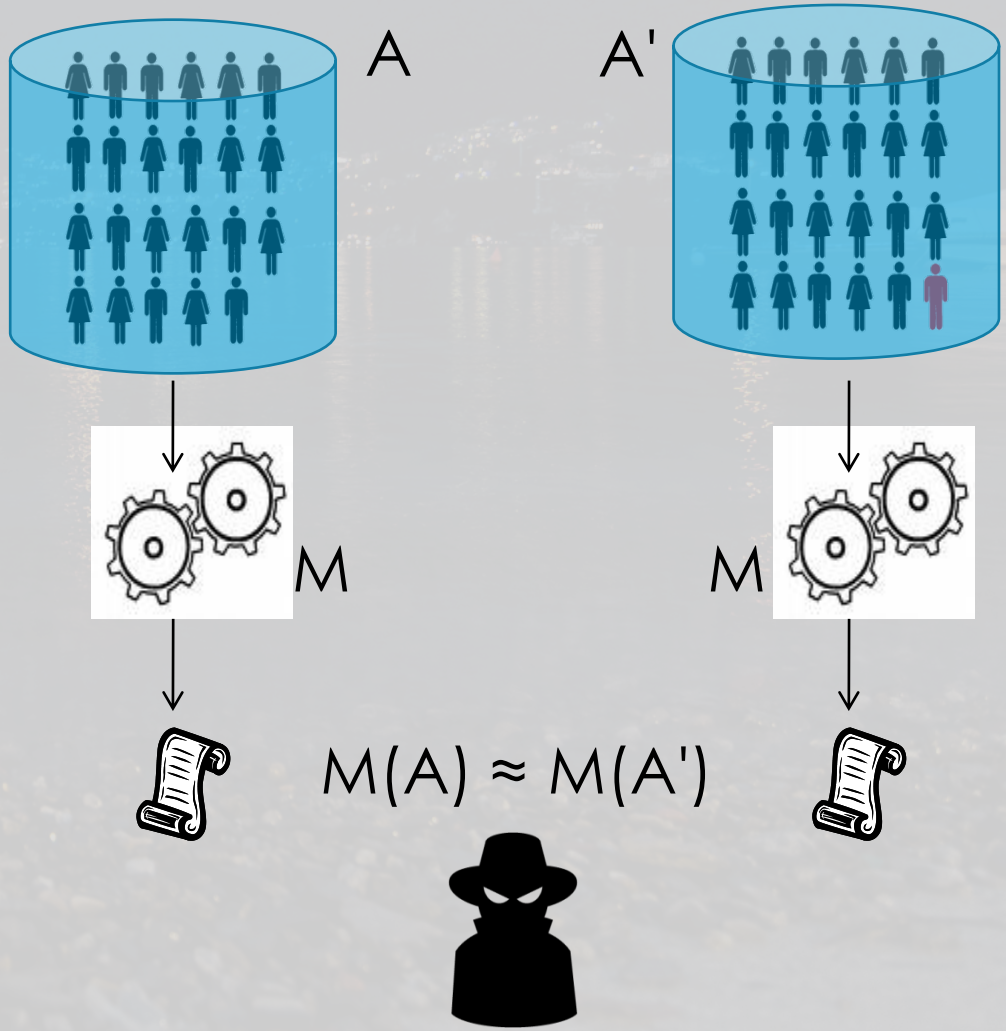


Extreme obfuscated – extreme predicted

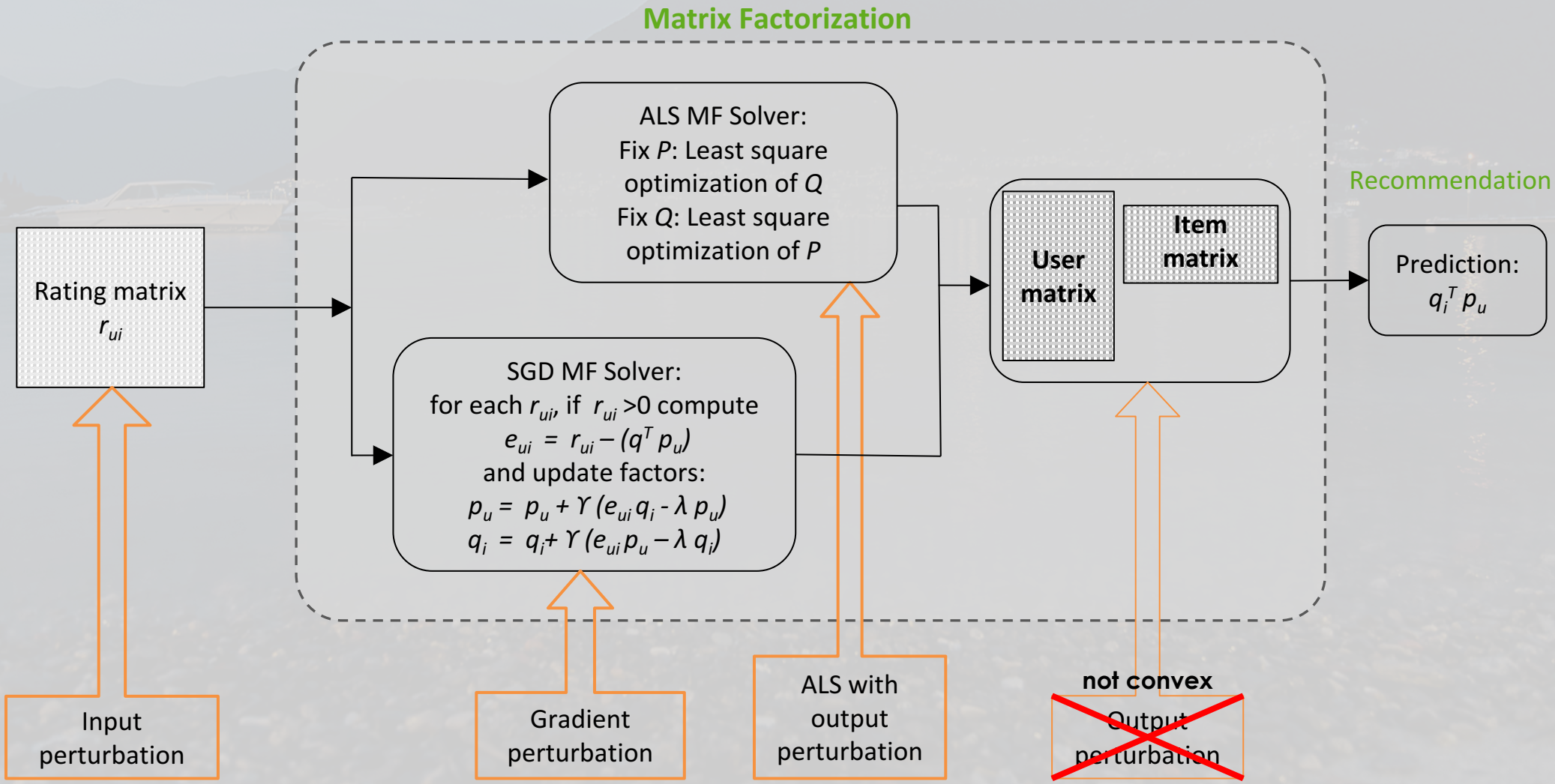
- Again distribution policy
- Normalized effect on error
  - With respect of obfuscation=0 that is flat
- Obfuscation effect
  - Is manageable on moderate and stronger on extreme ratings
  - Is stronger on negative than on positive extreme ratings
    - Observation: greater volume of positive than negative ratings
- Extreme ratings are important for accurate recommendations and they are deemed more sensitive
  - Even harder to optimize privacy and accuracy at the same time

Differential privacy - intuition

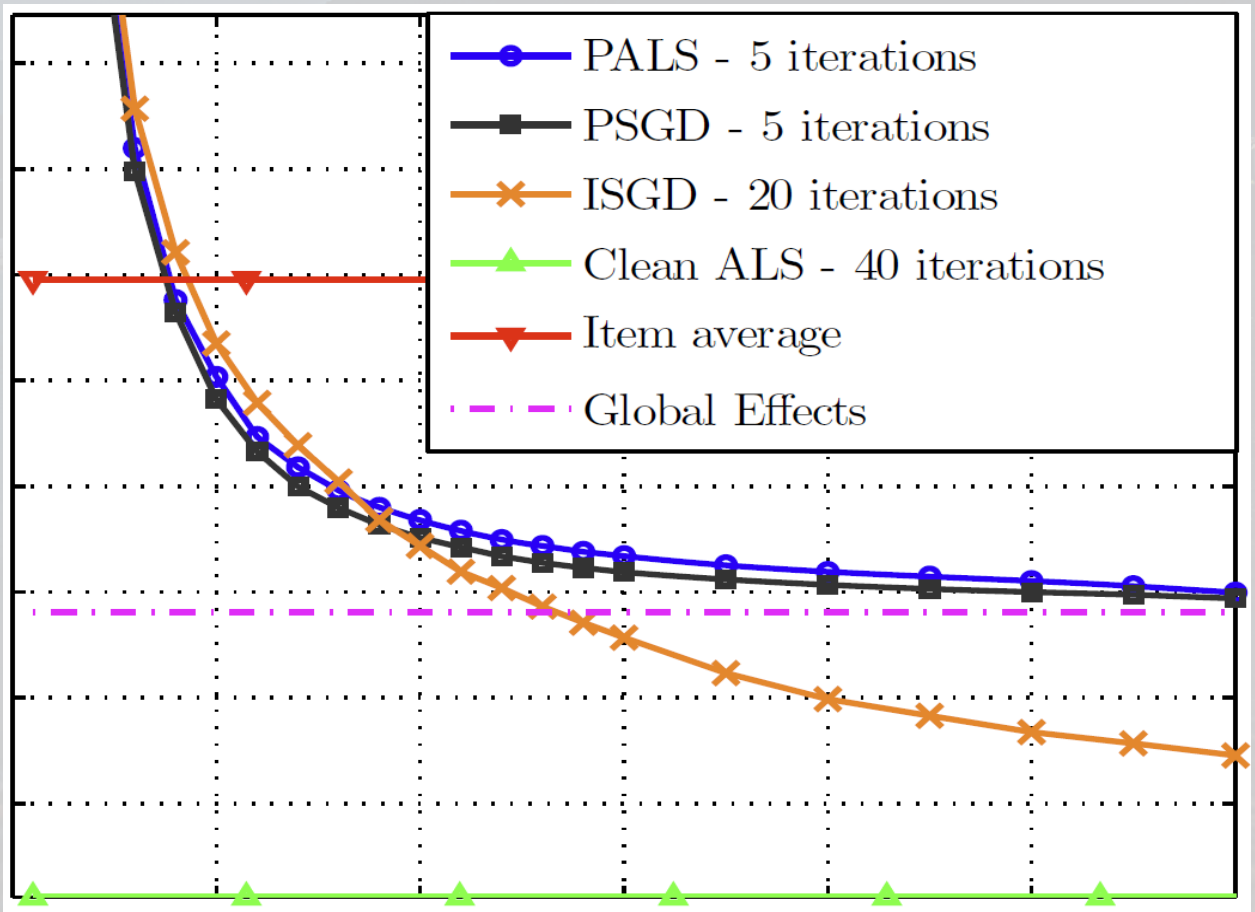
- Provable privacy property
  - Similar inputs induce similar distributions of outputs
- Presence of a single input in the data cannot be identified
  - Defined with probabilistic parameter ε
  - Privacy inversely correlated with ε
- Add noise to data
  - Tune the amount of noise to ε



Differentially private MF



Differentially private MF variants



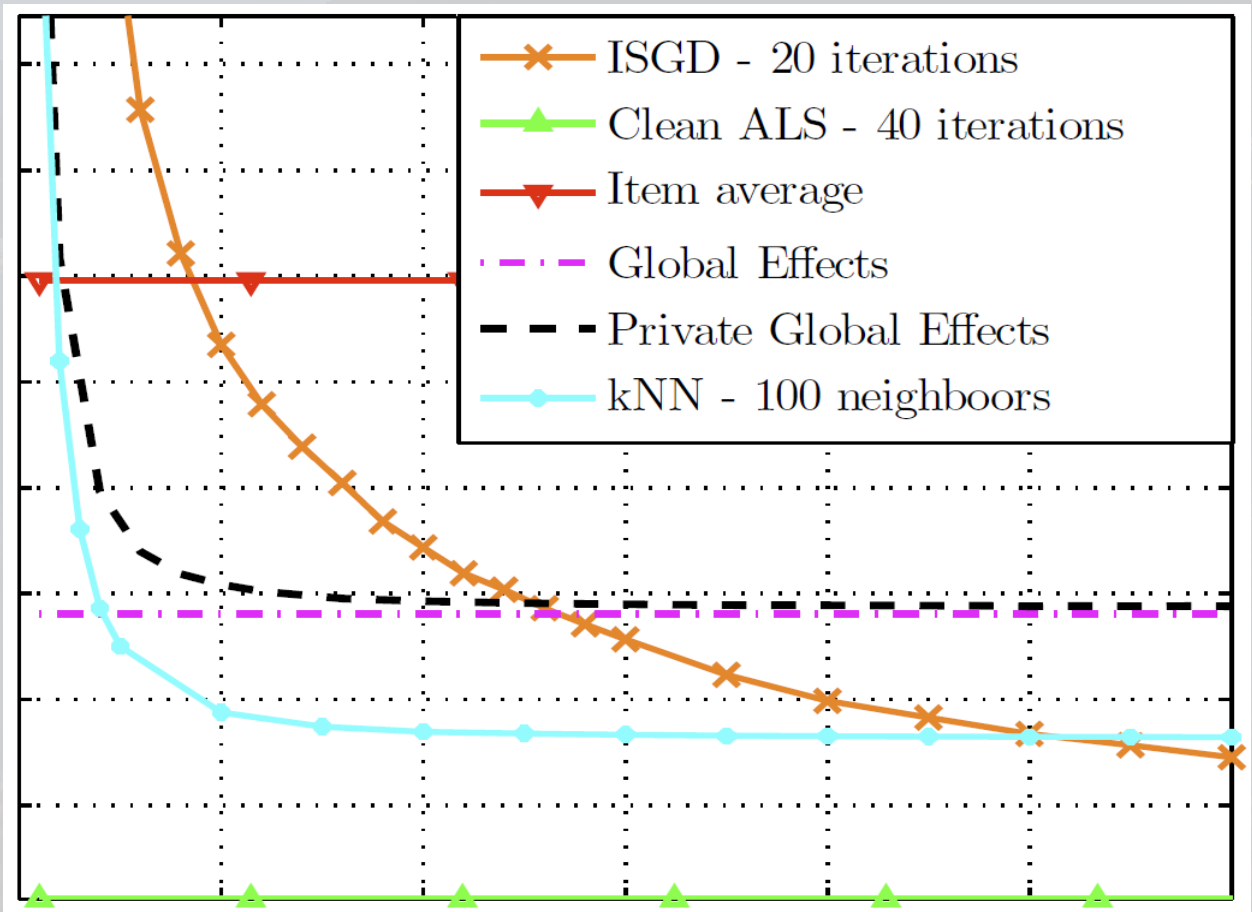
strong noise,
good privacy

weak noise,
poor privacy

Differentially private MF variants

- Error decreases as the noise gets weaker
  - Error decreases as privacy deteriorates
- error(input-perturbation) < (error(private-SGD) ≈ error(private-ALS))
- error(input-perturbation) is approaching error(non-private-MF)
  - Achieved at high values of ε
  - Weak noise, poor privacy levels

Comparison with private CF



strong noise,
good privacy

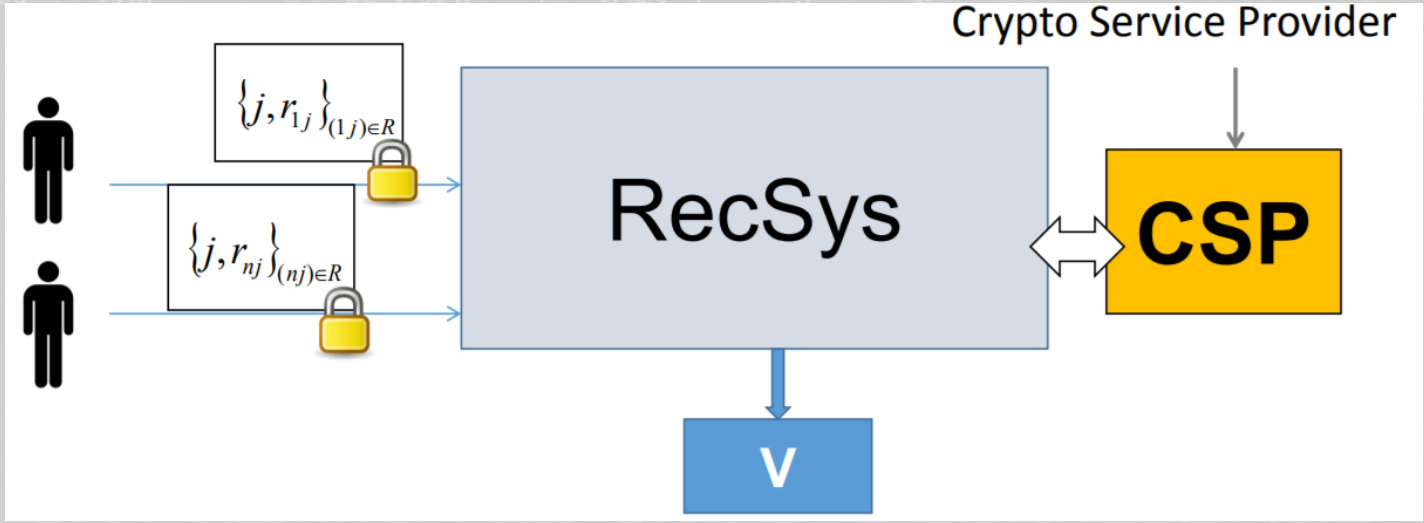
weak noise,
poor privacy

Comparison with private CF

- Regular non-private comparison
  - MF is superior to CF
  - error(non-private-MF) < error(non-private-CF)
- With strong noise, good privacy
  - error(private-MF) > error(private-CF)
- With weak noise, poor privacy
  - error(private-MF) < error(private-CF)
- Hard to optimize privacy and accuracy at the same time

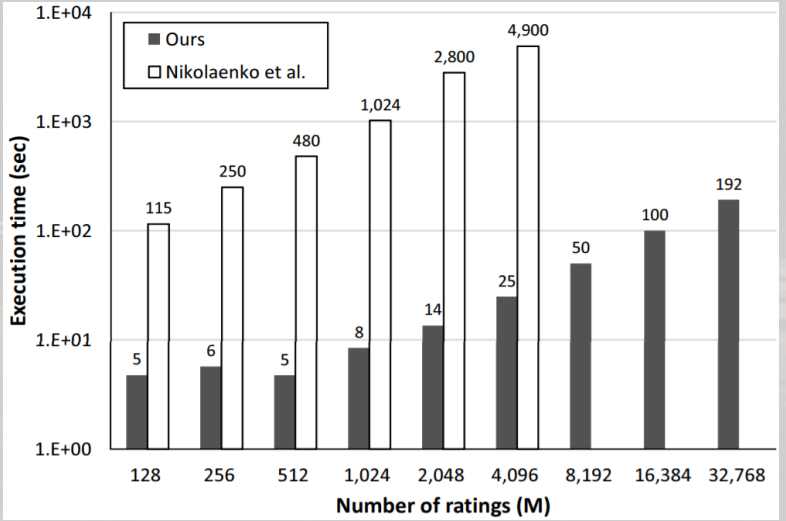
Cryptography

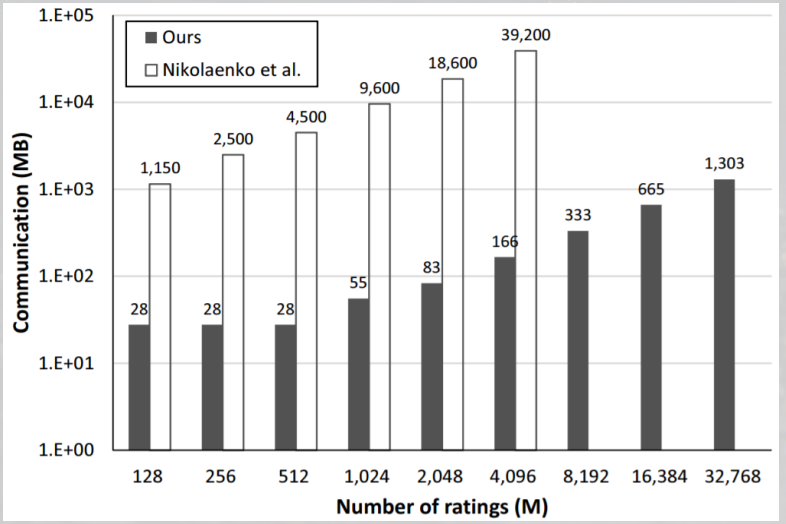
- Several works on encrypted MF
  - Implemented using homomorphic and fully homomorphic encryption
    - Garbled circuits
  - Formal guarantees
- Key characteristics
  - Recommender learns
    - latent vectors
    - number of user ratings
  - Recommender does not learn
    - individual ratings
    - items rated by user



Cryptography – evaluation

- High accuracy
  - Very close to non-private MF
- But very expensive
  - Even for toy datasets
    - Long execution time
    - Heavy communication overheads
- Inapplicable in practical recommenders
- Privacy and accuracy are optimized
  - But recommender is not practical





User-centric solutions

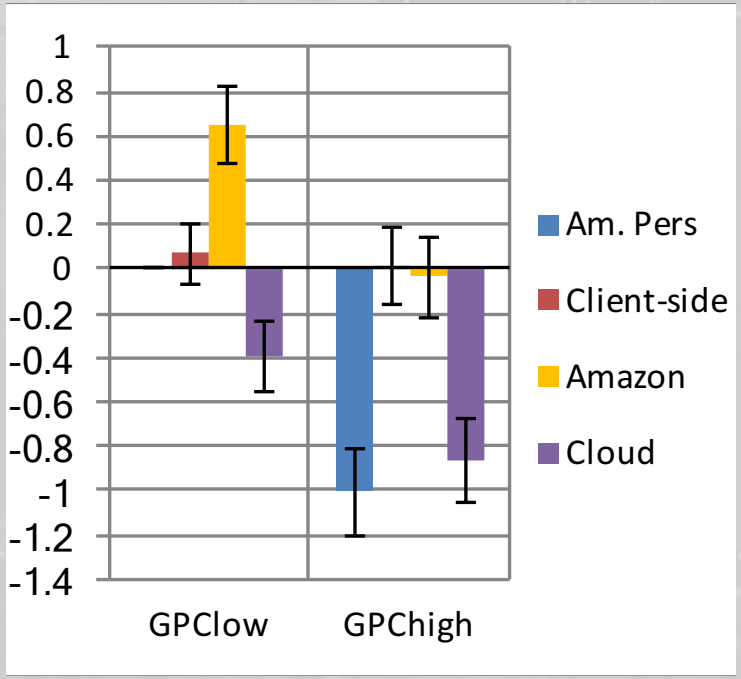
- As engineers, we would first go for a **technical** solution
  - Homomorphic encryption
  - Differential privacy
  - Data obfuscation
  - Client-side personalization
- Problem: How to **convince users** that this is better?

Why technical solutions don't always work

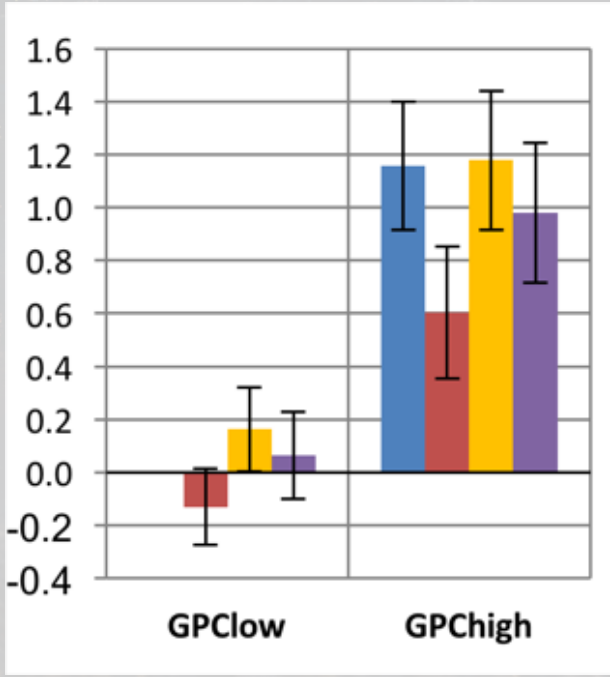
- Let's say you want to use users' data for personalization in a mobile app recommender…
- …which type of situation are users most comfortable with?
  - "Your data will be sent to, and analyzed by a company called American Personalization."
  - "All personalization happens client-side."
  - "Your data will be sent to, and analyzed by Amazon."
  - "Your data will be sent to, and analyzed in the cloud."

Why technical solutions don't always work

Perceived Privacy Protection



System-Specific Privacy Concerns



Why technical solutions don't always work

- Client-side personalization works, but only for people with **high overall privacy concerns**
- Client-side personalization brought up **other issues**:
  - "What happens if my phone gets stolen?"
  - "Can I lock my user model remotely so the thief can't get to it?"
  - "Is there a backup of my user model?"

Privacy is a human issue

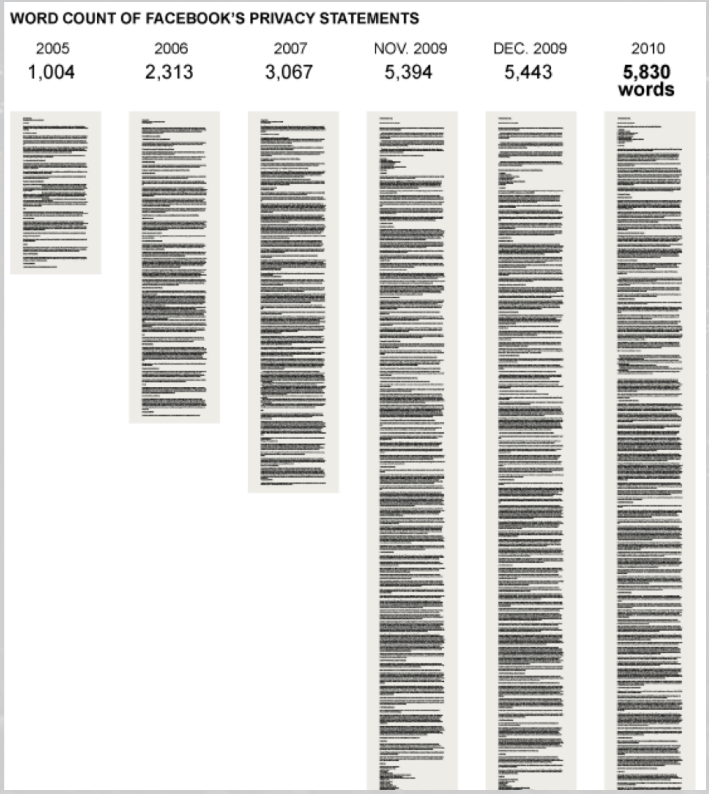
- The concept of privacy is an inherently human attitude
  - People make assumptions and form opinions about the collection, distribution and use of disclosed data
- Disclosure itself is a human behavior
  - In the end, people decide whether they want to disclose or not
  - Regardless of our technical solutions, we should help them with that decision
- So how can we help them?
  - Transparency and control
  - Privacy nudging
  - User-tailored privacy

Transparency and control

- **Privacy Calculus:** People weigh the risks and benefits of disclosure
  - Dr. Loewenstein: Is this really true?
- Prerequisites of the privacy calculus are:
  - being able to **control** the decision;
  - having adequate **information** about the decision.
- Transparency and control **empower** users to regulate their privacy at the desired level.

Why this doesn't work

- **Transparency paradox:**
- Simple privacy notices **aren't useful**, but detailed notices are **too complex**.
  - (Nissenbaum 2011)
- Dr. Loewenstein: $653 B opportunity costs of reading privacy policies

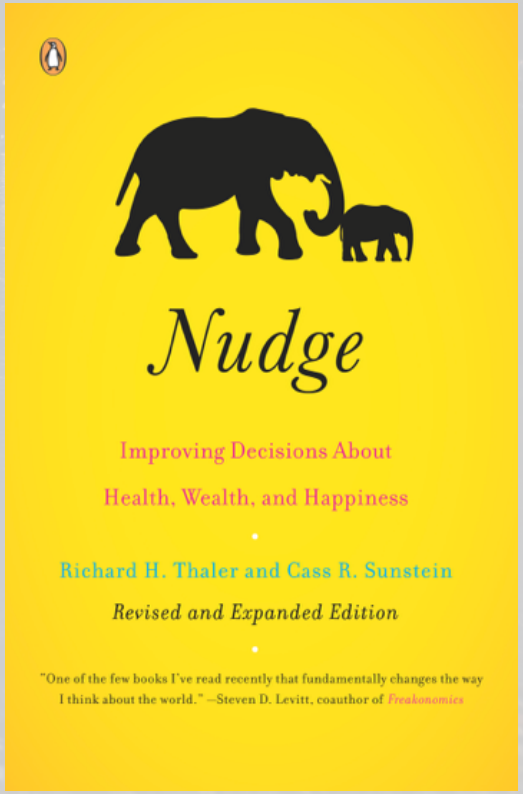


Why this doesn't work

- **Control paradox:**
- Consumers claim to want full control over their data, but they **eschew the hassle** of actually exploiting this control!
  - (Compañò and Lusoli 2010; Knijnenburg et al. 2013)
- **People are not good at control:**
- Dr. Loewenstein: misplaced confidences, strangers on a plane

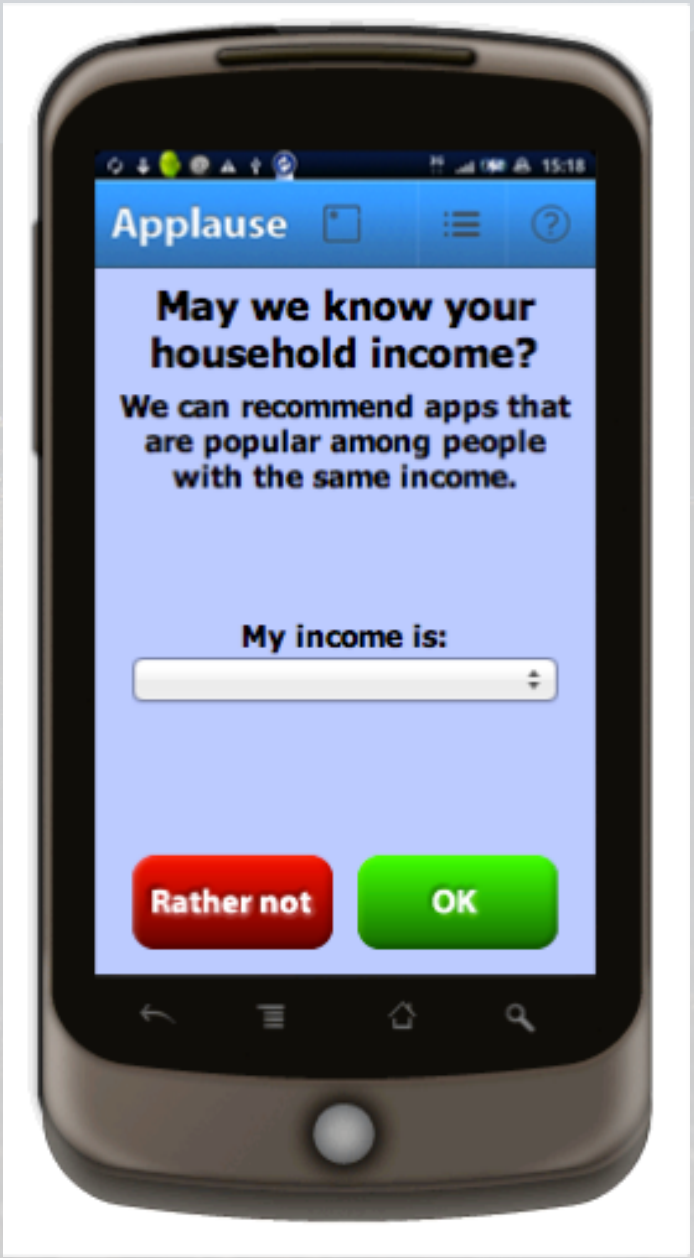
Privacy nudging

- Subtle yet **persuasive cues** that makes people more likely to decide in one direction or the other.
  - (Thaler and Sunstein 2008)
- Examples of nudges:
  - **Justification:** a succinct reason to disclose or not disclose a certain piece of information.
  - **Default:** make the best action the easiest to perform.



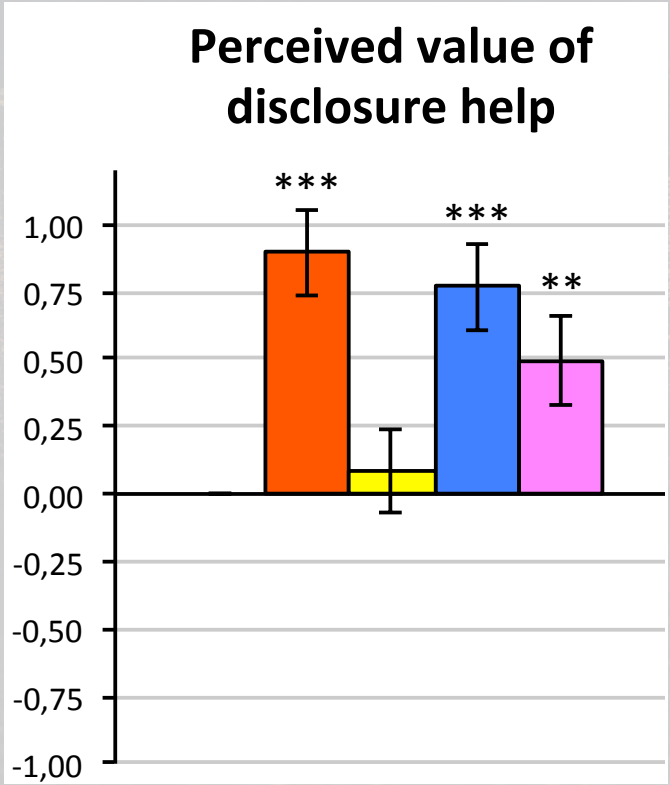
Testing justifications

- 5 justification types
  - None
  - Useful for you
  - Number of others
  - Useful for others
  - Explanation



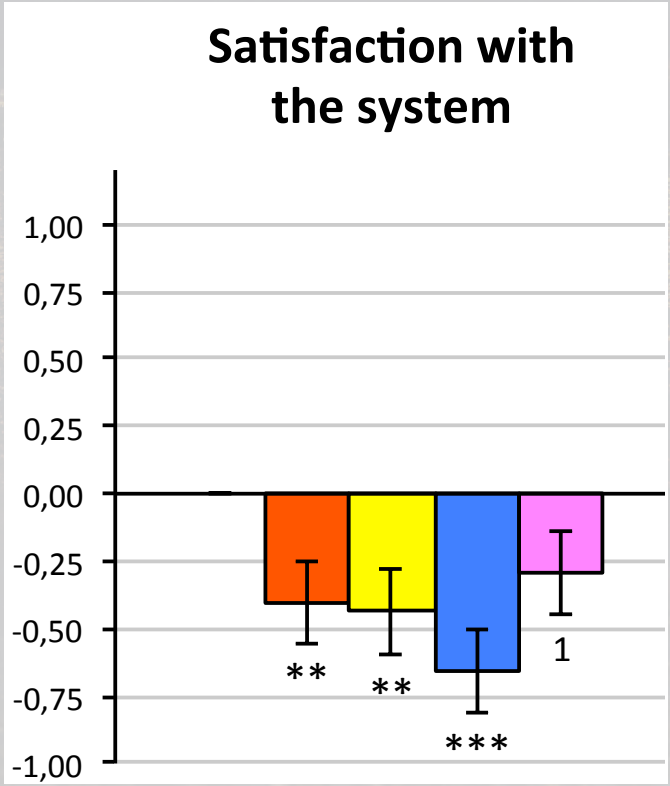
Results

- Perceived value of disclosure help:
  - 3 items, e.g. "The system helped me to make a tradeoff between privacy and usefulness"
- Higher for all except "number of others"



Results

- Satisfaction with the system:
  - 6 items, e.g. "Overall, I'm satisfied with the system"
- Lower for any justification



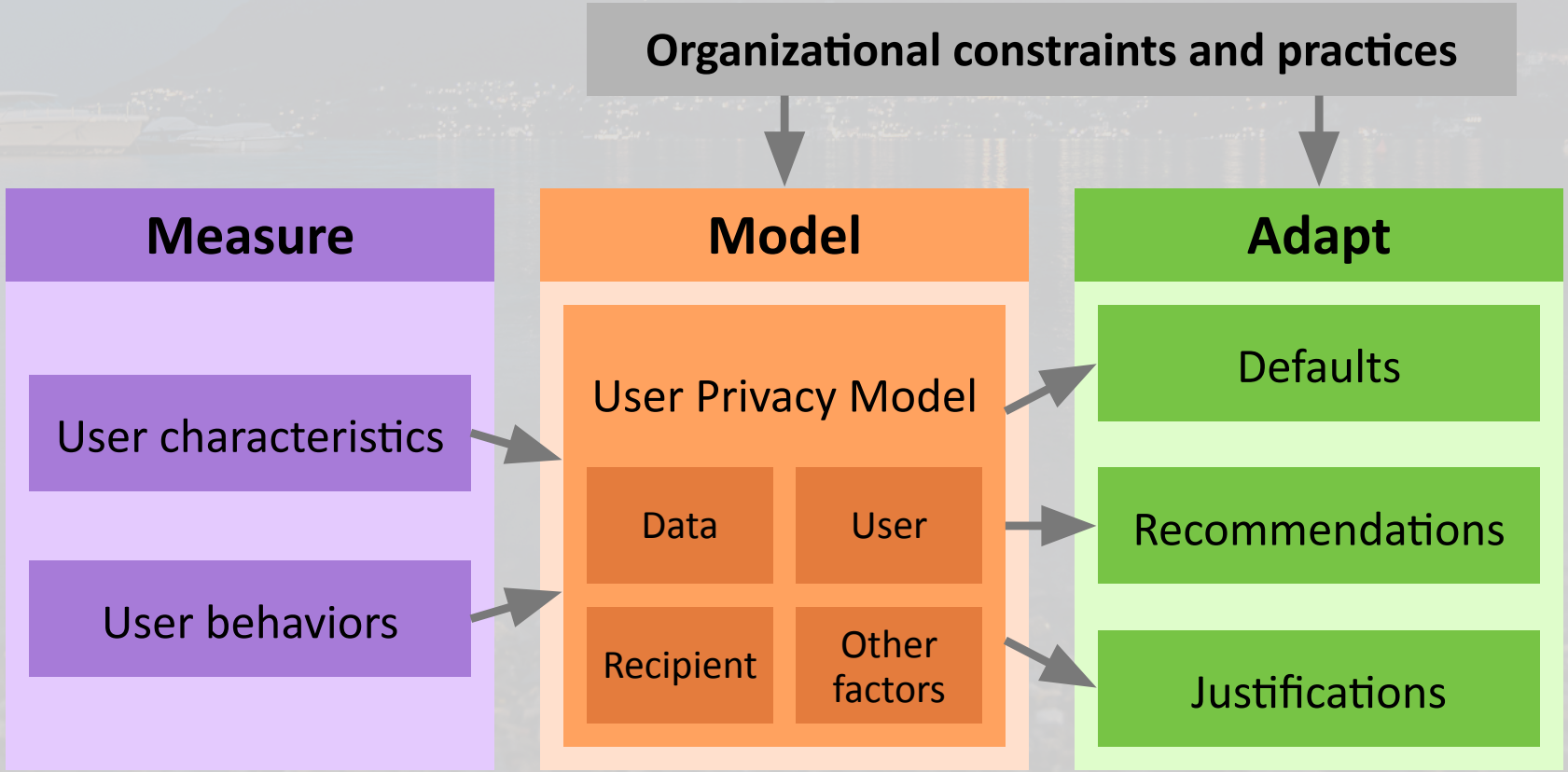
Why nudges don't work

- What is the **"right" direction** of a nudge?
  - **More disclosure:** better personalization, but some may feel tricked.
  - **More private:** less threat, but harder to enjoy the benefits of disclosure.
  - Going for the **average** (e.g. "smart default", Smith et al. 2013): impossible, because people vary too much.
- **Solution: move beyond the one-size-fits-all approach!**

User-tailored privacy

- Idea: Give people privacy **recommendations**!
- "Figure out what people want, then help them do that."

User-tailored privacy



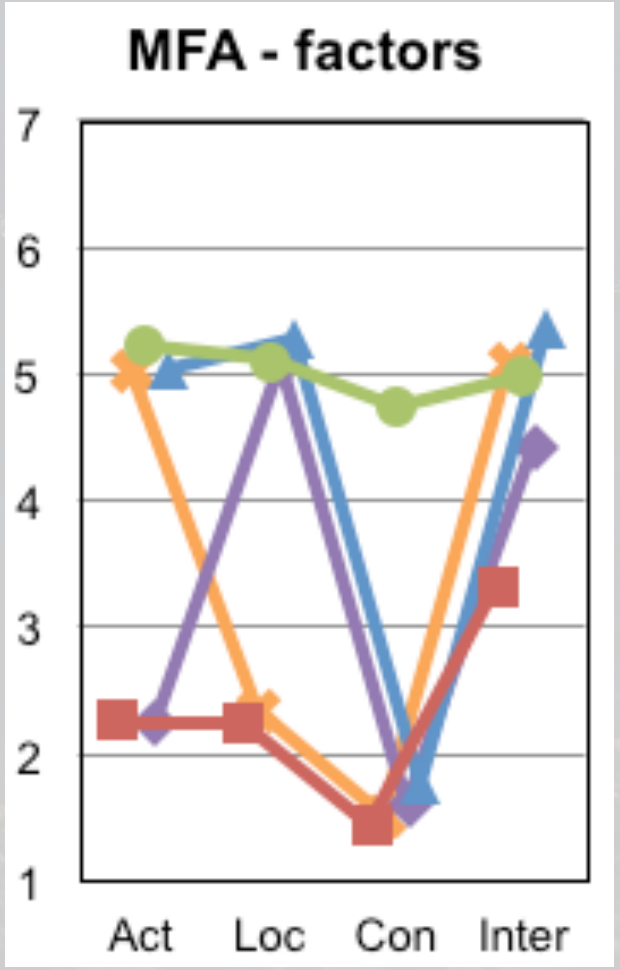
Measure: information ("what")

Type of data	ID	Items
Facebook activity	1	Wall
		updates
		d links
		tes
		tos
Location		town
		n (city)
		te/province)
Contact info		reet address)
		number
		address
Life/interests		s views
		te movies, etc.)
		k groups

"What?"
=
Four
dimensions

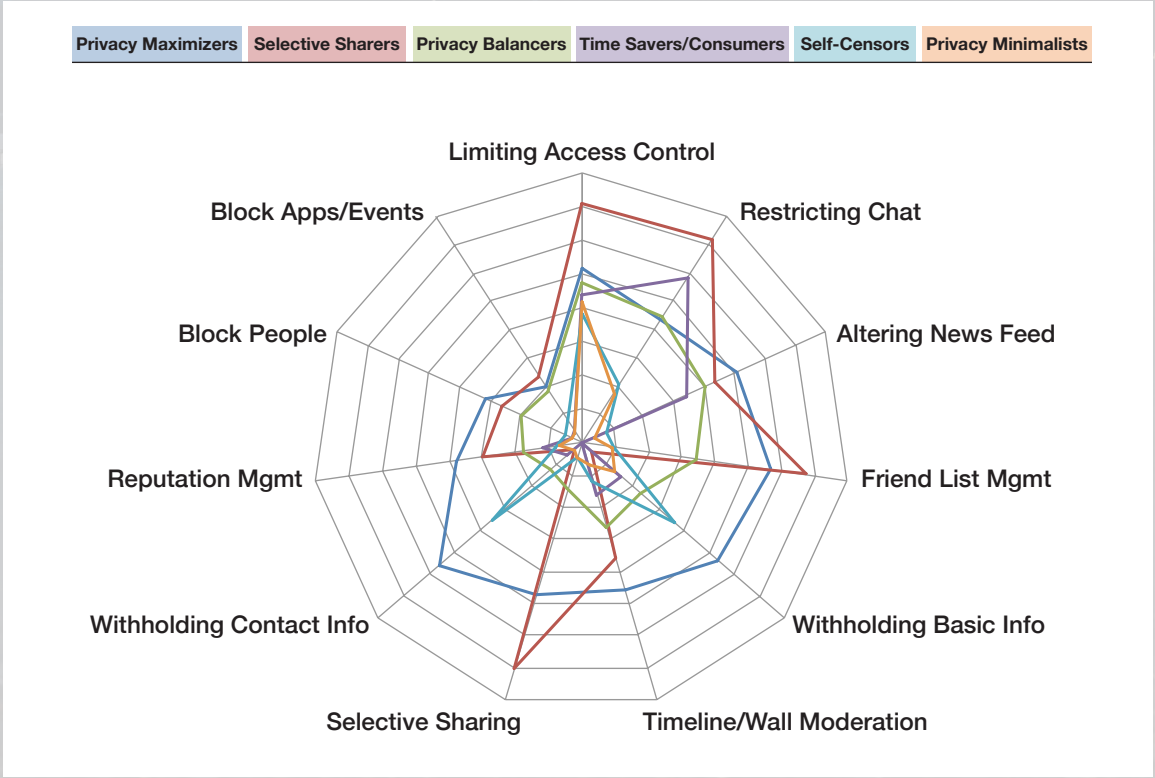
Measure: user ("who")

"Who?"
=
Five
disclosure
profiles

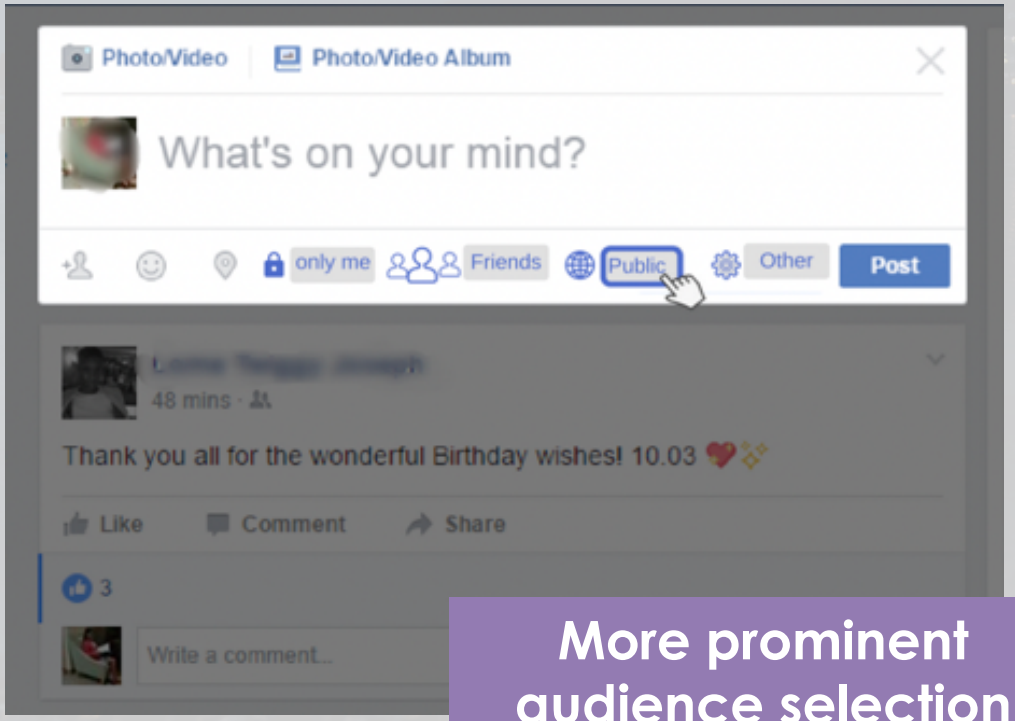


- 159 pps tend to share little information overall (LowD)
- 26 pps tend to share activities and interests (Act+IntD)
- 50 pps tend to share location and interests (Loc+IntD)
- 65 pps tend to share everything but contact info (Hi-ConD)
- 59 pps tend to share everything

Measure: user ("who")



Adapt: the interface



More prominent audience selection (selective sharers)

More prominent timeline moderation (time savers)

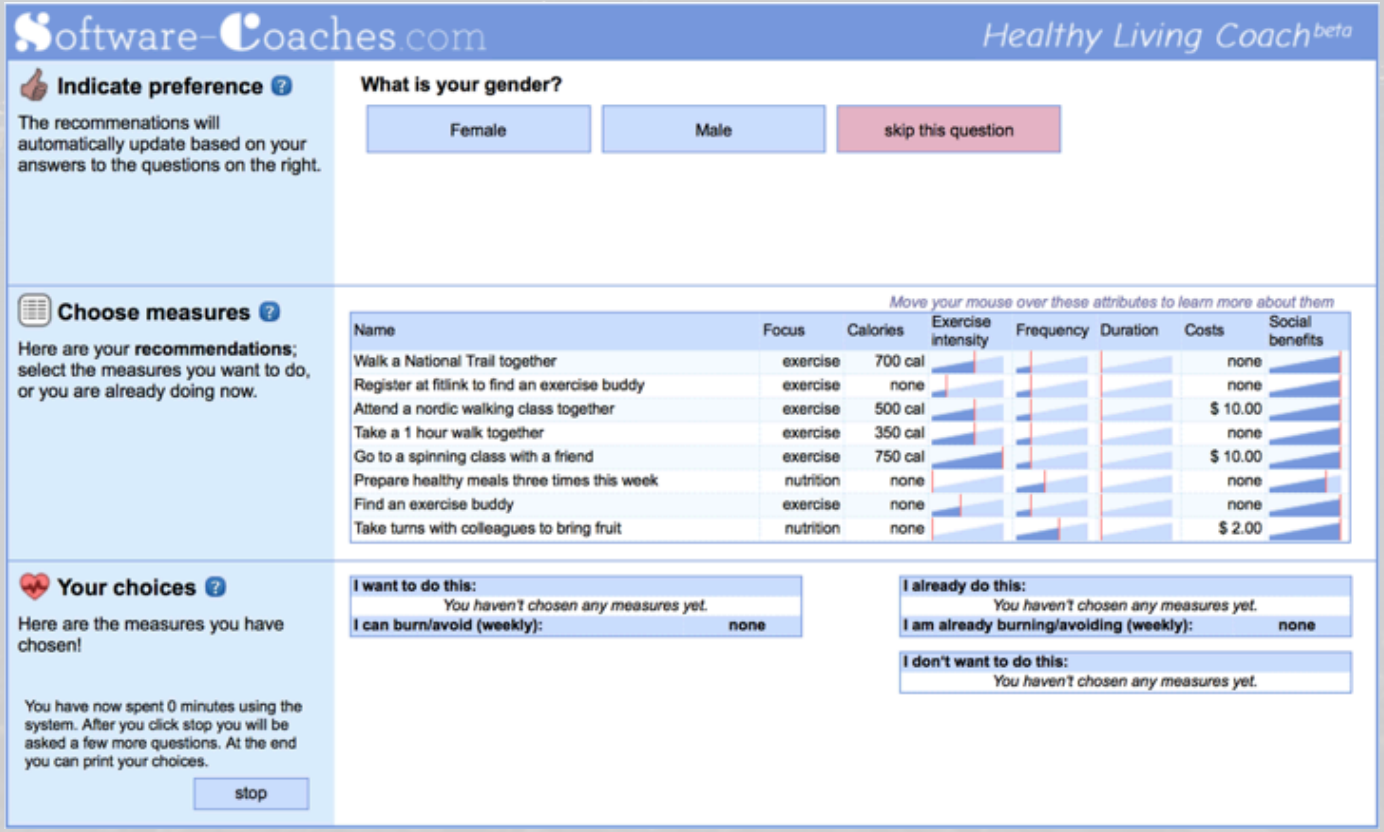
Adapt: the justification

- What if we gave different types of users different types of justifications?

Adapt: the visible options

- What if we showed a subset of location-sharing options based on the user's evaluation of the activity?

Adapt: the request order



Privacy-personalization paradox revisited

- Demographics can be used as a proxy for preferences.
  - Needed: an algorithm that translates answers to demographic questions into attribute weights.
  - Based on these weights I can then recommend items as usual.
- Demographics-based PE:
  - May be most beneficial for domain novices (known and easy to report).
  - May be more privacy-sensitive than other PE-methods (Ackerman et al. 1999).
- Which item to ask first?
  - Not all items are equally **useful** to the recommender.
  - Not all demographic items are equally **sensitive**.
  - Not everyone is equally **private** regarding their demographics.

Adaptive request order

$$u_o = \sum_{r_{oa}} \frac{v_r}{d_{an}}$$ where $$d_{an} = abs\left(w_{an} - \bar{w}_n\right) + .0001$$

$$p_{ni} = \frac{e^{\beta_n - \delta_i}}{1 + e^{\beta_n - \delta_i}}$$

Useful -ness

$$r_i = \begin{cases} u_i & if \quad \delta_i < \alpha, \\ -\delta_i & if \quad \delta_i > \alpha. \end{cases}$$

Sensi -tivity

Trade-off

Tendency

$$\beta_n = \text{mean}_n(\delta) + \sqrt{1 + \text{var}_n(\delta)/2.9} * \ln\left(\frac{|D_n|}{|L_n| - |D_n|}\right)$$ and $$\alpha_n^H = \beta_n - 1.5$$

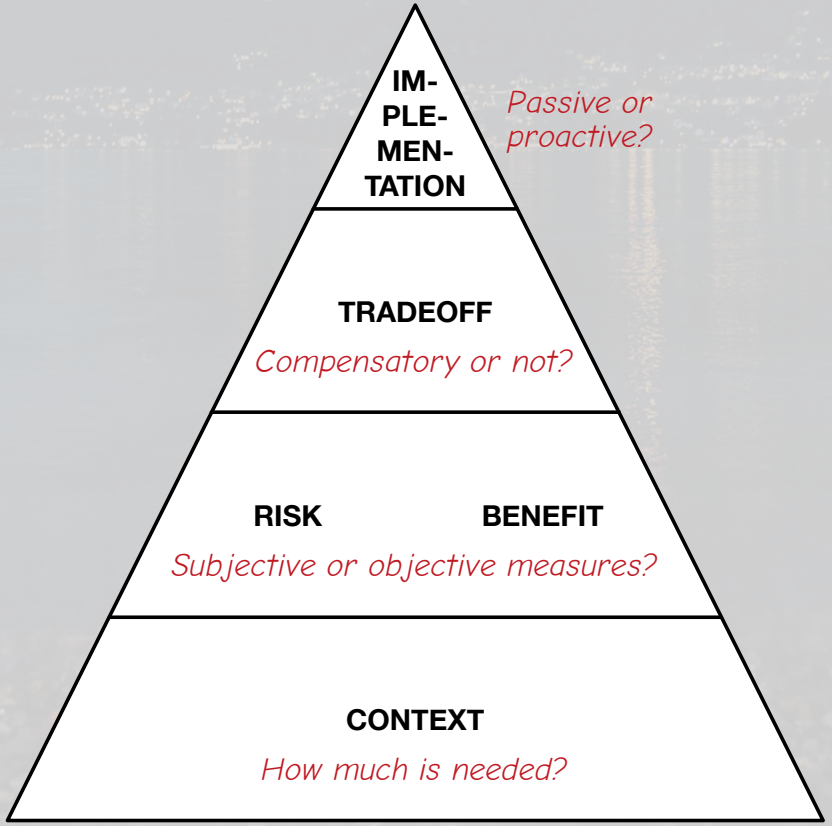
Outcome

- The adaptive request order did not result in the hypothesized benefits.
- However, other (static) versions that automatically traded off usefulness and sensitivity **did** improve users' experience.
  - Reserved optimism: Automatic means to relieve some of the burden of controlling one's privacy settings are still promising.
- Future work may further improve the truly adaptive versions.
  - Goal: a universal method that works for **all kinds of users**.

Societal impact

- User-tailored privacy:
  - Relieves some of the burden of controlling privacy, while at the same time respecting each individual's preferences
  - Provides **realistic empowerment**: the right amount of transparency and the right amount of control
  - Refrains from making moral judgments about what the "right" level of privacy should be
- The best way forward to support people's privacy decisions!

Future research questions



Summary

- Conflict between privacy and personalization goals
- Take home messages
  - No magical solution on the algorithmic front
  - Since there is no magical solution and since everyone is different, we have to help people in making this trade-off
- 40K feet future directions
  - Algorithmic solutions are not orthogonal
  - Educate users to value and preserve their privacy

Emerging topics/areas

- Social Web and information leaks
- Personality and emotion modelling
- Cross-domain / multi-source recommendations
- Mobile recommendations and smart environments
- Explanations that leak
- Group recommenders that leak
- Usable privacy
- Holistic privacy solution
- Real-world feasibility

References

- Differentially private MF: Berlioz et al. “Applying differential privacy to MF”, RecSys 2015
- Hierarchical CF: Berkovsky et al. “Hierarchical neighborhood topology for privacy enhanced CF”, PEP 2006
- Privacy-personalization paradox: Berkovsky et al. “The impact of data obfuscation on the accuracy of collaborative filtering”, ESWA 2012
- User communities: Canny. “CF with privacy via factor analysis”, SIGIR 2002
- Differential privacy: Dwork et al. “Differential privacy under continual observation”, STOC 2010
- **Book chapter: Friedman et al. “Privacy Aspects of Recommender Systems”, RecSys Handbook 2015**

References

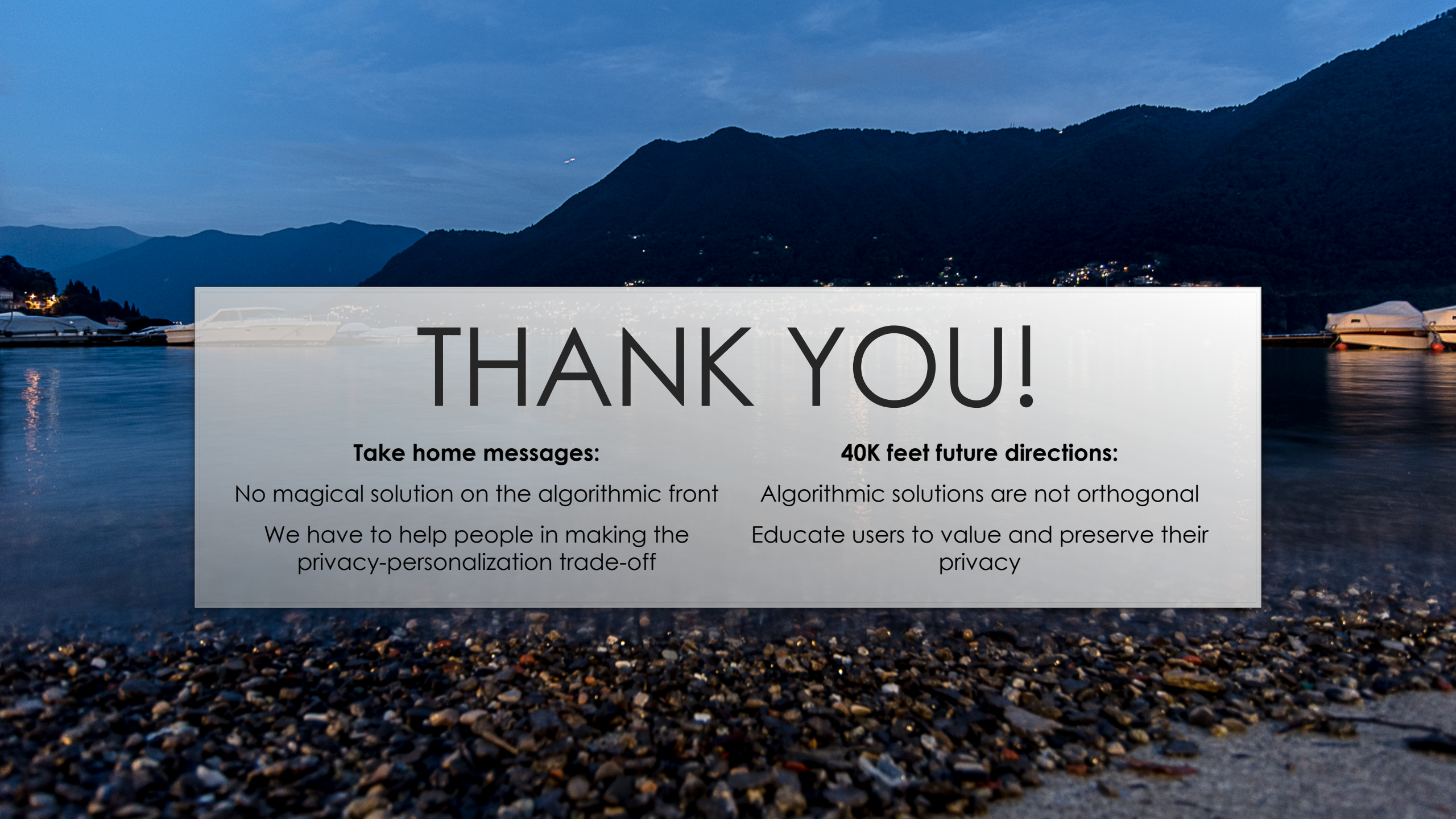
- FB likes: Kosinski et al. “Private traits and attributes are predictable from digital records of human behaviour”, PNAS 2013
- Differentially private CF: McSherry-Mironov. “Differentially private RecSys: Building privacy into the Netflix prize contenders”, KDD 2009
- Netflix de-anonymization: Narayanan-Shmatikov. “Robust de-anonymization of large sparse datasets”, In: IEEE S&P 2008
- Encrypted MF: Nikolaenko et al. “Privacy preserving MF”, CCS 2013.
- Privacy-personalization survey: Toch et al. “Personalization and privacy: a survey of privacy risks and remedies”, UMUAI 2012
- MF with no user data: Vallet et al. “MF without user data retention”, PAKDD 2014

References

- Privacy-personalization paradox: Knijnenburg et al. “Receiving Recommendations and Providing Feedback”, EC-Web 2010
- Client-side personalization: Kobsa et al. “The effect of personalization provider characteristics on privacy attitudes and behaviors”, JASIST 2016
- Testing privacy nudges: Knijnenburg and Kobsa “Making decisions about privacy”, TiiS 2013
- User-tailored privacy: Knijnenburg “Privacy? I can’t even!”, IEEE S&P 2017
- Privacy profiles (1): Knijnenburg et al. “Dimensionality of information disclosure behavior”, IJHCS 2013
- Privacy profiles (2): Wisniewski et al. “Making privacy personal”, IJHCS 2017

References

- Adapt the interface: Wilkinson et al. “User-Tailored Privacy by Design”, USEC 2017
- Adapt the justification: Knijnenburg and Kobsa “Helping users with information disclosure decisions”, IUI 2013
- Adapt the visible options: Knijnenburg and Jin “The persuasive effect of privacy recommendations for location sharing services” SigHCI 2013
- Adapt the request order: Knijnenburg “A user-tailored approach to privacy decision support” dissertation 2015
- Future research questions: Knijnenburg et al. “Death to the privacy calculus?” CSCW workshop 2017



THANK YOU!

Take home messages:

No magical solution on the algorithmic front

We have to help people in making the privacy-personalization trade-off

40K feet future directions:

Algorithmic solutions are not orthogonal

Educate users to value and preserve their privacy