privacy

EEECE 571B "Computer Security"

Konstantin Beznosov

what privacy is and is not?

what is it not?

- the right to be left alone
  - I don’t want to be alone, but I still want privacy
- anonymity
  - If I’m anonymous, I don’t need privacy. It’s when I’m ME that I’m worried.
- security
  - The security of my insurance company can be perfect, and my claims adjuster can still gossip about my medical condition.
- me controlling information about myself
  - I have no right to do this if the information is true, so I have no recourse absent prior consent, so I can only control information that comes from me, which is not sufficient to protect my privacy.
- secrecy

what is privacy then?

“the ability* to lie about yourself and get away with it.”

- Bob Blakley

*Not “Right”

what about <u>privacy right</u> then?

"The right to the ability to lie about yourself and get away with it."

- Bob Blakley

a more operational definition

"... a process of interpersonal boundary control that paces and controls interaction"

Altman, I. 1975. The Environment and Social Behavior. Privacy - Personal Space - Territory - Crowding. Brooks-Cole Publishing Company, Monterey, CA, USA.

boundaries

- disclosure, identity, temporality

involve

- privacy-publicity balancing
- management of self-presentation
- the sequence disclosures form over time

Palen, L. & Dourish, P. 2003. Unpacking "privacy" for a networked world. Proc. CHI'03. ACM Press, Fort Lauderdale, FL, USA.

boundary regulation of privacy and publicness in OSNs

Airi Lampinen, Vilma Lehtinen, Asko Lehmuskallio, and Sakari Tamminen, "**We're in it together: interpersonal management of disclosure in social network services**," In Proceedings of the 2011 annual conference on Human factors in computing systems (CHI '11), pp. 3217-3226.

background

- users cannot control the content others disclose about them
- “research questions”
  - what kind of interpersonal boundary regulation concerns OSN users have?
  - what kind of strategies they apply?
  - how do individuals manage not only their own privacy and publicness but also that of their peers?
- OSN vs. SN
  - interactions in OSNs differ from face-to-face settings in their persistence, replicability, scalability, and searchability
  - instead of being fleeting and offering the possibility to forget, interactions in SNSs leave enduring traces

methodology

- data collection
  - semistructured individual interviews (11+13)
  - 5 focus groups (18 participants)
    - probes based on individual interviews and press stories
- data analysis
  - focus on 1) concerns related to and 2) strategies for interpersonal boundary regulation
  - open-coding of concerns to interpersonal boundary regulation
  - grounded theory with prior key findings “as loose interpretive anchors”
- participants
  - 27
    - undergraduates in technology studies
    - (mostly international) graduate students in industrial arts an design
  - age: early 20’s & 30‘s
  - 17 males
  - regular users of FB and other OSNs
  - good enough?

examples of strategies from the data

Strategy Type	**Preventive**	**Corrective**
Individual	• Creating separate audience zones (sharing content groupwise, sharing content according to proximity category, or using multiple accounts – in one or more services) • Adjusting privacy settings to disable disclosure (of certain types of content and/or to certain people) • Choosing a private communication channel (private messages) • Using deliberate wordings and tones in (semi-)public posts • Avoiding publicizing content that could be problematic • Withdrawing from publicizing altogether • Regulating one's behavior offline • *Considering trust and trustworthiness* • *Applying rules of thumb in decisions on sharing*	• Deleting comments (in one's profile and/or comments one has posted elsewhere) • Untagging photos • *Interpreting a potentially problematic issue to be non-serious*
Collaborative	• Negotiating and agreeing on "rules of thumb" concerning sharing with other SNS-users • Asking for approval before disclosing content from those involved	• Asking another person to delete content • Reporting inappropriate content to service administrators • Supporting a non-serious interpretation • *Interpreting content to be non-serious*

mental strategies presented in italics

results & discussion

- augmentation of the prior set of dimensions of strategies
  - behavioural and mental
  - preventive and corrective
  - **individual and collaborative**

- majority of collaborative were corrective
- (support for) collaborative, preventive strategies needed
- corrective strategies risk not being socially feasible or effective
  - socially awkward
  - ineffective (the open barn door phenomena)
  - can even draw extra attention to the exact thing that was supposed to be swept under the carpet

conclusions

- predicting the effects of one's disclosure on another SNS-user's boundary regulation can be practically impossible
- Blunders in boundary regulation seem to derive often from the <u>difficulty of estimating how something would be interpreted</u> in others' varied networks.
- the strategies are are often tightly enough integrated with routines of everyday interaction to be employed in an almost automatic manner
  - not necessarily reflexively pondered
- it is not sufficient to focus on how individuals manage what they disclose of themselves online
  - disclosing content related to others
- possible improvement to technology
  - preview space wherein boundaries could be negotiated collaboratively within a group whom the content concerns

privacy risks in collaborative filtering

Calandrino, J.A.; Kilzer, A.; Narayanan, A.; Felten, E.W.; Shmatikov, V.; , "“**“You Might Also Like:” Privacy Risks of Collaborative Filtering**,” Security and Privacy (SP), 2011 IEEE Symposium on , pp.231-246, 22-25 May 2011

background

- recommendations by recommender systems
  - user-to-item: suggests items to an individual user based on its knowledge of the user's behavior
  - user-to-user: helps users find similar users
  - item-to-item: given an item, the system suggests similar items
  - item-to-user: list users who are strongly associated with a given item
- collaborative filtering
  - identifies relationships between items based on the preferences of all users
  - traditional: item-based
  - popular: user-based
    - generates recommendations using <u>item similarity scores</u> for <u>pairs of items</u>, which are based on the likelihood of the pair being purchased by the same customer

attack model

- passive inference attack
- attacker
  - has access to the public outputs of the recommender system
    - item similarity lists, item-to-item covariances, and/or relative popularity of items
  - observes the system over time and can thus capture changes in its outputs
    - Note: each update incorporates the effects of many transactions
- no access to PII or individual transactions
- auxiliary information
  - for some users, a subset of their transaction history is available to the attacker
  - sources: target system, users revealing the information via third parties, other sites leak partial information about users' transactions
- success criterion
  - **an inference attack is successful if it enables the attacker to learn transactions which are not part of the auxiliary information**

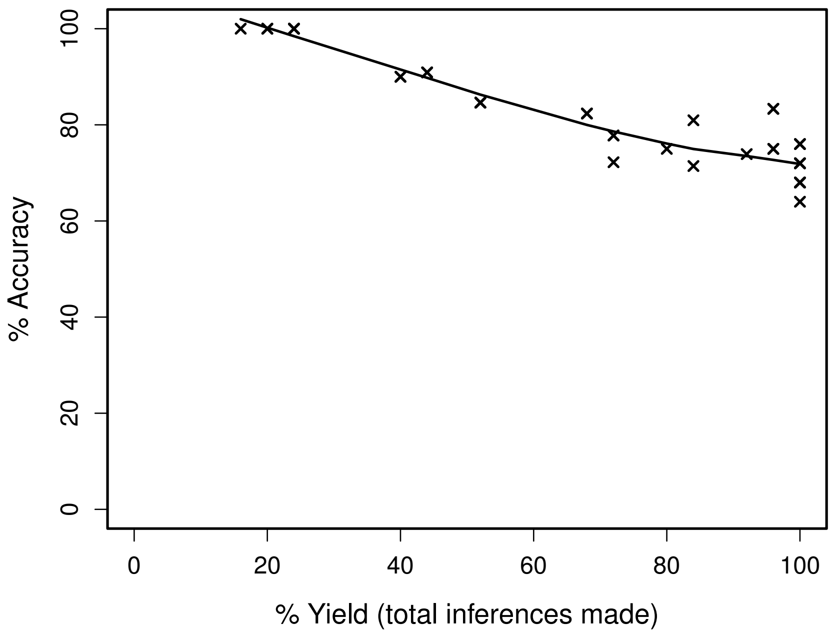
inference attack on related-items lists

- monitor the similarity list(s) associated with each auxiliary item (i.e., item that he knows to be associated with the target user)
- look for items which either appear in the list or move up, indicating increased "similarity" with the auxiliary item
- If the same target item *t* appears and/or moves up in the related-items lists of a sufficiently large subset of the auxiliary items, then *t* has been added to the user's record
- movements of obscure items give more information

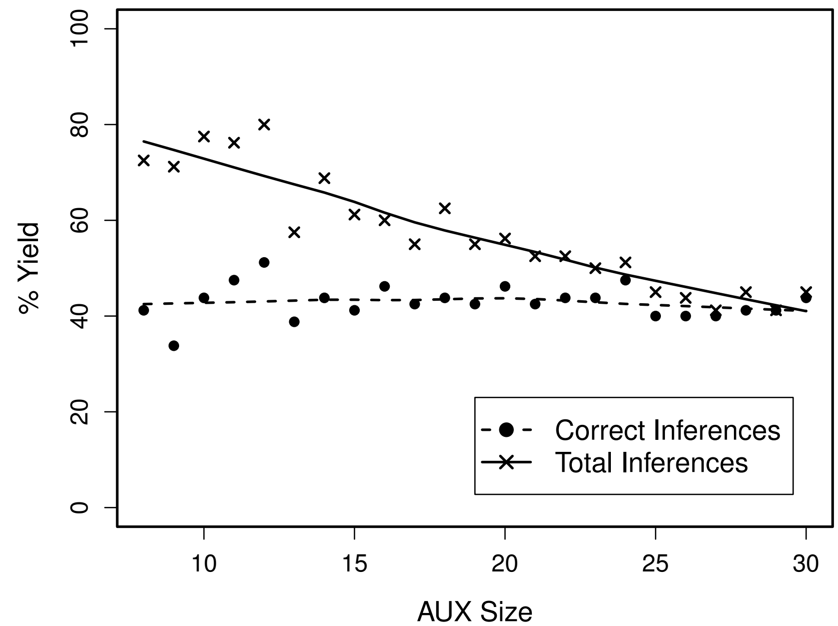
inference attack on kNN recommender systems

- active attack on
- the k-nearest neighbour (kNN) recommendation algorithm
  - for each user *U*, it finds the *k* most similar users according to some similarity metric
  - ranks all items purchased or rated by one or more of these k users according to the number of times they have been purchased and recommends them to U in this order
- the recommendation algorithm and its parameters are known to the attacker
- auxiliary information
  - U's partial transaction history, i.e., attacker already knows *m* items that *U* has purchased or rated
- attack
  - creates k sybil users
  - populates each sybil's history with the *m* items present in U's history ($m \approx O(\log N)$)
  - *k* nearest neighbors of each sybil will consist of the other $k - 1$ sybils and *U*
  - any new item on the list and is not one of the *m* items from the sybils' artificial history must be an item that U has purchased

results: Hunch

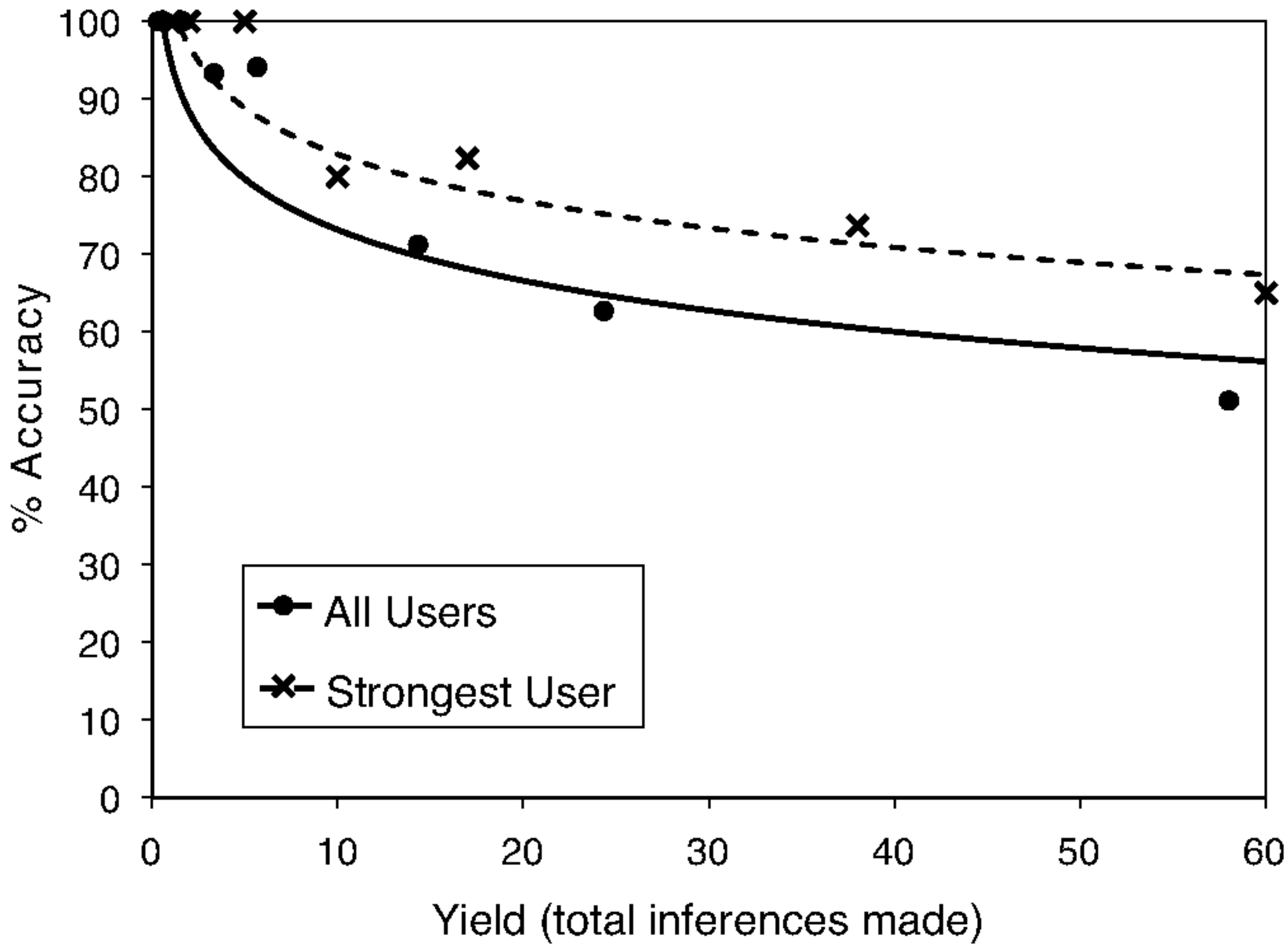


real users

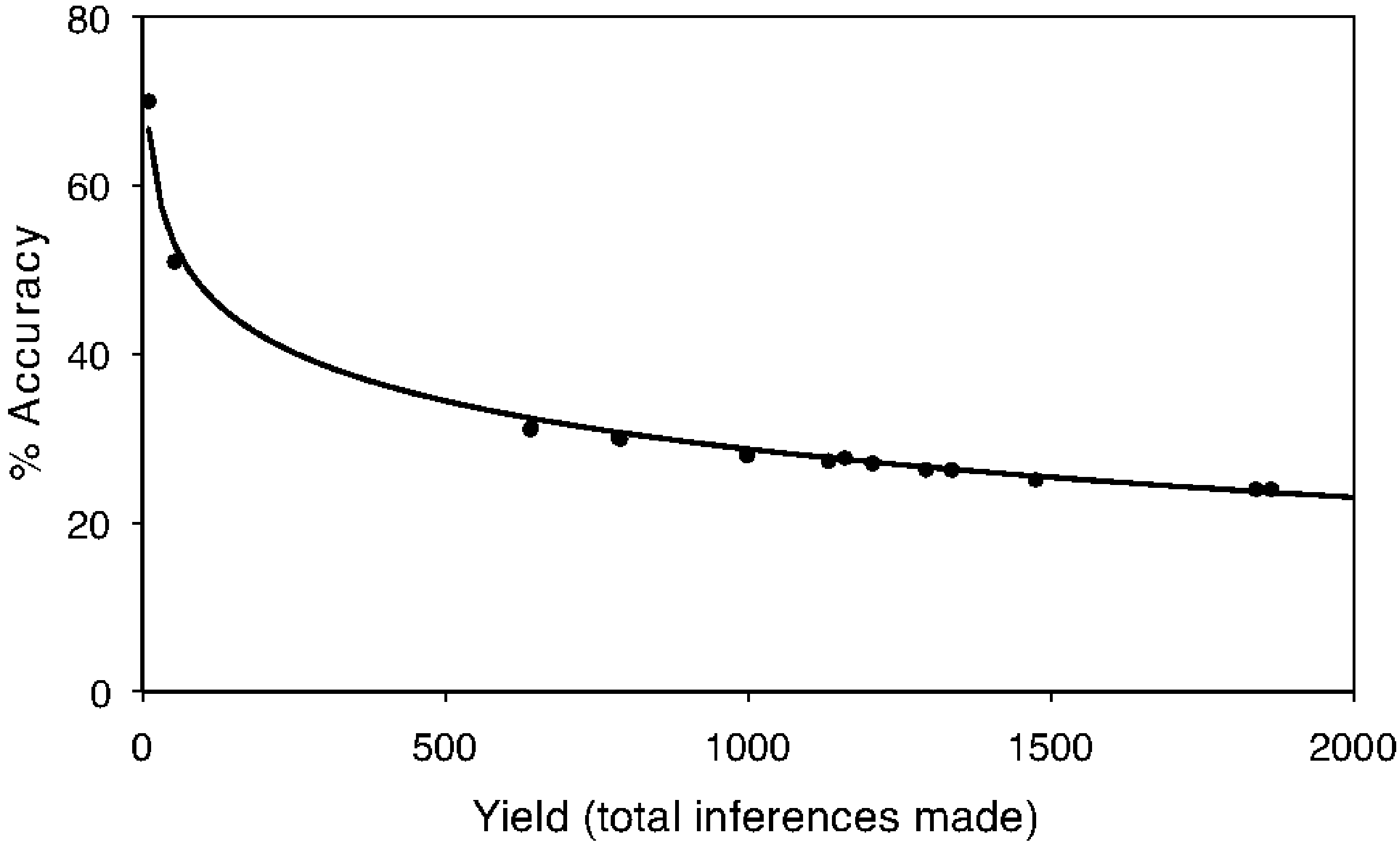


simulated users

results: LibraryThing



results: Last.fm



for an example user

suggested countermeasures

- limit the length of related-items list
  - the bottom items ordering reveal more information
- factor item popularity into update frequency
- avoid cross-genre recommendations
  - customers with interests in multiple genres tend to be at higher risk
- limit the speed and/or rate of data access
- user opt-out

conclusions

- public recommendations by recommender systems based on collaborative filtering may leak information about the behaviour of individual users to an attacker with limited auxiliary information
- customers of larger sites are generally safer
  - smaller datasets increase the likelihood of privacy risks
- undermine dichotomy between PII and large-scale aggregate statistics
  - dynamics of aggregate outputs constitute a new vector for privacy breaches

UBC

ece