On a Scale from 1 to 10, How Private are You? Scoring Facebook Privacy Settings
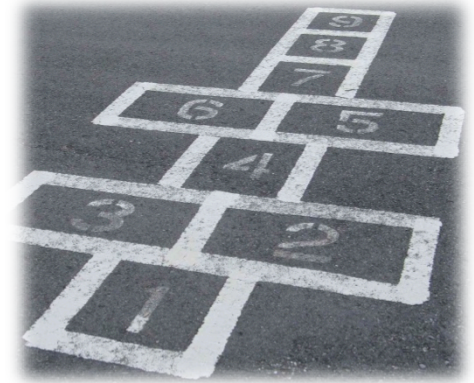
Tehila Minkus and Nasir Memon

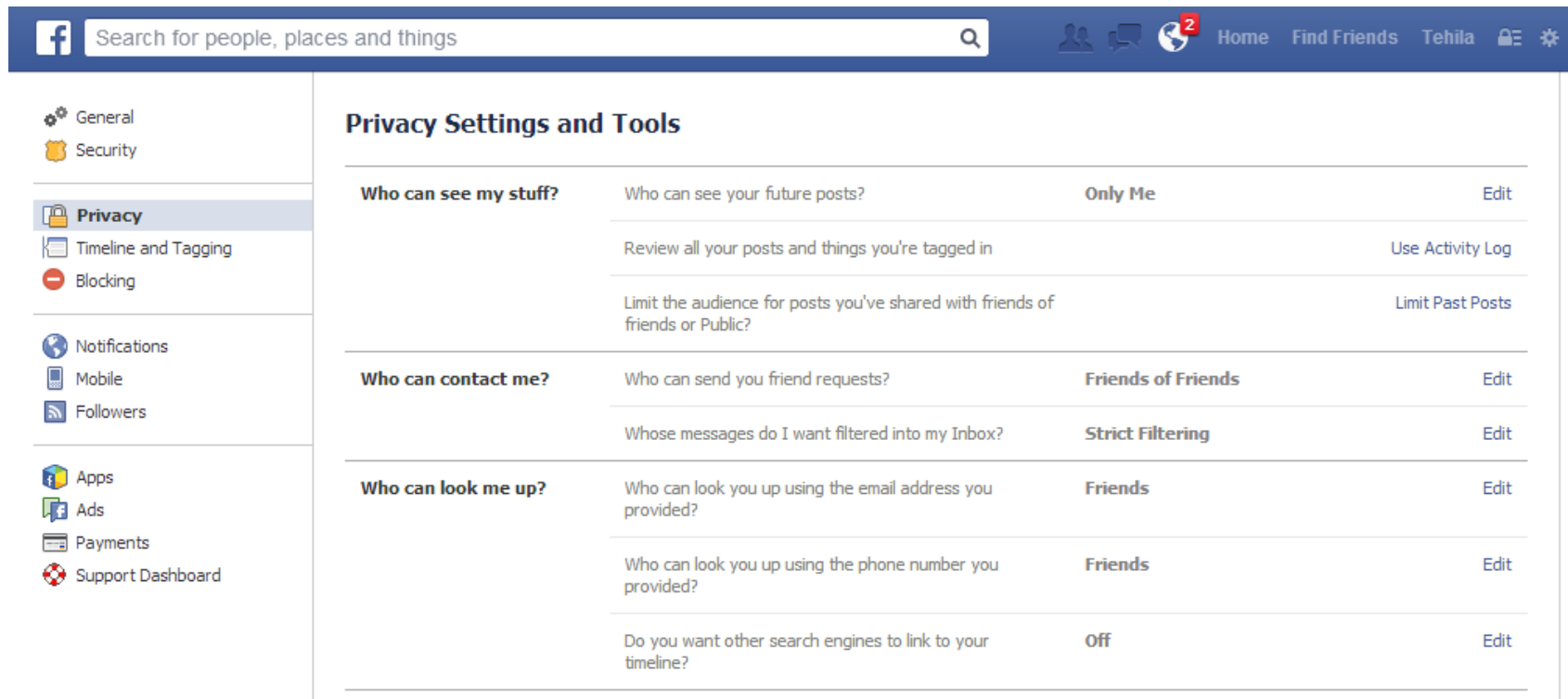New York University

USEC 2014

Facebook by the Numbers

- Number of users: 1.23 billion active monthly
- Time spent by each user: 8.3 hours monthly
- Average friends per user: ~300 among teens
- Number of privacy settings: 17
- Privacy of a given configuration: **unknown**

A Tale of Two Facebooks

Is it the best of privacy, or the worst of privacy?

Why quantify privacy?

- For users – instant feedback on choices

- For designers – prioritizing settings with high privacy risk

- For researchers – comparing configurations to learn more about peoples' privacy decisions

Outline

- Preliminaries:
  - List of settings
  - Terminology
- Two methods of scoring privacy settings
  - Naïve method
  - Weighted method
- User survey and results
- Potential applications
- Discussion and questions

NYU POLYTECHNIC SCHOOL OF ENGINEERING

Facebook Privacy Settings

Privacy	Timeline and Tagging	Apps	Ads
1) Who can see your future posts? 2) Who can send you friend requests? 3) Whose messages do you want filtered into your inbox? 4) Who can look you up using the email address or phone number you provided? 5) Do you allow other search engines to link to your timeline?	6) Who can add things on your timeline? 7) Review posts friends tag you in before they appear on your timeline? 8) Who can see posts you've been tagged in on your timeline? 9) Who can see what others post on your timeline? 10) Review tags people add to your own posts on Facebook? 11) When you're tagged in a post, who do you want to add to the audience if they aren't already in it? 12) Who sees tag suggestions when photos that look like you are uploaded?	13) What personal information goes into apps others use? 14) What is instant personalization set to? 15) Who can view your posts from old versions of Facebook for mobile?	16) Ads shown by third parties. 17) Ads and friends.

TABLE I

CURRENT FACEBOOK PRIVACY SETTINGS, DIVIDED IN FOUR CATEGORIES.

Terminology

- Privacy setting: a specific choice offered
  - e.g., "Who can see your future posts?"
    - Public
    - Friends
    - Custom
    - Only me
- Privacy configuration: the set of all privacy choices that a user has selected

Terminology

- Privacy score: the metric used to measure the privacy of a given setting's choice
  - e.g., "Who can see your future posts: Only me" earns what score?
- Total privacy score: the metric used to measure the privacy of a user's configuration
  - e.g., what overall score does Laura earn for the total set of privacy choices she made?

NYU | POLYTECHNIC SCHOOL OF ENGINEERING

Notation

- C := list of privacy settings

- $C(x)$:= privacy configuration of user x

- $C(x)_i$:= the option that user x chose for setting i

- $S(C(x)_i))$:= the privacy score assigned to the choice user x made for setting i

- $S(C(x))$:= the total privacy score earned by user x
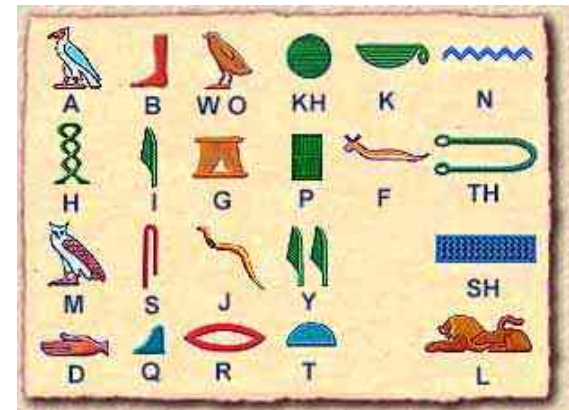
Naïve Method
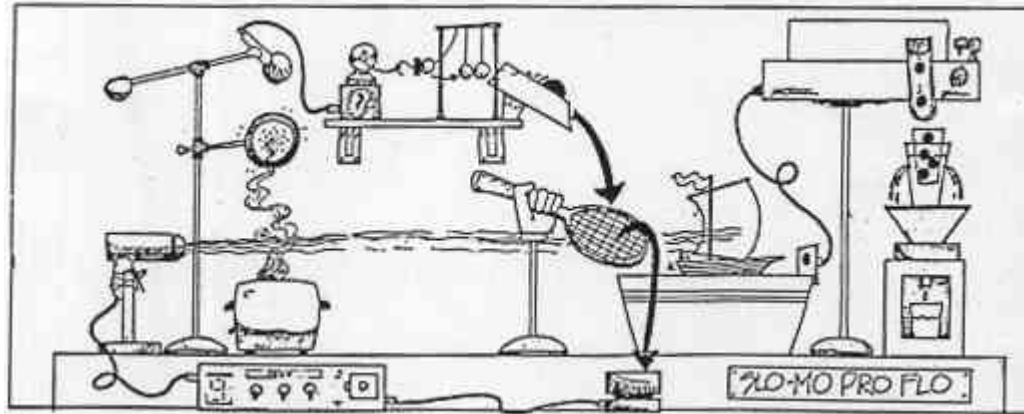
- General approach: assign scales of privacy for each option, then combine them into a holistic privacy score

Naïve Method – details

- Initialize a score for each available option:
  - For each setting, sort options in order of least to most private
  - Assign least private option a score of 0, and increment by 1 successively
  - Divide each score by the maximum score available
- Based on above, assign a score for each choice made.
- Total score: sum of the individual scores (scaled onto range from 0 to 10)

Naïve method, in notation:

$$S(x) = \sum_{i=1}^{n} \frac{s(C(x)_i)}{\max\limits_{k} s(C_{i,j})}$$

Pros and Cons of the Naïve Method

- Pros:
  - Easy to calculate
  - Robust to changes in the privacy settings

- Cons:
  - Assumes that all privacy settings are equally important

Weighted Method

- Can we more accurately capture the relative importance of each setting?

- What is privacy? [Maximilien et al, W2SP 2009]
  - *Sensitivity:* how embarrassing is the content?
  - *Visibility:* how public is the medium?

- Combining these two metrics yields a *privacy index* or a weight for a given privacy setting

Weighted Method: Detail

- For each setting:
  - Determine a weight (more detail later) expressing importance of setting in overall privacy of configuration, based on *sensitivity × visibility*
  - Assign a scaled score for the option chosen
  - Multiply these two to yield a weighted subscore
- Combine all the subscores to yield an overall score, and scale to get a numeric score in range 0:10

Weighted Method, in notation:

$$S(x) = \sum_{i=1}^{n} s(C(x)_i) * w(i)$$

Pros and Cons of Weighted Method

- Pros:
  - Accurately reflects priority in privacy settings
- Cons:
  - More sensitive to changes in Facebook privacy
  - Dependent on user base and time of sample

Determining Weights

- Survey on Amazon Mechanical Turk
- N = 250 before filtering, 189 after filtering
- Instrument design:
    1. Demographics
    2. Description and instructions
    3. Rating privacy settings

Sample questions on the survey

*9. Whose messages do you want filtered into your inbox? *(strict filtering, basic filtering)*

	Not at all	A little bit	Somewhat	Rather	Extremely
How sensitive is this?	○	○	○	○	○
How visible is this?	○	○	○	○	○

*10. Who can look you up using the email address or phone number you provided? *(everyone, friends of friends, friends)*

	Not at all	A little bit	Somewhat	Rather	Extremely
How sensitive is this?	○	○	○	○	○
How visible is this?	○	○	○	○	○

*11. Who is paying attention? Choose 'extremely' for the first part and 'rather' for the second. *(everyone, friends of friends, friends)*

	Not at all	A little bit	Somewhat	Rather	Extremely
How sensitive is this?	○	○	○	○	○
How visible is this?	○	○	○	○	○

NYU | POLYTECHNIC SCHOOL OF ENGINEERING

Weights as assigned by respondents

Privacy Setting	Sensitivity	Visibility	Sens.*Vis.
What personal information goes into apps others use?	2.82	2.16	6.0912
Who can see what others post on your timeline?	2.17	2.59	5.6203
Who can see posts you've been tagged in on your timeline?	2.39	2.35	5.6165
Who can look you up using the email address or phone number you provided?	2.42	2.17	5.2514
Who can add things on your timeline?	2.16	2.32	5.0112
Who can see your future posts?	1.97	2.41	4.7477
Review posts friends tag you in before they appear on your timeline?	2.29	1.96	4.4884
Who can view your posts from old versions of Facebook for mobile?	2.08	2.08	4.3264
When you're tagged in a post, who do you want to add to the audience if they aren't already in it?	1.84	1.94	3.5696

|			
What is instant personalization set to?	2.12	1.65	3.498
Who sees tag suggestions when photos that look like you are uploaded?	1.89	1.79	3.3831
Ads and friends. Pair my social actions with ads for whom?	1.89	1.78	3.3642
Review tags people add to your own posts on Facebook?	1.81	1.77	3.2037
Ads shown by third parties. Show my information to whom?	1.8	1.76	3.168
Do you allow other search engines to link to your timeline?	1.86	1.63	3.0318
Who can send you friend requests?	1.09	2.04	2.2236
Whose messages do you want filtered into your inbox?	1.47	1.47	2.1609

TABLE II
WEIGHTS OF CURRENT FACEBOOK PRIVACY SETTINGS, SORTED BY THE PRODUCT OF SENSITIVITY AND VISIBILITY IN AN INCREASING ORDER.

NYU | POLYTECHNIC SCHOOL OF ENGINEERING

Observations

- Wide range in assigned values confirms that not all privacy settings are equally important

- These priorities can be used in design of FB interface
  - Overall scoring of privacy
  - Privacy shortcuts
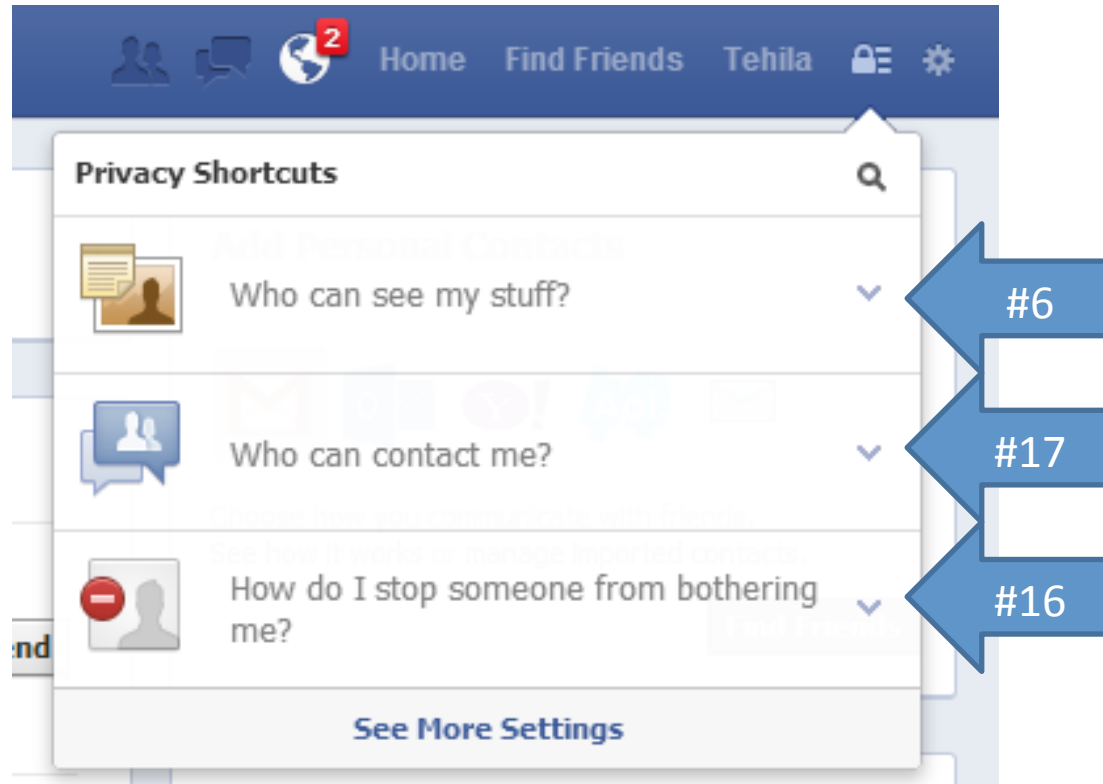
Privacy Shortcuts

Privacy Shortcuts

- Recommended settings, as rated by users:
  - "What personal information goes into apps others use?"
  - "Who can see what others post on your timeline?"
  - "Who can see posts you've been tagged in on your timeline?"

Recap of Contributions

- Are all privacy settings equal? NO
- Two methods for calculating privacy:
  - Naïve scoring method
  - Weighted scoring method
- An ordered list of privacy settings
- Applications for privacy scores:
  - Users
  - Designers
  - Researchers

Thank you!

- Any questions?

email: tehila@nyu.edu