Social Media, Usability, and Privacy

Thorsten Strufe, Thomas Paul, Daniel Puscher
Chair for Privacy and IT-Security

Padova, 06.09.2016
A Short History of Networked Services...

FREE ACCESS to...
- Tradewars 2002
- Lord2
- FIP...
- And Much More...

Legend of the Red Dragon ...
Lunatix...
Telnet...

Running On: Wildcat v5/winserver
Connected To node:2
Your Name: Aquaman
Today's Date is: 3/26/2002
Telnet: bbs.froglan.net
URL: http://www.froglan.net
Note: Check Out http://www.hn.com slashdot for BBS's!!!!!!!
[Only One Account Per Person]

What is your user name?
...around came the World Wide Web...

Major announcement

My book is now on sale! The page detailing about my book has been updated as well, now including the full size version of the front cover image.

Welcome

Welcome to Utilitas City! In Utilitas City, you'll find tips and tricks that'll help you with math, computer operation, and a few others. I have three games you can play: rattles, the doce, or mathties. My version of the last's journal with over 100 entries is probably one of the biggest online journals. From school and my game, I have some stories available. Along with that, I have two computer games I'm working on: "The Supernatural Olympics" and a 2D RPG game. Both websites, the only ones, my great creations from 2002, are in one beautiful area. My site is best viewed at 1024x768 resolution at true color (24 or 32-bit color). 1024x768 is optimal.

Lou's page

1995 Annual Report

- Letter to Investors/Financials
- The Future of Computing
  - 1994 Annual Report Review

Financial information

- IBM stock quotation
- Quarterly results
- Stockholder services

Employment

IBM Planetwide

Other corporate activities

- IBM and the environment
- IBM and the Global Information Infrastructure
- Philanthropy


06.09.2016
...with more info than we really wanted...

Unsere Feste in Bildern ...

Da es sich hierbei um ganz private Feste handelt, sind diese Seiten nur für Familienmitglieder und Freunde zugänglich.

An alle Verwandten und Freunde!
Bitte schickt mir eine Mail, falls ihr eine Zugangsberechtigung wollt.
Ich werde euch dann umgehend die Daten zumailen.

JA!!!
Ich will ein Passwort!!!

Momentan könnt ihr Bilder folgender Feste betrachten:

- Sebastian's Konfirmation
- Joachim's 50½ Geburtstag
- Otto's 65. Geburtstag

Auf zu den Bildern!
...with high barriers.

Quite costly
Difficult to set-up and maintain

Initially mainly
- a few “nerds”
- universities (research centers)
- large companies and only

http://www.cern.ch
http://www.berkeley.edu
http://www.bbn.com
http://www.fiat.it
http://www.cocacola.com
1969: Birth
1972: First Email
1978: First Spam
1979: First MUD (online game)
1980: First Virus
1983: Introduction of TCP/IP
1991: WWW and first Web pages

Growth of the Internet

http://www.zakon.org/robert/internet/timeline/
http://www.isc.org/solutions/survey
Reach

Mass Media/
Information at your fingertips!

Communication:
Skype/VoIP!
chat, text, discussions, mail!

User generated content,
live- and on demand
multimedia streaming

Shopping!
(around the world)

Online games...

http://www.internetworldstats.com

06.09.2016

SNSPT '16 – Thorsten Strufe
But What is it, That Moves all Those People on the Web?

Facebook (Home and Work)

<table>
<thead>
<tr>
<th>Person(s)</th>
<th>MOM UA % Change</th>
<th>MOM Time % Change</th>
</tr>
</thead>
</table>

Facebook

Rest of the Web

06.09.2016
So Today *everybody* Shares Some Data…

Linkedin

My Travel

Your upcoming trips
December 2008
New York City, NY — Dec 13 - Dec 23
Flying to NYC for work, then down south to visit friends and family on the east coast.

Who’s close to you
Scott Hintz
Tripit
Nov 18 - Nov 20, 08

Andy Denmark
Tripit

See where your entire professional network is traveling and when you will be in the same city as your colleagues. Meet up at the next industry event or re-connect with old friends. Add the My Travel application to display your current location, upcoming trips and travel stats within your network.

Raves about Tripit:
• “Tripit’s straightforward approach makes it addictive…” – Walt Mossberg Solution, Wall Street Journal
• “Best for organizing travel details…” – CNN, Travel + Leisure
• “Easy and damn useful…” – DailyCandy
• “A terrific site that has changed my travel life…” – WNYC New York TV

MNL
...with calculated side effects...

A Quebec woman is suing her employer for firing her after she accepted a Facebook friend request from her boss. It is alleged that after accepting the request, she was called into work and asked to install a program on her computer, which led to her computer being stolen.

The comic strip illustrates the potential consequences of accepting such requests and installing programs on personal devices.

Sources:
...increasingly immersive to daily life...
Penultimate public census
Scheduled for 1981 (delayed to 1983)

Significant public opposition
- Fear of a surveillance society
- The transparent citizen ("gläserner Mensch")
- Bounty for discovered GDR citizens and esp. foreigners
- Appeal for civil disobedience
- Finally accomplished in 1987

Consequence: "25% inherent error"
- Significant gap between census and community register
  (So let's just get a unifying tax number... ;)

http://upload.wikimedia.org/wikipedia/commons/4/43/Volkszaehlung.JPG
06.09.2016
... completely willingly...
Online Social Networks

Simplified, walled-garden version of „the Web“:
• Easy to set-up pages („profiles“) of individuals (... and companies...)
• Links reflecting real-world relations between individuals
• Possibility to share user generated content

...including messaging
• “Guest book” / “Wall” (asynchronous broadcast)
• Email (asynchronous unicast)
• Chat (~ synchronous unicast)

Collaborative applications / games

! Different target audience / application domain
• Private and personal OSN
• Public and professional OSN (business-oriented)
Target Audience and Domain

Professional business services

Private and personal services

Other services
• “Micro blogging”: Twitter
• Business trips and meeting service: Dopplr, TripIT
• Location-based achievement systems: foursquare, gowalla

Target Audience and Domain

06.09.2016
Two words on: *Information Sovereignty*

- The Bible
- Newspapers
- Mass media
- The Internet (1.0)
- Web 2.0...

i.o.w: to which effect? aka. „Power to the people!“ (O RLY?)

So what do you do?
Go into the ammo business...

The deer now have guns...

06.09.2016
The Ammo Business...

Facebook is and will always be free. “At Facebook your data is yours.”

“Myspace, after failing to meet the expected income, starts selling their users‘ data.”
it’s known **who they are**
with **whom** they’re **friends**
whom they’re **talking to**
what they **think** and **want**

and specifically what they’re **aiming** at...

...and quite fortunately, **their ammo can be removed**,  
*should they fail to comply*
Can it Get Worse?  
*(a little polemic)*

**Homogeneity and control!**
Provide their Hardware  
Operating System  
„Applications“ (controlled, be the gatekeeper!)

Just license them, don‘t give them any ownership!  
Control/surveillance their music  
Their movies  
Their social network...  
...their life...
„And it’s a Unix under the hood!“

„Yeah, they’re so cool, no effort, no hassle!“

„Worst case: you can jailbreak...“

Think differently.

Enough polemics, back to the topic!
Let’s try and understand the behavior of users in some OSN

**ABSTRACT**

Analyzing Online Social Networks (OSN), voluntarily maintained and automatically exploitable databases of electronic personal information, promises a wealth of insights into users’ behavior, interest, and utilization of these currently predominant services on the Internet. To understand popularity in OSN, we analyzed a large sample of profiles from a highly popular network for three months, and analyzed the relation between profile properties and their impression frequency. Evaluating the data indicates a strong relation between both the number of accepted contacts and the difference of updating contacts versus the frequency of requests for a profile. Counter intuitively, the overall activity, gender, as well as participation gym of users have no remarkable impact on their profile’s popularity.

**Categories and Subject Descriptors**

1.4 [Computer Applications]: SOCIAL AND BEHAVIORAL SCIENCES

**Keywords**

Online Social Networks, Popularity, User Measurement, User Model, Behavior

1. INTRODUCTION

Social networking services (SNS) contain a wealth of information. Users voluntarily feed self-discovery details into the Online Social Network (OSN) and their utilization behavior is completely observable by the OSN provider. Analyzing this complex data facilitates understanding of the psychological and sociological properties of online social networks and their users. Understanding how users navigate online social networks, e.g., offers insights into how people become the presented profiles. It consequently allows for the identification of relations between selected profile characteristics and their request frequency. We will call this frequency popularity (P) for the remainder of this paper, which we will measure as the number of impressions of a profile P (the number of times it is viewed by another user) per time t (i.e., P(t)).

Identifying key properties of users and their profiles, which allows for the prediction of their popularity, however, is not only interesting for social scientists. Especially system designers and developers of social networking services may capitalize on the extracted knowledge. In case of centralized, service based systems, like Facebook, LinkedIn, or Twitter, the user experience may be enhanced. Being able to predict the popularity helps engineering tuning parameters for different profiles. Designers hence are able to decide, which profiles need to be promoted with very low delay under normal circumstances. The possibility to identify properties of profiles that may tolerate higher response times is an additional, beneficial effect. It additionally yields information, which profiles may more frequently tolerate temporary unavailability without causing a significant deterioration of the service experience.

Knowledge on the popularity of profiles may prove being even more valuable for the development of decentralized architectures [1, 2], for the parameterization of this entirely decentral approach to providing social networking services. Rough, intuitive beliefs about the profile popularity are abundant. Unverified rumors frequently have it that profiles of women are more often visited than profiles of men, that providing a profile picture will drastically increase the number of impressions of a profile, or that the unfortunate possession of a last name later in the alphabet, will inevitably condemn a user’s profile to eternal lack of popularity.

This work aims at exploring these conceptions. It analyzes data from a large, central European OSN with a main focus on professional and business centered utilization. The main contributions are:

- to correlate the fact that correlations between properties of profiles and their popularity do exist, and
- to expose, which properties of profiles have a considerable effect on their popularity, and which properties have less, or no effect at all.

A large set of profiles from the selected OSN are monitored for this purpose, and traces that were gathered over a period of three months in this course have been analyzed. Mann Whitney U tests have been used to identify significant differences in the popularity of profiles of different groups, and correlations between parametric variables and the profile’s

http://www.facebook.com/twitter.com
Understanding the behavior of users in OSN

Why?

• It’s interesting! 😊
• Plus: we need to know to build better (P2P) OSN...

Questions of interest

• Sessions (when, how long, - active, - often?)
• Preferences / services used
• Popularity of content / pages
• Scope of access / reciprocity?

Here: focus on profile popularity...
What’s in a “Professional” Profile

- Identifying info
  - Name
  - Photo
  - Address...
- CV
  - Current/prev. employments
  - Educational track
- Interests
  - Personal/professional
  - Wants/haves
  - Interest Groups
- Personal contacts
- Messaging
- Statistics
Popularity of Profiles – Intuitive Beliefs

Which profiles are “popular”?
- Measured in frequency of requests
- Possible to correlate to properties of user/profile?
- Which profiles do we have to keep available (and by which means?);-

Why?
Common beliefs...
- “Profile of women are much more often visited than profile of men”
- “Profiles with pictures are more interesting than profiles without”
- “Old/experienced profiles attract more views”
- “The profiles of active users are more attractive”
- “Users with many friends are sought and viewed more often”
- “Last name starting with a letter late in the alphabet sucks…” (c/list pages...)

Reflecting: how are users lead?
- Assuming the users generally follow links (rather than searching for content)
- What do they see as “home” – the front page of the OSN?
The Front Page (and how to get on it)

“Home page” (after login) usually almost identical
- Info on profile owner
- Updates from the provider (and advertisement)
- Feed of news from “f

Activity in the news feed:
- Changes to profile
- Status updates
- Birthdays
- Contact list maintenance (adding friends)

Note: having many friends leads to broad dissemination...
Data Collection

How can we gather the data?

- Access to server logs (Ha!)
- Surveys & Interviews (problems of scale)
- Traffic logging (problems of scope)
- Crawling/API access (problems of scale, incompleteness of information, sampling)

Crawls gather only limited data

- Does not sufficiently allow inference on sessions
- Mainly comprises of plain, static profile info and social graph
- Generally does not include data about popularity

Regular monitoring:

- Collect changes to profiles
- Frequent, regular measurements over long period of time needed
“xing” selected for the study
  • Business/professional OSN, similar to LinkedIn
  • 8 Mio users, mainly from central Europe (now: 13.8; 6.7 from DACH)
  • xing profiles include
    – Registration date
    – activity meter
    – hit counter (number of profile impressions for popularity)
    – Weak privacy settings (professional profiles are there to be seen)
  • Visitors to profile visible (to paying users – no stalking, unlike LinkedIn, facebook)

Crawling / monitoring the complete OSN is infeasible
  • 8mio profiles, most >15 pages of contact lists (up to 160k contacts!, 10 per page)
  • Access per page takes ~ .5s, complete crawl takes > 275 h (if all goes well)
  • each page > 150 KB, > 17 TB in total
  • Providers don’t like this much... (rate control, disabled accounts, blocked IPs)

Large, random sample needed for meaningful results
Random sampling
- Conducted random walks (25k, 5k, ~1k)
- Aim of the crawls:
  - Diverse graphs without overlap
  - Collect “john does” (no outliers, no abandoned profiles)
- Covered over 2Mio unique profiles in total
- Starting at diverse “edges” (AUS, DE, PL, RUS, TR, UK, US)
- all converged to D.A.CH

Selected sub graph without overlap
- 31.643 unique profiles (25k, 5k, 1.6k random walks)
- Gender automatically derived via website on international first names

Subsequently frequently monitored for a long period of time
- Since Nov 2009
- At least twice daily
- Only core data needed (no pictures, friend list not regularly since # on profile)

06.09.2016
Identifying Data
  - Name, Image, Gender

CV
  - Current employment, universities attended, claimed spoken languages

Interests
  - Interests as stated
  - Number of subscribed groups, subscribed groups, number of members in groups, number of messages in respective groups, languages of group

Contact list information
  - Number of contacts
  - Complete list of contacts gathered infrequently

Statistics
  - Registration date, number of profile impressions, activity meter
  - Timestamp of crawl

06.09.2016
Analyzing & Validating the Sample

3 months section of monitoring data analyzed (Nov ’09 – Jan ’10)
Some profiles removed (celebrities, abandoned profiles)
Remaining sample
- 25,274 (7,824 / 17,450) (31% vs. 34%)
- Degree dist. (~PL, min 5, max 12.332)
- Name frequency follows Zipf
  - First names: s = 1.67
  - Last names: s = 3.14
- Binned popularity dist. log-normal

06.09.2016
Activity

- “Activity meter” in profiles very coarse grained
- Derived “profile alteration frequency” as alternative
- Men are slightly more active than women (to both metrics)
- Profiles without image belong to inactive users

Membership in Groups

- Wide range of group membership
- Max 511, Mdn 3
- >5k profiles are not registered to any group
Testing Groups of users (non-parametric)

Testing popularity between groups (nominal data, Mann-Whitney U)

“Profile of women are much more often visited than profile of men”
- Male vs. female: 0.039 vs. 0.041 (Mdn)
- No significant difference

“Profiles with pictures are more interesting than profiles without”
- With picture significantly higher popularity (0.5 pi / d)
Correlational Tests (Pearson’s *r*)

“*The profiles of active users are more attractive*”
- Activity (parametric: interval) as given on the profiles (“*activity meter*”)
  - $r \approx 0.17$, no noteworthy correlation
- Activity (parametric: ratio) measured in *group memberships*
  - $r \approx 0.37$ (higher for men, lower for women)
- Activity meter is very coarse grained
- Activity measured in *profile alterations*
  - $r \approx 0.62$ ($0.61 < r < 0.63$) *high correlation*
→ Popularity correlates with activity of users (profile alterations/group activity)

“*Users with many friends are sought and viewed more often*”
- Correlating popularity to the degree of profiles
  - $r \approx 0.75$, *high correlation*
  - Stronger for women: $0.81 < r < 0.83$ vs. men: $0.74 < r \leq 0.75$

**In retrospect:** Combination somewhat unsurprising: changes are published at friend’s profiles...
“Old/experienced profiles are viewed more frequently”

- Preferential attachment / experience could lead to higher popularity
- Ho rejected, but \( r \approx 0.11 \), no noteworthy correlation

“Last name starting with a letter late in the alphabet sucks...”

- Ho not rejected, there is no correlation.
- *Taking the “rich-club”, however...*
- Top 5% profiles: \( r \approx -0.09 \) the better..)
- Top 2‰ profiles: \( r \approx -0.22 \)
- Top 1 ‰ profiles: \( r \approx -0.29 \)
- Top 10 profiles: \( r \approx -0.9!! \)
Summarizing the User Model

Selected large sample of profiles in business oriented OSN
Monitored profile properties and popularity (in pi/h)

Profile Popularity can be predicted. Relates to
- Providing image
- Activity (diligence of maintaining profile)
- Number of friends and contacts

What we take away (P2P OSN)
- Nice correlation with activity/friends (P2P & replicating at friends...)

Future Work
- Kept monitoring, but results quite stable...
- Analyze data from DB and server access logs “spi” (fb-like personal osn)
  - Are profile requests “local” (viewing friends...)?
  - Is interest mutual?
  - Can we learn more on the sessions?
- Struggle to get/analyze more data!

Shameless plug: FPA
So what about these privacy settings?

C4PS - Helping Facebookers Manage their Privacy Settings

Thomas Paul, Martin Stopczynski, Daniel Puscher, Melanie Vollmer, Thorsten Strufe

CfP, Technische Universität Dresden

Abstract

The ever increasing popularity of Online Social Networking has left a wealth of personal data on the web, accessible for broad and automatic retrieval. Protection from untrusted recipients and harvesting by crawlers is implemented by access control, usually configured by the user in his privacy settings. Privacy unfriendly default settings and the user unfriendly privacy setting interface cause an unnoticed sharing. We propose C4PS - Controls for Privacy Settings, a concept for future privacy setting interfaces. We developed a mockup for privacy settings in Facebook as a proof of concept, applying color coding for different privacy visibilities, providing easy access to the privacy settings, and generally following common, well known practices. We evaluated this mockup in a lab study and show in the results that the new approach increases the usability significantly. Based on the results we provide a Firefox plug-in implementing C4PS for the new Facebook interface.

1 Introduction

Over 500 million users allegedly share personal information, private photos, videos, opinions and discussions on Facebook. The shared personal information include their age, gender, sexual preferences, taste and hobbies. All this data stored in Facebook or any other Online Social Network (OSN) can be linked to the relating individual by their real names published in their profile.

Access to all this information is controlled by the OSN service provider, based on the user’s privacy settings. Studies have shown that despite increasing awareness, users due to the intricacy of the task are incapable of configuring their intended settings, and indeed do not understand their activities’ implications. Moreover, the fact that Facebook and other OSNs have modified the default privacy settings to be more and more open with each update, makes it very important that users can easily grasp and change their privacy settings.

Consequences of this situation span unintended over-sharing, and more serious threats, arising as spoofing and harvesting [3,7], automated social engineering [10], social phishing [11] as well as various further attacks. In face of this pervasive incomprehensibility, we go as far as proposing to abandon
Increasing awareness due to greater press coverage
Young people share more, but manage privacy better!

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>single change at least</td>
<td>30%</td>
<td>91%</td>
<td>98%</td>
</tr>
<tr>
<td>3 or more changes</td>
<td>/</td>
<td>24%</td>
<td>51%</td>
</tr>
</tbody>
</table>

Access Model

User

Grantable
- specific contact(s)
- contacts
- contacts of contacts
- service subscribers
- public

Implicit
- SNP
- Affiliates
- Extenders
- Advertisers

- ISP

Everything the installing user can see
Not much (aggregates) Unless they pay really well

Everything their subscribers see/write (until Nov 21st '12)
Facebook’s Privacy Evolution (2005)

Facebook is now the most successful social networking service provider today:
- >600 mio users
- Integration of new services:
  - Status updates and news feed (similar to Twitter)
  - Facebook apps
  - Places (similar to Foursquare)

Successful as a voyeur service (you can’t know who’s watching you).

…living of the availability of personal data

Source: Matt McKeon
Facebook’s Evolution: 2007

2007

Availability of your personal data on Facebook (default settings)
Number of People

Matt McKern, May 2016
Facebook’s Evolution: 2009

2009 (Nov)

Click the chart to advance, or click on a year

- 2005
- 2006
- 2007
- 2009 (Nov)
- 2009 (Dec)
- 2010 (Apr)

Availability of your personal data on Facebook (default settings)

Number of People

Matt Mckeon, May 2016
Help users manage their privacy settings
  ● Apply well known control methods
  ● Proximity between controls and data / decrease overhead

Help them understand what they've done!
  ● Display settings directly
  ● Color coding helps humans understand a situation better

Quick walk through the concepts...
Low Overhead and Easy Access

Settings reached easily via centrally-mounted buttons

Settings integrated directly in profile, can be changed with a single click
Background colors indicate visibility of the entries

Help and explanations
Well-known Control Methods

Drag n’ Drop
Inactive elements greyed out

Help! (tooltips)

Für alle Freunde sichtbar machen
Aggregating Access-Management using Groups

Sichtbarkeit für "Beziehungsstatus" einstellen

Gruppen:
- Schulfreunde
- Arbeitskollegen

Alle Freunde:
- Andrea Schweizer
- Anja Mayer
- Christian König
- Christina Kuhn
- Claudia Bauer

Wer darf "Beziehungsstatus" sehen?
- Schulfreunde
  - Marco Hartmann

Schließen
Evaluating an Interface

Properties to evaluate, hypotheses to test
1. Is it easier to find out who can see what? (setting)
2. Can the user easily find out how an arbitrary other sees her? (visibility)
3. Do groups make life easier, faster, more precise? (groups)
4. Is the entire new interface more effective than facebook? (effectivity)

Experiment
- Controlled in-depth user study with 20 (18) participants
  - General questions
  - Extensive set of tasks to solve (alternating order of systems)
  - Standardized satisfaction scores
- Metrics: precision, overhead (clicks), time needed

Expectation: advantage of our interface, especially for laypeople
67% visit social networks daily

67% were Facebook users

94% have already changed their privacy settings

76% found the settings to be confusing

22% create groups of friends

22% do not know exactly what parts of data they have shared
Precision

- Grasp settings
- Checking visibility
- Group settings
- Complex tasks (effectivity)

Hypothesis

- H1
- H2
- H3
- H4

Precision:

- Facebook
- New Interface
Testing the Audiences

Facebook users

Users with no fb experience
Effectivity / Number of Clicks

![Box plot showing the number of clicks for Facebook and New Interface across different tasks.](image)
Effectivity / Time Needed

![Box plot showing time needed for different tasks using Facebook and new interface. The x-axis represents different tasks, and the y-axis represents time in seconds. The box plots compare the efficiency and time needed for the two interfaces.]
94% found Facebook settings confusing
89% rated new solution as "much better"
100% rated coloring “good” to “very good”
82% think chosen colors are useful / clear.
Some issues with the choice of colors
Profile preview not found
Notion of “Selected Friends” somehow difficult to grasp
Follow-up of the study...
Facebook Privacy Watcher

Study Setting

- Browser extension (plug-in) for Firefox and Chrome
- More than 44,800 downloads from 102 countries
- Based on data, embedded in feedback requests (*informed consent*)!

Collected data:
- User Demographics
- Completeness of profile fields
- Friend lists (hashed)
- Visibility of profile fields
- Privacy changes that have been made with the FPW
- Number of plug-in activations

---

<table>
<thead>
<tr>
<th>Country</th>
<th># Feedback responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Germany</td>
<td>7,581</td>
</tr>
<tr>
<td>Egypt</td>
<td>272</td>
</tr>
<tr>
<td>Austria</td>
<td>218</td>
</tr>
<tr>
<td>United States</td>
<td>150</td>
</tr>
<tr>
<td>Switzerland</td>
<td>147</td>
</tr>
<tr>
<td>France</td>
<td>94</td>
</tr>
<tr>
<td>Spain</td>
<td>72</td>
</tr>
<tr>
<td>Netherlands</td>
<td>62</td>
</tr>
</tbody>
</table>
Facebook Privacy Watcher
Alteration Matrix

From
- Public
- Friends
- Only me
- Custom

To
- Public
- Friends
- Only me
- Custom

- 26.5%
- 23.2%
- 2.8%
- 9.2%

- 4.4%
- 8.8%
- 4.8%
- 9.9%

- 1.5%
- 4.6%
- 1.6%
- 2.7%
### Facebook Privacy Watcher

Re-Authorizing Specific Attributes

<table>
<thead>
<tr>
<th>Attribute</th>
<th>More Private</th>
<th>Less Private</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Timeline-Entry</td>
<td>0.42</td>
<td>0.58</td>
<td>7133</td>
</tr>
<tr>
<td>Friendlist</td>
<td>0.55</td>
<td>0.45</td>
<td>835</td>
</tr>
<tr>
<td>Employer</td>
<td>0.45</td>
<td>0.55</td>
<td>433</td>
</tr>
<tr>
<td>Current city</td>
<td>0.48</td>
<td>0.52</td>
<td>422</td>
</tr>
<tr>
<td>Relationship Status</td>
<td>0.58</td>
<td>0.42</td>
<td>401</td>
</tr>
<tr>
<td>Family</td>
<td>0.50</td>
<td>0.50</td>
<td>358</td>
</tr>
<tr>
<td>Hometown</td>
<td>0.56</td>
<td>0.44</td>
<td>356</td>
</tr>
<tr>
<td>Photo album</td>
<td>0.50</td>
<td>0.50</td>
<td>340</td>
</tr>
<tr>
<td>School</td>
<td>0.53</td>
<td>0.47</td>
<td>337</td>
</tr>
<tr>
<td>Birthday</td>
<td>0.47</td>
<td>0.53</td>
<td>273</td>
</tr>
<tr>
<td>University</td>
<td>0.45</td>
<td>0.55</td>
<td>205</td>
</tr>
<tr>
<td>About you</td>
<td>0.55</td>
<td>0.45</td>
<td>151</td>
</tr>
<tr>
<td>Quotations</td>
<td>0.50</td>
<td>0.50</td>
<td>116</td>
</tr>
<tr>
<td>Religious Views</td>
<td>0.50</td>
<td>0.50</td>
<td>75</td>
</tr>
<tr>
<td>Instant Messenger</td>
<td>0.50</td>
<td>0.50</td>
<td>73</td>
</tr>
<tr>
<td>Interested in Languages</td>
<td>0.50</td>
<td>0.50</td>
<td>70</td>
</tr>
<tr>
<td>Mobile phones</td>
<td>0.50</td>
<td>0.50</td>
<td>58</td>
</tr>
<tr>
<td>Political views</td>
<td>0.50</td>
<td>0.50</td>
<td>58</td>
</tr>
<tr>
<td>Website</td>
<td>0.50</td>
<td>0.50</td>
<td>55</td>
</tr>
<tr>
<td>Address</td>
<td>0.50</td>
<td>0.50</td>
<td>35</td>
</tr>
<tr>
<td>Emails</td>
<td>0.50</td>
<td>0.50</td>
<td>29</td>
</tr>
<tr>
<td>Other phones</td>
<td>0.50</td>
<td>0.50</td>
<td>6</td>
</tr>
</tbody>
</table>

**Changing the setting:**
- **More private**
- **Less private**

06.09.2016
Sampling was not unbiased. But:
- Not everybody wants *more* privacy
- Users distinguish between types of attributes

Facebook is an international system. But:
- Extreme cultural differences in authorizations

By far not satisfying data
- We need more, to understand better.
Meet FPA: Facebook Privacy Analyzer
Browser extension for Firefox and Chrome
Endowed with user's access rights

Observation on users' own devices and own user profiles

Observation period:
• 123 days; started on 1st of January 2014
• Flexible start and end → average observation period: 34 days

Participants: 2071 (with informed consent!)

Collected data:
• User Demographics
• Performed actions
• Friend lists (hashed)
• Activity logs
Temporal properties:
• When are users active
• What are the churn models

Interaction with content:
• What are the properties of content users interact with? (Where do they stay?)
FPA: Measuring Churn, Method

Precise measurement reflects user attention better
FPA: Measuring Churn, Results

2:16 minutes average session duration

Facebook sessions are much shorter than assumed in the literature!
25.77% of all accessed content is younger than 1 h
84.79% younger than 24 h

Accessed content is very fresh
FPA: Measuring Retention Time

![Box plot showing relative retention times for different categories.](image)
Users communicate only with a minor subset of their friends

Percentage of Facebook friends that was communicated with:
Stale user models become obsolete

FPA: Behaviour over Time

Share of the action in the activity log

Year


n = 960 n = 1574 n = 2668 n = 2597 n = 1428 n = 496

Comment  New friend  Share  Post to timeline
Like      Photo    Status-Update

06.09.2016
Dozens of additional analyses...

General findings:
• Facebook is a mature OSN
• Users find less new friends
• Fraction of low effort actions (likes, shares) rises; the fraction of high effort actions decreases (e.g. comments, status updates)

And...
• By far not satisfying data
  • We need more, to understand better.
References

Leyla Bilge, Thorsten Strufe, Davide Balzarotti, and Engin Kirda. “All Your Contacts Are Belong to Us: Automated Identity Theft Attacks on Social Networks.” In WWW, 2009


All pictures credit wikimedia, unless stated differently

06.09.2016