



Peer-to-Peer Networks

Chapter 5: Applications on P2P, Online Social Networks
Thorsten Strufe

Chapter Outline



- Overview
- Communications (Online Social Networks)
- Content Distribution Revisited (P2P Streaming)
- Distributed Backup and Storage
- Gaming

P2P: More Than Filesharing?



- P2P so far focussed on finding stuff
- Main application: File Sharing (exchange of data objects between end hosts)
- It works well and is very popular, since
 - Resources (bandwidth/selection of data) in client-server is limited, scales better with P2P
 - Users have an incentive to participate (download sth fast, it's free as in beer)
 - Almost pervasive broadband access, „always on“
→ Difference between clients and servers decreases (again...)
- So what about other applications of the P2P paradigm?

More than File-Sharing!



- Reversing the paradigm:
 - Distributed backup and storage
- Pushing content distribution a bit further:
 - Overlay- / P2P streaming, Application Layer Multicasting
- Distribution of control
 - Groupware / Collaborative work
 - Decentralized online social networks
- Centralized use of P2P
 - Enhancing centralized systems using P2P

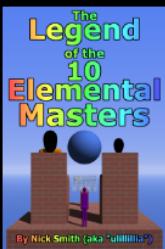


- What are Online Social Networks and how have they evolved?
 - A short overview of the history of networked services
- Why are they interesting for research?
 - A peek at their prevalence
- How do people actually use them?
 - Characterizing user behavior in OSN
- The privacy problem
- Movements to decentralize
- Safebook

A Short History of Networked Services...



...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.

Announcement

It's now official - my book is going into book stores. When you'll see it, that I don't know. Lulu states 6 to 8 weeks so this implies a time from April 13 to April circumstances - I couldn't order my second proof copy. **Updated Mar 3, 2010**

Uillillies City Software logo - a space city drifts freely among the background stars with colorful 3D text in front

1 Welcome!

Welcome to Uillillies City! In Uillillies City, you'll find tips and tricks that'll help with math, computer operation, and a few others. I have free games you can make materials like dice or marbles. My extensive dream journal with over 750 entries is probably one of the largest online dream 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. Best website. The only ads, my own creations from 2002, are in one isolated area. My site is best viewed at 800x600 resolution at true color (24 or 32-bit color) 1024x768 is optimal.

Lost? Don't know how to navigate? I've explained [the basic design](#) of my site (needs to be updated) and how it's organized from page to page. The category

Note: My website is currently undergoing a redesign so some pages, especially those that haven't been updated in at least one year, will have a somewhat ending with ".html" are the very old documents not having been updated since July of 2005 at the latest.

2 News and latest updates

My site contains the latest news and updates in more detail than described in this page and events going back when I began it in late 2004. This page contains



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](#)

[[IBM home page](#) | [Order](#) | [Search](#) | [Contact IBM](#) | [Help](#) | [\(C\)](#) | [\(TM\)](#)]

http://www.ibm.com/IBM/ Last visited: 1996-10-21

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



THE ROOF

Welcome to the new, improved Roof!

Hi! I'm the monkey who lives on the Roof, and this, obviously, is my roof! I audience gaps in amazement? Yup! I sure did!

And now, the revised Table of Contents:

DOOM!!!

PAGE ONE: "SIGHT" I WILL EXPLAIN THIS one more TIME! THE PAGE YOU' Duh!

PAGE TWO: KUNST

PAGE THREE: MISCELLANEOUS

PAGE FOUR: INTERESTING WHOSE HEAVY WHAT HOPES AND HOPES

PAGE FIVE: FANTASY AND FLAMING STUFF



0306
www.UltimateCounter.com





This space for rent..

[Home](#)

[Joachim](#)

[Susanne](#)

[Sebastian](#)

[Christopher](#)

[Daniel](#)

[Kimba](#)

[Tammy](#)

[Max](#)

[Unsere Feste](#)

[Fundgrube](#)

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.



World Wide Web

The WorldWideWeb (W3) is a universe of documents.

Everything there is online about a project, [Mailing lists](#), [Policy](#)

What's out there?

Pointers to the world's

Help

on the browser you are

Software Products

A list of W3 project co

[Mail robot](#), [Library](#))

Technical

Details of protocols, fo

Bibliography

Paper documentation o

People

A list of some people

History

A summary of the history of t

How can I help?

If you would like to support t

Getting code

Getting the code by [anonymo](#)

Take a look at a demo of one of our latest technologies, ProVision.

BBN ON THE WORLD WIDE WEB

We get the Internetworking for your business

[Press Releases](#) [Investor Relations](#) [Jobs](#)
[Customer Connection](#) [Business Partners](#) [BBN Features](#)

BBN PROVISION [ALSO INSIDE](#)
[BBN Planet](#)
[Customer Support](#)
[Online!](#)

[BBN Corporation 1996 Annual Meeting](#)
[BBN Planet Network Map](#)

Welcome to FIAT

Benvenuti sul Web-server FIAT !

Il server è attualmente in sviluppo: la home-page definitiva, per ora riservata ai gestori del server, sarà disponibile prossimamente.

Per ulteriori informazioni Vi invitiamo a visitare il server di ITS s.r.l. oppure potete scrivere a webmaster@its.it

FIAT Copyright © 1995.

<http://www.fiat.it>

All contents © 1995

• **Student Information**
Information of interest
[Undergraduate Admiss](#)

[Done](#)

<http://www.cocacola.com>

So Today *everybody* Shares Some Data...



book Home Profile Friends Inbox Thorsten Strufe

DOPPLR FOR YARICK Overview | Your connections | Your activity to V!!! on Wednesday Top Friends

Where Next? { You have no trips planned.

LinkedIn People | Jobs | Answers | Companies | Account & Settings | Help | Sign Out | Language

Account Type: Basic | Upgrade »

Search People Search Advanced

Applications BETA

FAQ | Feedback | Browse More Applications

My Travel

Application Preview

Your upcoming trips
December 2008
New York City, NY — Dec 13 - Dec 28
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 channeled my travel life..." – WNRC New York TV

half of 2009, be sure to add any past trips during 2009 that you might have missed. Add your past trips here.

View your full journal

Your Personal Annual Report

Search People Search Advanced

Application Info

Developed By TripIt Inc.

Category Utility

When you add an Application, you are allowing it to access profile information about you and your connections which the Application uses to provide the service. This and other information you provide to the Application may be displayed publicly depending on the Application's functionality.

By proceeding, you confirm your acceptance of the [LinkedIn Terms of Use](#) and you acknowledge that your use of the Application is subject to the Application developer's terms of use and to the Application developer's privacy policy.

Amadeus

MUSIC LOVERS, SNL, I France à l'ONU direct loser we love to hate,

TU Darmstadt, FG P2P, Th. Strufe P2P 5-1: Online Social Networks 11

...with calculated side effects...



The Joy of Tech™

by Nitrozac & Snaggy

CBC news

Home World C
BC Calgary Ed

LICENSE | EMAIL

Depre
over 1

Last Updated: Sa
CBC News

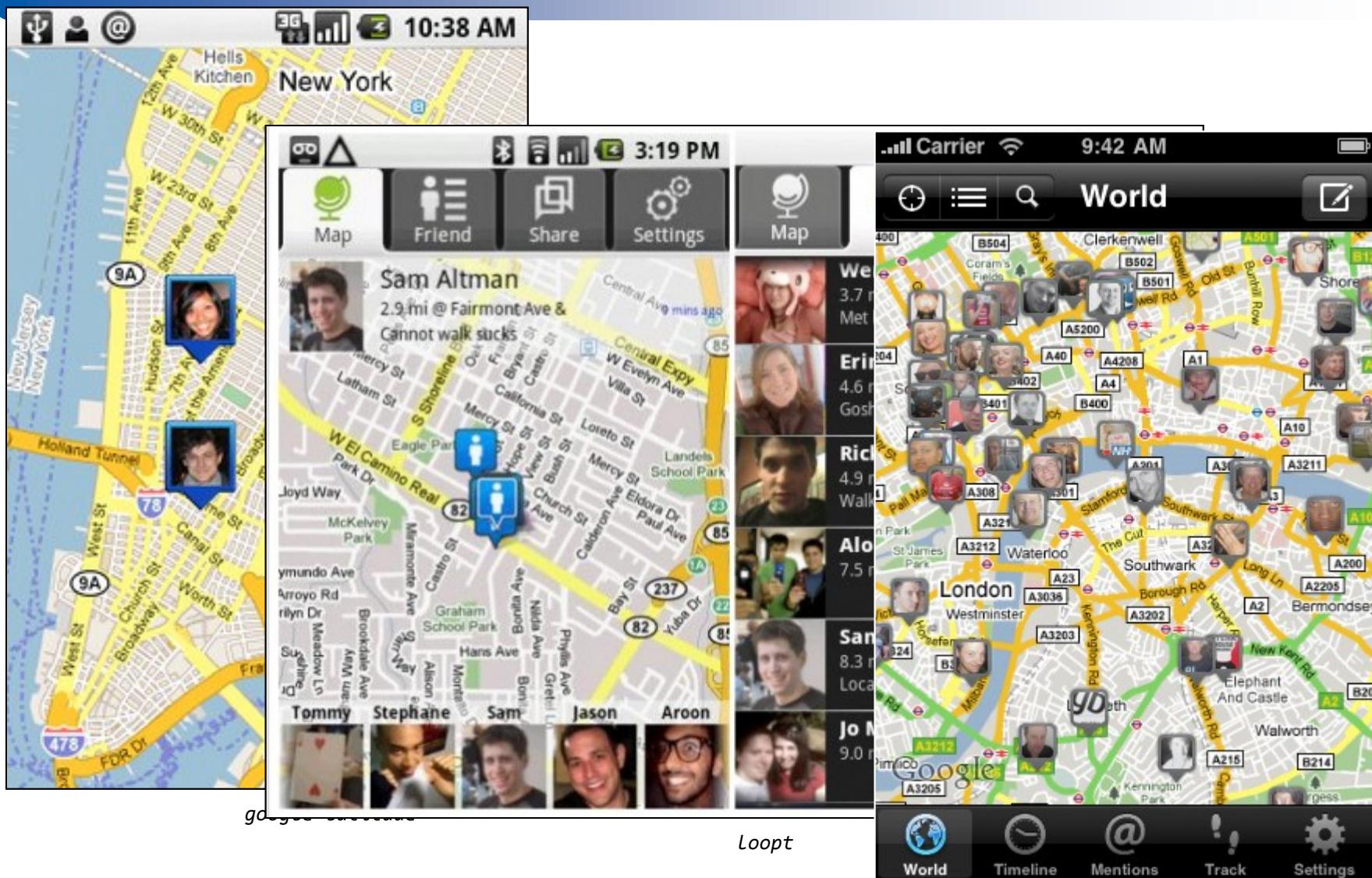
A Quebec W
employer's i

Nathalie Blan
holiday durin

pictures E
having a
— eviden



...increasingly immersive to daily life...



„Volkszählung“ 1987

- Penultimate public census in Germany
- Scheduled for 1981 (delayed)
- Significant public opposition
 - Fear of a surveillance society
 - The transparent citizen („gläserne Gesellschaft“)
 - Bounty for discovered Germans
 - Appeal for civil disobedience
 - Finally accomplished in 1987
 - Consequence: „25% inherent gap“
 - Significant gap between census and reality
 - (So let's just get a unifying tax)

VOLKSZÄHLUNG 1987
Personenbogen

Bitte so markieren:
Belegnr. 34 131 505 Lfd. Nr. der Person 5

Rechtsgrundlage:
Siehe Haushaltsumbogen oder Erläuterungsblatt,
die Bestandteile der Erhebungsvordrucke sind.
Stichtag: 25. Mai 1987

Falls Sie einen Abschluß an einer allgemeinbildenden bzw.
berufsbildenden Schule/Hochschule haben:
Falls Sie eine praktische Berufsausbildung
(z. B. Lehre) abgeschlossen haben:
Falls Sie zur Zeit tätig als

1 Geburtsangaben

a) Geburtsjahr
b) Geburtsmonat 1. Januar bis 24. Mai
25. Mai bis 31. Dez.

2 Geschlecht

männlich
weiblich

3 Familienstand

ledig
verheiratet
verwitwet
geschieden

4 Rechtliche Zugehörigkeit zu einer Religionsgesellschaft

Römisch-katholische Kirche
Evangelische Kirche
Evangelische Freikirche
Jüdische Religionsgesellschaft
Islamische Religionsgemeinschaft
andere Religionsgesellschaften
keiner Religionsgesellschaft rechtlich zugehörig

5 Welche Staatsangehörigkeit haben Sie?

deutsch
griechisch
italienisch
übrige EG-Staaten
jugoslawisch
türkisch
sonstige/keine

6 Wird von Ihnen noch eine weitere Wohnung (Unterkunft/Zimmer) in der Bundesrepublik Deutschland einschließlich Berlin (West) bewohnt?

nein
ja

7 Sind Sie erwerbstätig?

ja: Falls ja
Mehrere Antworten möglich
Falls ja:
a) Für Verheiratete, die nicht dauernd getrennt leben:
Ist die hierige Wohnung die vorwiegend benutzte Wohnung der Familie?
nein
ja
b) Für alle übrigen Personen:
Ist die hierige Wohnung die vorwiegend benutzte Wohnung?
nein
ja
c) Außerdem für Erwerbstätige, Schüler/Studenten:
Gehen Sie vorwiegend von der hierigen Wohnung aus zur Arbeit oder Schule/Hochschule?
nein
ja

nein
ja

Volzeit (über 36 Std.²) in der Woche
Teilzeit (bis zu 36 Std.²) in der Woche
arbeitslos, arbeitsuchend
nicht erwerbstätig
Hausfrau, Hausmann
Schüler(in), Student(in)

8 Leben Sie überwiegend von

Erwerbs-, Berufstätigkeit
Arbeitslosengeld-, hilfe
Rente, Pension
eigenem Vermögen, Vermietung, Verpachtung, Anteil an
Zuwendungen, Unterhalt durch Eltern, Ehegatten usw.
sonstigen Unterstützungen (z. B. Sozialhilfe, BAföG)

NUR VOM STATISTISCHEN LANDESAMT AUSZUFÜLLEN

Geburtsjahr (1 a)
2. 3. u. 4. Stelle
Hauptfachrichtung (10 b)
Erlebter Beruf (11 a)
Dauer (11 b)

Arbeitsstätte, Schule/Hochschule (Pendler)
Land (12) Gemeinde (12) Straße (12) Hausnummer (12)

Wirtschaftszweig (16)
Ausgeübte Tätigkeit (17)

FÜR ERWERBSTÄTIGE

15 Sind Sie zur Zeit tätig als

Facharbeiter(in)
sonstige(r) Arbeiter(in)
Angestellte(r)
Ausbildende(r)
Beamter/Beamtin, Richter(in), Soldat, Zivildienstl.
(auch Beamtenank.)
Selbständige(r)
mit bezahlten Beschäftigten
ohne bezahlte Beschäftigte
mitteilende(r) Familienangehörige(r)

16 Zu welchem Wirtschaftszweig (Branche, Behörde) gehört der Betrieb (Firma, Dienststelle), in dem Sie tätig sind?

17 Welche Tätigkeit, welchen Beruf üben Sie aus?

18 Falls Sie eine Nebenerwerbstätigkeit ausüben, handelt es sich um eine

landwirtschaftliche
nichtlandwirtschaftliche

... completely willingly...



HVA/HAs. Abl.	A K G	Ge
BVNr	G e r a	M 1 S
— 333 —		
18. MAI 1982		
Tgb Nr. 552		
Weltzeit		

Erfassungsbeleg

(Mit Schreibmaschine oder in Druckform ausfüllen)

012742

1. Grunddaten zur Person

1.1. W e b e r , Martin :

Name, Vornamen, Geburtsname, weitere Namen

1.2. Staatsangehörigkeit

1.3. Staatsangehörigkeit, Pseudonym, Spitzname

1.4. Zeitraum der Dienst und Dienststelle des Dienstes

Januar 1940

1.5. Deutschland, Ottendorf

1.6. männlich

1.7. deutsch

1.8. BRD

1.9. 1940-1982

1.10. 1940-1982

1.11. Weber, K

Ehefrau

2. Zuordnung zu der (den) Personenkategorie(n)

PK. 1 Wird der Verbindung zum Geheimdienst

PK. 4 Teilnehmer an der Tätigkeit einer so

Organisation

BStU	
000005	
I. Beziehung zu Geheimdiensten und Zentren der politisch-ideologischen Diversen	
II. Beziehung zu Geheimdiensten und Zentren der politisch-ideologischen Diversen	
BS	
0000	

6.6. Dienststelle/Funktion	6.7. Dienststellen Hauptaufgabe Hauptperson	6.8. Zeitraum der Zugänglichkeit
7. Wohnt		
7.1. BRD		
7.2. Bayern, Oberfranken		
7.3. 8642 Ludwigsstadt		
7.4. Am Sommerberg Nr. 12		
7.5. 1600		
7.6. -1977-1981-		

BStU	
000007	
10. Gegen die Person durchgeführte Maßnahmen	
10.1. Charakter der Maßnahme	
10.2. Von wem wurde die Maßnahme durchgeführt	
10.3. Wie wurde die Maßnahme durchgeführt	
11. Zusätzliche Angaben	
3.5. Übergibt Angaben zu Fahnenflüchtigen an eine amerikanische Dienststelle	
5.5. Führt Erstvernehmungen durch und organisiert Intensivkontrollen gegen DDR-Bürger	

12. Die Angaben beziehen sich auf das Jahr 1981

13. Die Person ist erfaßt für die DE BV Gera, Abteilung VI

14. Ausgabeverfahrensweise: A, B, C, D, E*

15. Abstimmung zur Nutzung: erforderlich/nicht erforderlich*

16. Dok.-Nr. der Personeninformation DUG

Bestätigt

Leiter der AID AGG/Beauftragter der HVA



- What is Social Media anyways?
 - ***Web-based (user generated) content*** sharing
 - Knowledge about the posting ***individual***
 - Plus the ***relationship*** between the users
- Relationship?
 - „Friend“, „Acquaintance“, „Idol“...?
 - Uni- vs. bi-directional („connect“ vs. „subscribe“)

System Classes



- Content Communities
- Blogs/Microblogs
- Virtual Gaming Worlds
- Virtual „Social“ Worlds
- Collaborative Projects
- Online Social Networks
(read: all of the above ;)

- youtube

- Twitter

- WoW

- 2nd life

- Wikipedia

- FB, g+ (?),

*Are last.fm, pandora
social media?*

*What about
amazon/ebay,
then?*



Online Social Networks (*Social Networking Services*)

“web-based services that allow individuals to

- (1) construct a public or semi-public profile within a bounded system,*
- (2) articulate a list of other users with whom they share a connection, and*
- (3) view and traverse their list of connections and those made by others within the system.”*

- Is this really all?

*“the **communication** of participants through **message exchange**, **commenting** on the profiles of others (or previous interactions, e.g. in recommendations), which merely is a message exchange with the aim to annotate the addressed profile, and the wealth of **applications** (starting from simple ‘‘poking’’ mechanisms to a variety of ‘‘gift’’ and ‘‘likeness’’ applications for interactions between users)”*

Source1: boyd et al.: Social Network Sites

Source2: cutillo et al.: Privacy Preserving Social Networking through Decentralization



- 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 service

- Private and personal

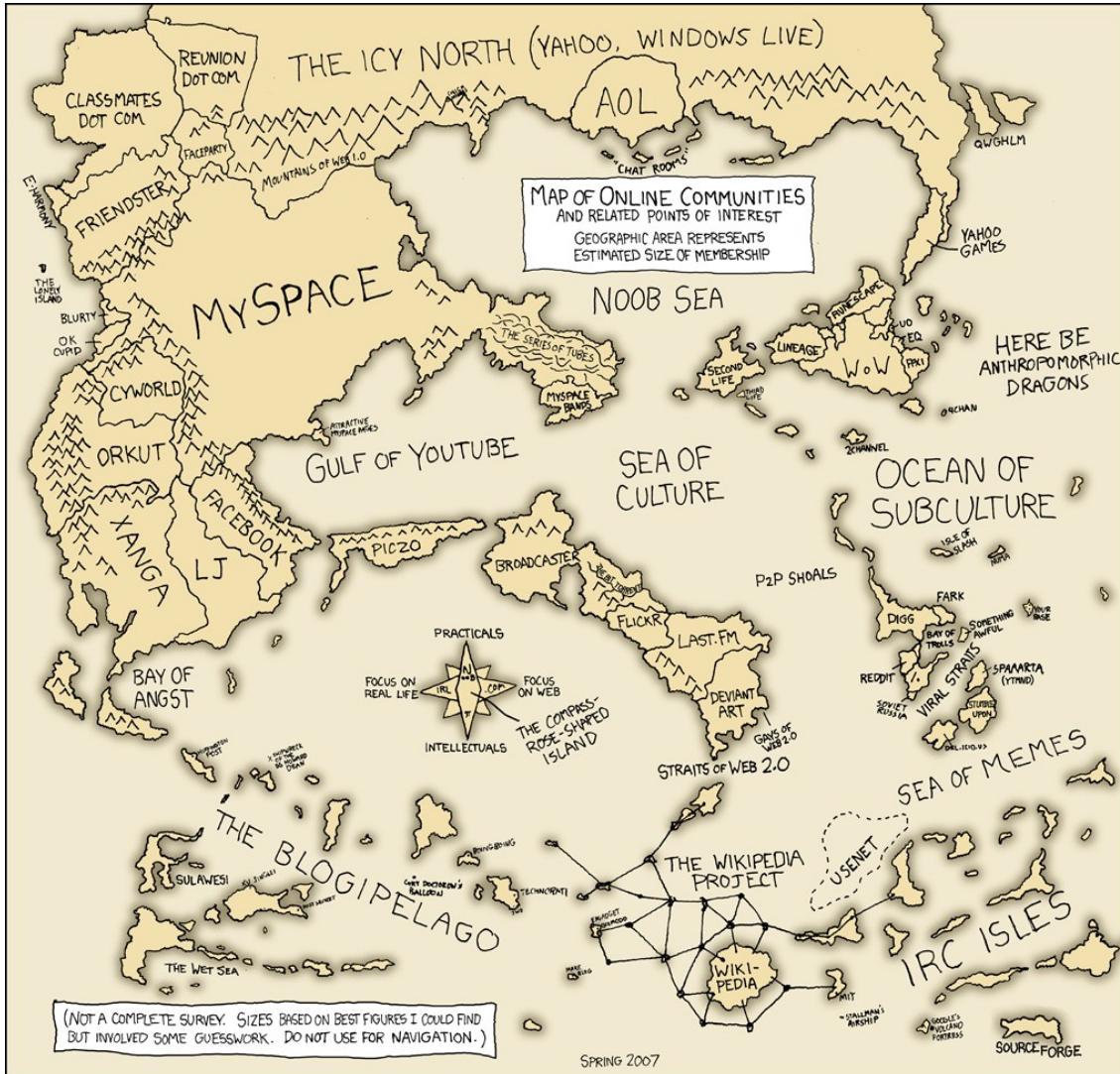
- Niche/tailored service
 - “Micro blogging”: Twitter
 - Business trips and meetings
 - Location-based achievements

The screenshot shows a Gowalla user profile for "Mr. Crown" from Hamburg, Germany. The profile includes a profile picture of a person on a horse, a summary of 2412 stamps and 79 pins, and a bio mentioning they are a software solution developer and architect, a voluntary member of the International Search and Rescue Advisory Group (INSARAG), and a member of the German Federal Agency for Technical Relief (THW). The user has an "Add as Friend" button. The main feed displays a timeline of recent check-ins:

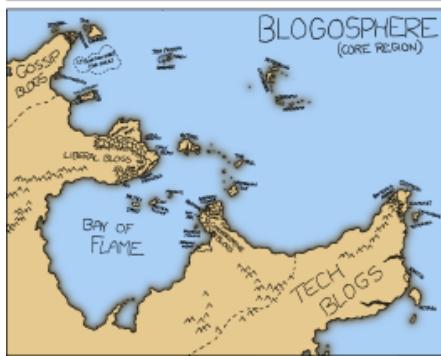
- Mr. Crown checked in at Funkturm Berlin, received some Bats, and swapped some Bats for a Yo-yo.
- Mr. Crown checked in at Rafih, swapped a Yellow Kayak for a Fieldnotes & Books.
- Mr. Crown checked in at IFA Berlin, Hall #20, and took a photo.
- Mr. Crown checked in at IFA Berlin, Hall #20, and took a photo.

On the left, a sidebar shows a navigation menu with categories like Business, Personal, Want, Have, Interest, Organization, and Professional, with Professional currently selected. Below the sidebar, a timeline shows check-ins from April (1 year ago) to September (6 years ago). The right side of the screen shows a "Top Places" list with 10 entries, including Faisalabad International Airport, SFO San Francisco International Airport, Central Park, Apple HQ, Adobe Systems, IFA Berlin, LIVESTRONG Challenge, Apple Flagship Store, Fenwick & West, and Bank of America Building.

Map of Online Communities



Source: Randall Munroe, XKCD.org



ABOUT THIS MAP

COMMUNITIES RISE AND FALL, AND TOTAL MEMBERSHIP NUMBERS ARE NO LONGER A GOOD MEASURE OF A COMMUNITY'S CURRENT SIZE AND HEALTH. THIS UPDATED MAP USES SIZE TO REPRESENT TOTAL SOCIAL ACTIVITY IN A COMMUNITY — THAT IS, HOW MUCH TALKING, PLAYING, SHARING, OR OTHER SOCIALIZING HAPPENS THERE. THIS MEANT SOME COMPROMISING OF APPLES AND ORANGES, BUT I DID MY BEST AND TRIED TO BE CONSISTENT.

ESTIMATES ARE BASED ON THE BEST NUMBERS I COULD FIND, BUT INVOLVED A GREAT DEAL OF GUESSWORK, STATISTICAL INFERENCE, RANDOM SAMPLING, NONRANDOM SAMPLING, A 20,000-CELL SPREADSHEET, EMAILING, CADDING, TEA-LEAF READING, GOAT SACRIFICES, AND GUT INSTINCT (I.E. MAKING THINGS UP).

SOURCES OF DATA INCLUDE GOOGLE AND BING, WIKIPEDIA, ALEXA, BBS-BOARDS.COM, STUMBLEUPON, WORDPRESS, ANSWER, EVERYWEB, STATISTICA, PAGE 1 (OULD FIND), PRESS RELEASES, NEWS ARTICLES, AND INDIVIDUAL SITE EMPLOYEES. THANKS IN PARTICULAR TO FOLKS AT LAST.FM, LIVEMUSIC.JOURNAL, REDDIT, AND THE NEW YORK TIMES, AS WELL AS SYSTEMADMIN AT A NUMBER OF SITES WHO SHARED STATISTICS ON CONDITION OF ANONYMITY.



Two words on: *Information Sovereignty*



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



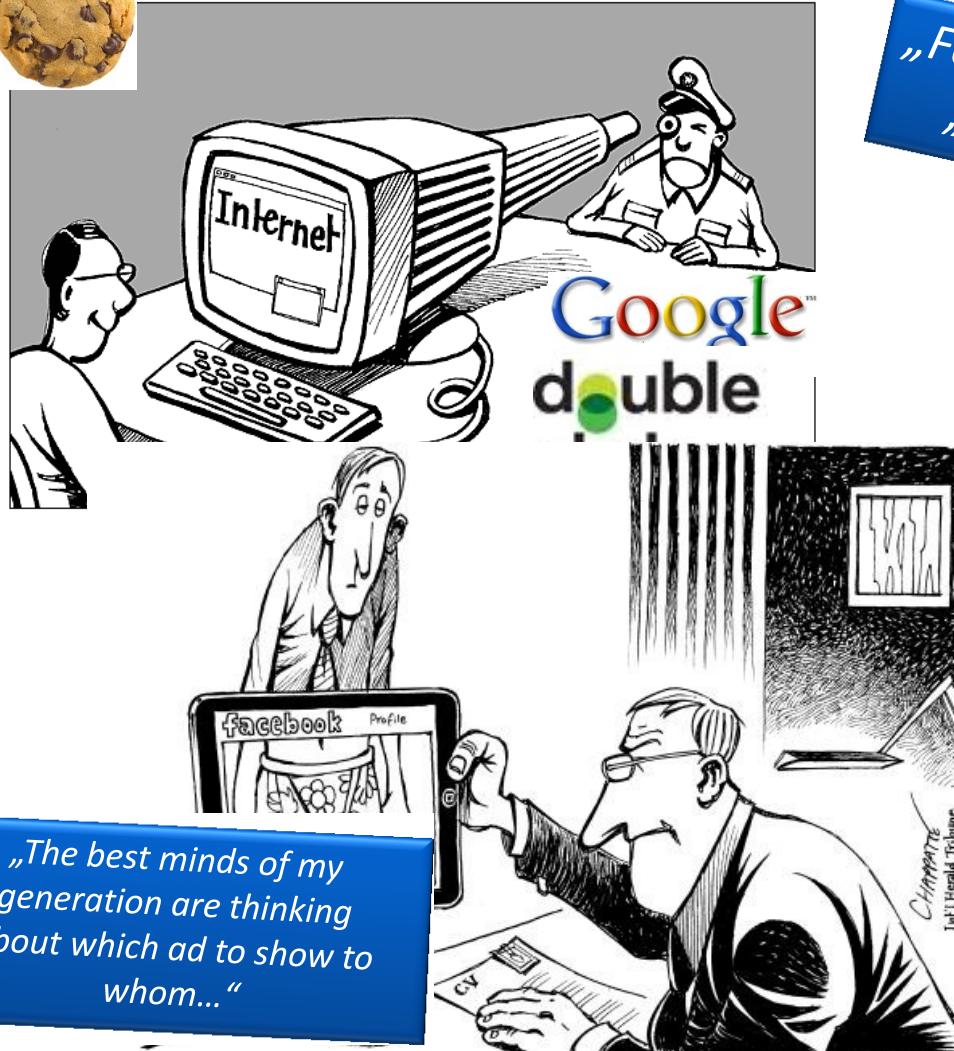
*i.o.w: to which effect? aka.
„Power to the people!“ (O RLY?)*



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

The deer now have guns...

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.“

The Deer May have Guns, but...



it's known ***who they are***
with ***whom*** they're ***friends***
whom they're ***talking to***
what they ***think*** and ***want***

even 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!)

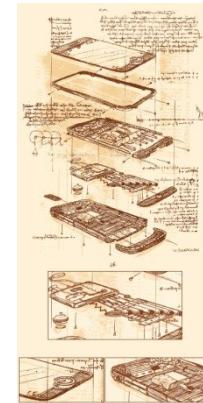
Just license them, don't give them any ownership!

Control/surveil their music

Their movies

Their social network...

...their life...





*„And it's a Unix
under the hood!“*



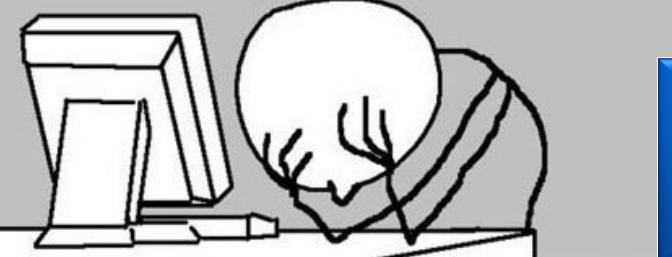
*„Yeah, they're so cool,
no effort, no hassle!“*



*„Worst case:
you can jailbreak...“*



Think differ



*Enough polemics,
back to the topic!*



Back to Social Networks – and: Why Bother?



Top 10 Web Brands for January 2010 (U.S., Home and Work)

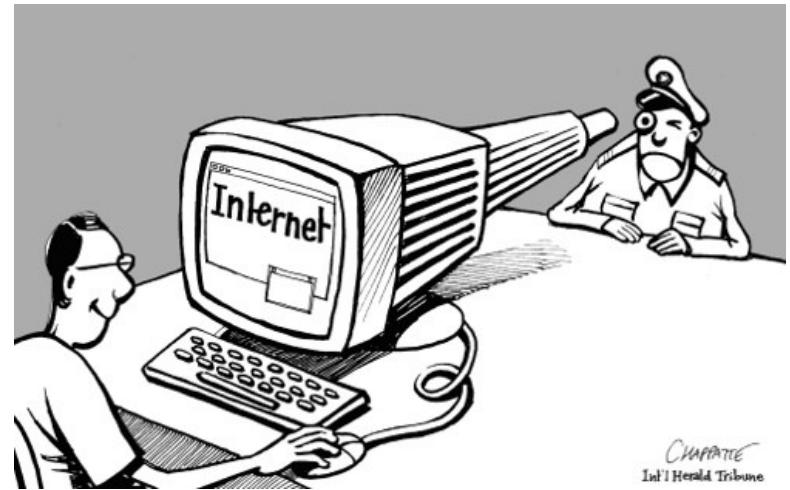
RANK	Brand	Unique Audience (000)	Time Per Person (hh:mm:ss)	MOM UA % Change	MOM Time % Change
1	Google	152,708	1:23:54	4.10%	-16.90%
2	Yahoo!	134,561	2:00:11	4.30%	-26.80%
3	Facebook	116,329	7:01:41	5.80%	9.70%
4	MSN/WindowsLive/Bing	109,425	1:25:22	1.20%	-18.10%
5	YouTube	99,525	1:02:27	7.60%	-10.30%
6	AOL Media Network	82,306	1:01:14	-6.80%	-57.80%
7	Wikipedia	64,917	0:15:59	10.70%	-2.70%
8	Fox Interactive Media	62,112	1:23:28	1.00%	-9.10%
9	Amazon	60,772	0:22:34	-8.60%	-32.90%
10	Ask Search Network	57,776	0:12:35	10.70%	-11.40%

Source: The Nielsen Company

Characterizing User Behavior in OSN



- 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



XING

Logout | Invite contacts | Help & Contact | English ▾

Search by name, company, and more

Home Search Messages Contacts Groups Events Jobs Companies Best Offers

Thorsten Strufe
Prof. Dr.-Ing. (Dipl.-Inf.)
Fachgebiet Peer-to-Peer Netzwerke
TU Darmstadt
Hochschulstraße 10, 64289 Darmstadt, Germany
+49-6151-164557
Local time: 04:48 PM

Business details Confirmed contacts 83 About me Guestbook Applications

Personal information

Wants PhD students

Haves PhD positions in the fields of networking / distributed systems, especially security / DoS resilience of P2P systems and privacy in social networking services

Interests P2P online social networks, skiing, diving, kicker/footsie!, politics

Organizations GI, IEEE, Servas, Studienstiftung des deutschen Volkes

Professional experience

+ Add entry

Apr 2009 - present **Fachgebiet Peer-to-Peer Netzwerke** (Full time employee)
(1 year, 1 month)
TU Darmstadt, <http://www.p2p.tu-darmstadt.de>
Industry: Academia
JP

Feb 2008 - Mar 2009 **PostDoc Research Engineer** (Full time employee, Manager (Manager/Supervisor))
(1 year, 2 months)
Eurecom, Sophia-Antipolis, France, <http://www.sosoc.org>
Industry: Research, 51-200 employees
+ Add company size

Sep 2001 - Jan 2008 **wissenschaftlicher Mitarbeiter** (Full time employee)
(6 years, 5 months)
Technische Universität Ilmenau, <http://www.tu-ilmenau.de/fakia>
/Fachgebiet_Telematik.317.0.html
Industry: Research, Wissenschaft, Aus- und Weiterbildung
+ Add company size

Teilhaber, Leitung IT
Multiple-Choice GmbH
Industry

Searches related to my profile

- Members who recently visited my profile
- Members whose “wants” match my “haves”
- Members whose “haves” match my “wants”
- Members who know several of my contacts
- Current and former colleagues

Status update
Premium Members can enter a status update here (to let your network know you're at a conference or on vacation, etc.).

I'm a XING member because I want to ...
Let others know why you are on XING, and boost your networking potential.

Settings

Options

Confirmed contacts: 283

Thorsten Strufe's statistics
No Premium Membership
Member since: Jan 2005
Profile hits: 3,534
 Activity meter: 90%

My Groups

Visitors to my profile
(only visible to you)

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 “friends”/contacts

The screenshot shows the XING homepage. At the top, there is a navigation bar with links for Home, Search, Messages, Contacts, Groups, Events, Jobs, Companies, and Best Offers. To the right of the navigation bar, there are links for Logout and In. Below the navigation bar, there is a search bar with the placeholder "Search by name, company, and more". The main content area is titled "What's new in your network" and displays a list of recent activity items. Each item includes a small profile picture, the user's name, a brief description, and a "Send message" link. The items listed are:

- Tahssin Asfour präsentiert das neue design von www.pr-bote.de
Today - Status update · [Send message](#)
- Mario Melle joined the "Neuer Personalausweis (nPA) für ..." group.
Today
- Detlef Mämpe joined the "Army Network" group.
Today
- Dr. Thomas Rücker will be attending the "Netzwerkgruppe Innovation" event.
Today · [Send message](#)
- Stella Pieri is now connected to Julian Hühnermann (Bigpoint GmbH).
Today
- Only Premium Members joined the "Innovationsfinanzierung durch Business ..." group.
Yesterday
- Only Premium Members has changed her private address.
Yesterday · [Send message](#)
- Only Premium Members is now connected to Christoph Lühr (basilicom gmbh).
Yesterday

At the bottom right of the news feed area, there is a "more »" link. To the right of the news feed, there is a sidebar with various links and a "Grow your network" section.

- 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

Studying a Prominent OSN



- “xing” selected for the study
 - Business/professional OSN, similar to LinkedIn
 - 8 Mio users, mainly from central Europe
 - 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

Sampling and Monitoring



- Random sampling
 - Conducted random walks (25k, 5k, ~1k)
 - Wish for 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)

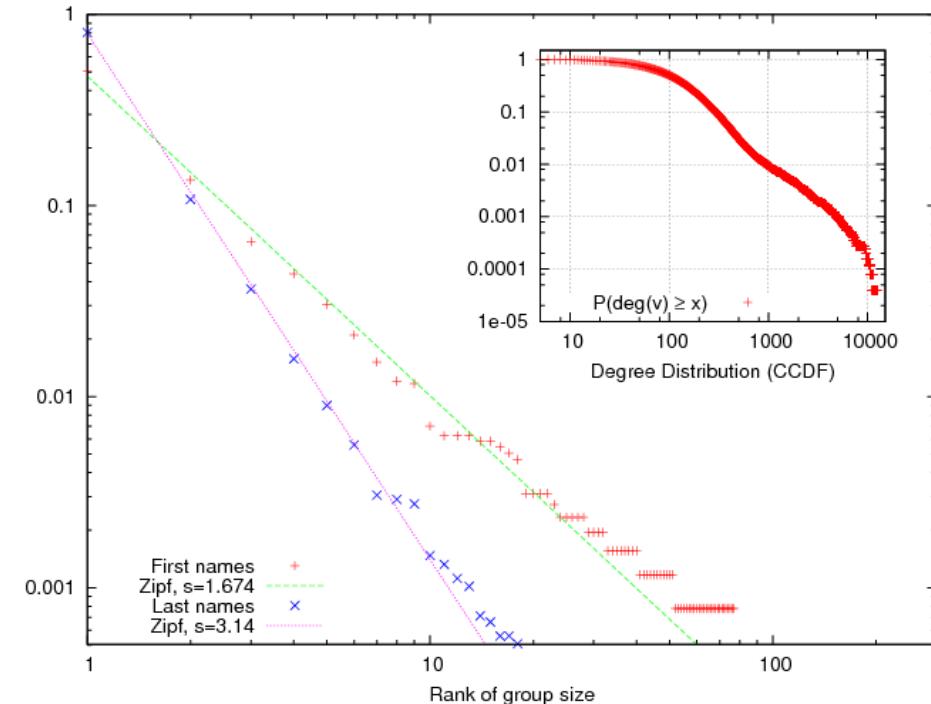
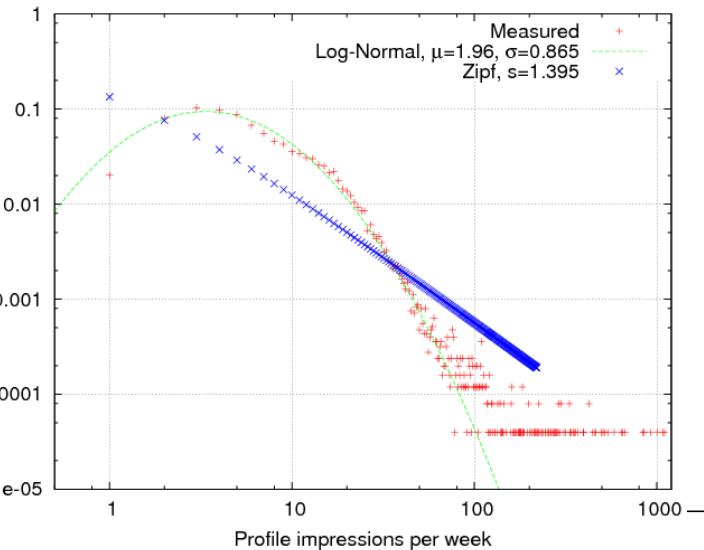


- Identifying Data
 - Name, Image, Gender
- CV
 - Current employment, universities attended, claimed spoken languages
- Interests
 - Interests as stated
 - Number of subscribed groups, subscribed groups, number of members in groups, number of messages in respective groups, languages of group
- Contact list information
 - Number of contacts
 - Complete list of contacts gathered infrequently
- Statistics
 - Registration date, number of profile impressions, activity meter
 - Timestamp of crawl

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. (\sim 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



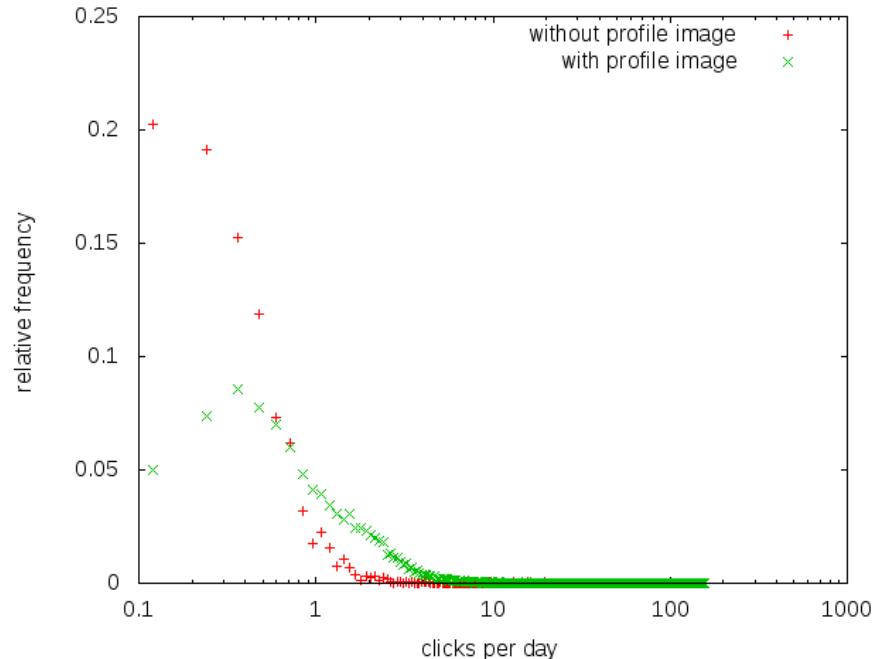
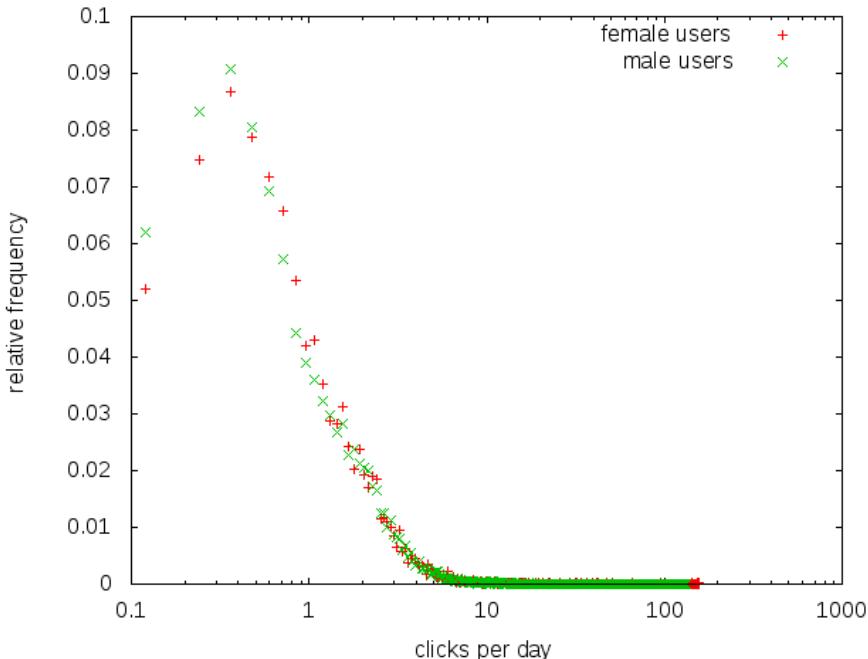


- Activity
 - “Activity meter” in profiles very coarse grained
 - Derived “profile alteration frequency” as alternative
 - Men are slightly more active than women (to both metrics)
 - Profiles without image belong to inactive users
- Membership in Groups
 - Wide range of group membership
 - Max 511, Mdn 3
 - >5k profiles are not registered to any group

Testing Groups of users (non-parametric)



- Testing popularity between groups (non-parametric)
- *“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)

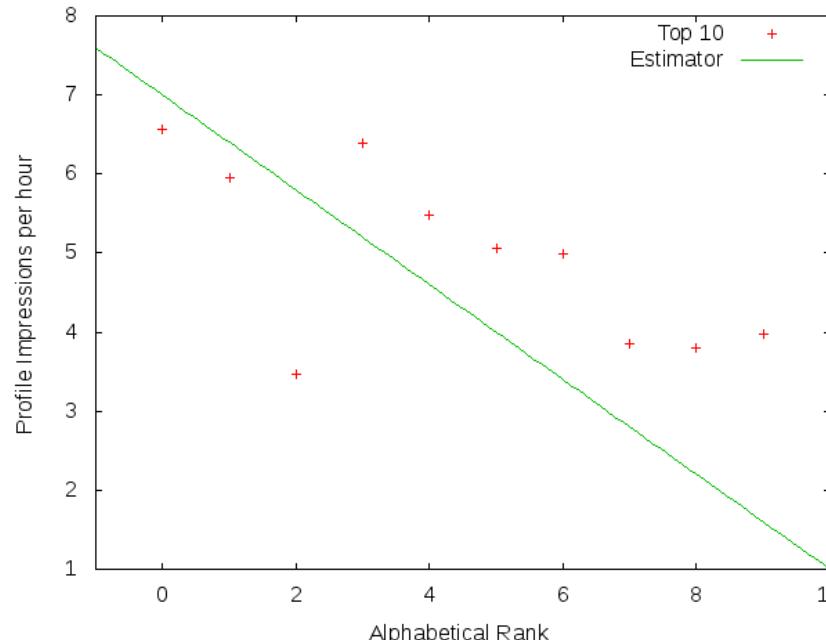


- “The profiles of active users are more attractive”
 - Activity as given on the profiles (“**activity meter**”)
 - $r \approx 0.17$, no noteworthy correlation
 - Activity measured in **group memberships**
 - $r \approx 0.37$ (higher for men, lower for women)
 - Activity meter is very coarse grained
 - 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 retrospective: Combination somewhat unsurprising: changes are published at friend's profiles...**

Correlational Tests (odds n' sods ;-)



- “Old/experienced profiles are viewed more frequently”
 - Preferential attachment / experience could lead to higher popularity
 - H_0 rejected, but $r \approx 0.11$, **no noteworthy correlation**
- “Last name starting with a letter late in the alphabet sucks...”
 - H_0 not rejected, there is no correlation.
 - *Taking the “rich-club”, however...*
 - Top 5% profiles: $r \approx -0.09$
 - 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!



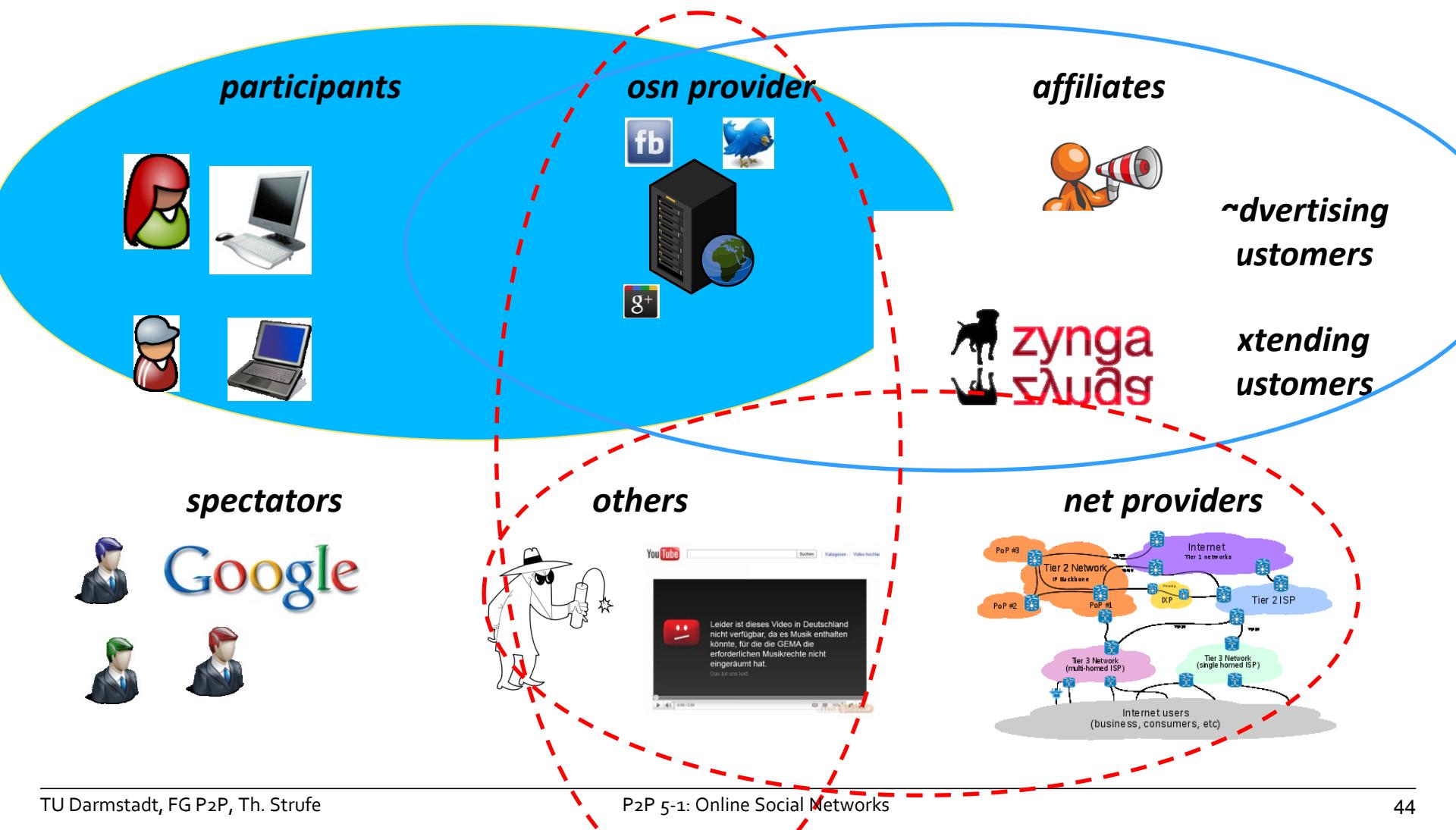
What is „***trust***“, in the first place?

- *Trust is the certainty that another party (a subject) is going to act, operate, behave **as expected**.*

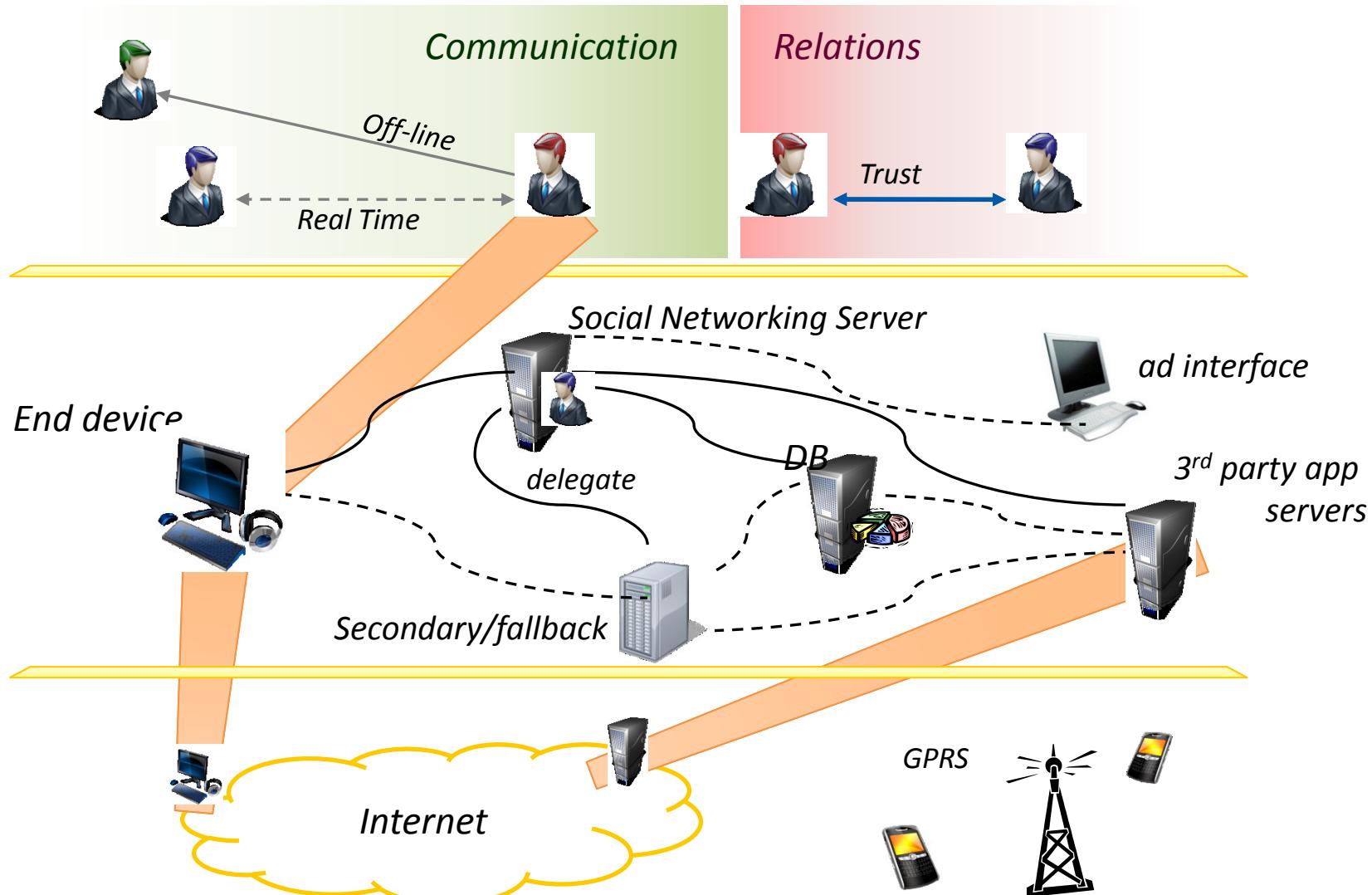
What is so special about trust in OSN?

- Service focuses on ***individuals*** (whom the user knows)
- and their ***relations*** (seeming correspondence to RL trust)
- ***Object*** of trust:
 - wealth of personally identifiable information (PII)
 - RW reactions to seemingly personal communications

Stakeholders in OSN



A Little Model



Access Model



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

Implicit

- SNP



Everything the installing user can see

- Affiliate



- Extended
- Advertising



*Not much (aggregates)
Unless they pay really well*

- ISP



*Everything their subscribers see/write
(until Nov 21st '12)*

Ignoring downstream abuse...

Types of Information



■ Explicit

- created content (+self descriptive)
- annotations/comments
- preferences/structural interaction (contacts, +1, etc)



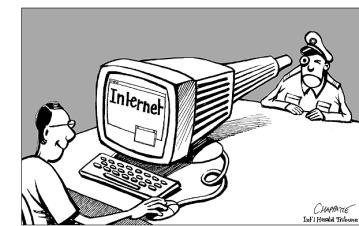
■ External

- interest/preferences (clickstreams through adnetworks, fb-connect)



■ Implicit

- inferable from environment (homophily)
- observable
 - *session artifacts* (time of actions), *interest* (retrieved profiles; membership in groups/ participation in discussions), *influence* (users)
 - *clickstreams*, ad preferences, exact sessions, *communication* (end points, type, intensity, frequency, extent), *location* (IP; shared; gps coordinates), *udid*



■ Processed/derived

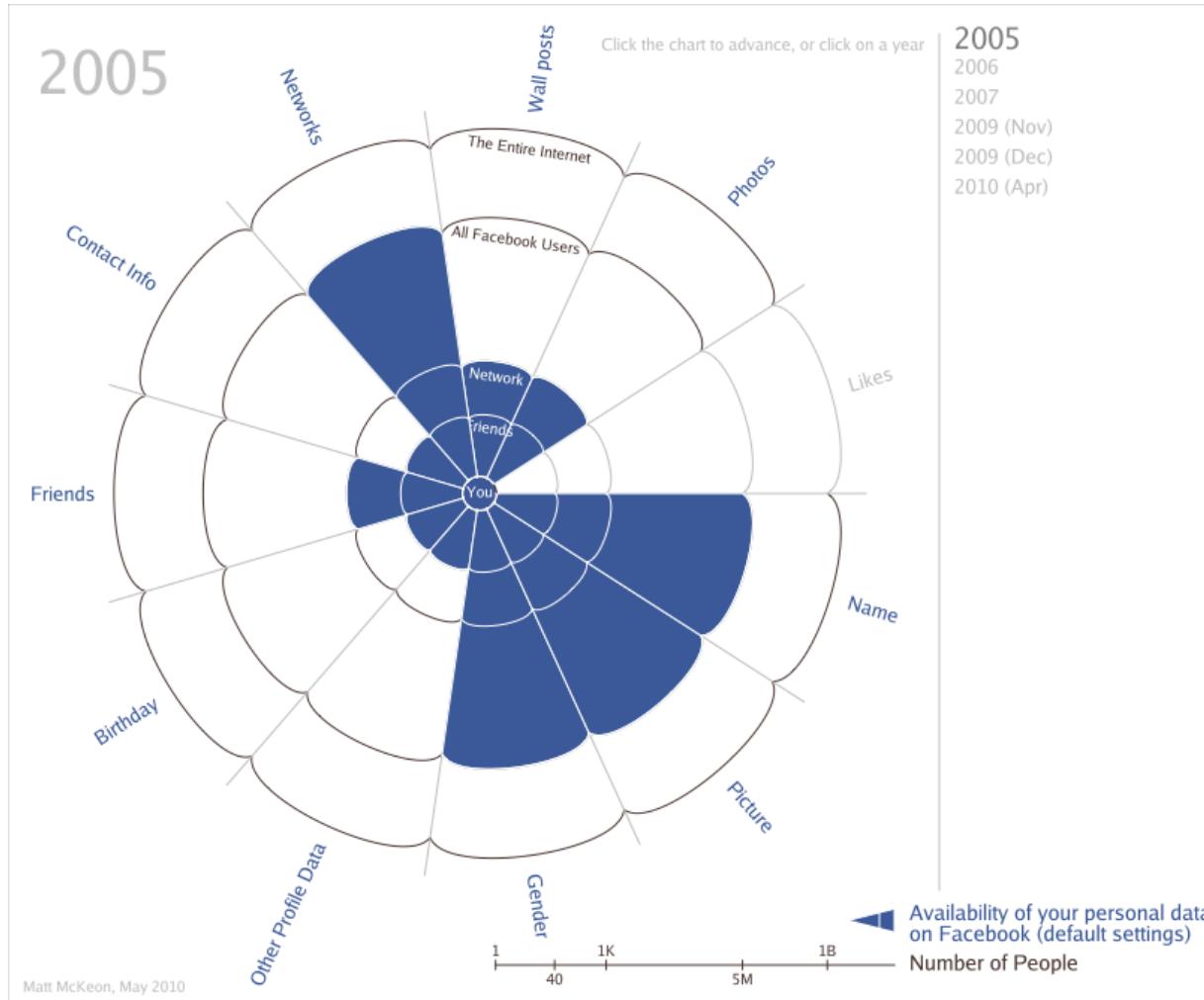
- preference models
- **image recognition models**
- aggregates (characteristic properties/ descriptors of user groups)

Facebook and Privacy Settings



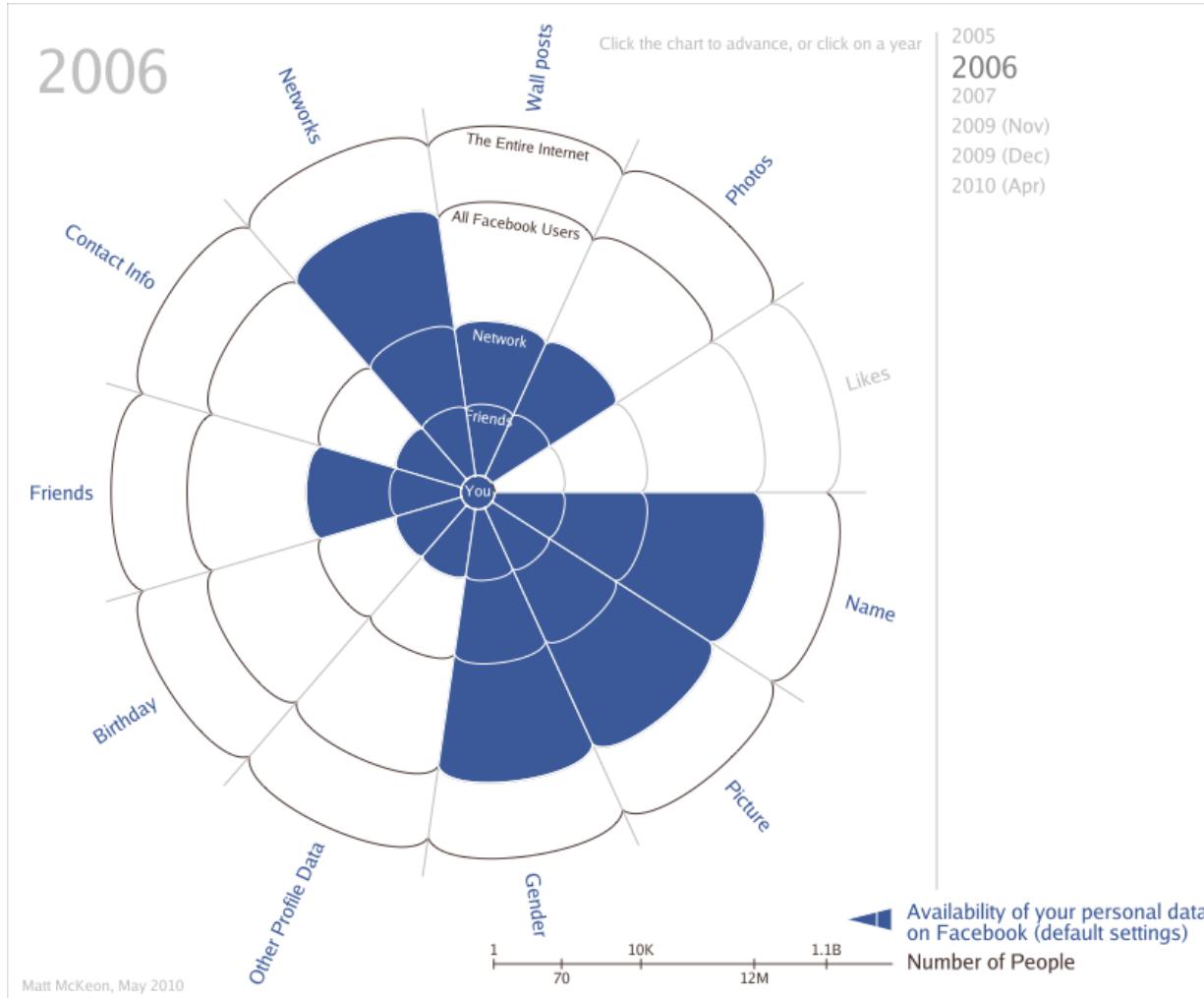
- Facebook largest and most successful OSN:
 - Over 800 Mio active users
 - 50 Mio user after 3 years (2004-2007)
 - 50 mio users per quarter since 2007
 - ½ Mio registrations ***daily***
 - De-facto reference
- Potential reasons for the dominance
 - Quick following of new hypes
 - “Apps” (Extensions to service, integration of 3rd party services)
 - “Friend feed”: construct virtual village (following twitter’s success)
 - Facebook places (following foursquare and gowalla’s success)
 - Stalking service (you don’t know who watched your profile => if this means success, users need to make their private data public!)

Facebooks Privacy-Evolution: 2005

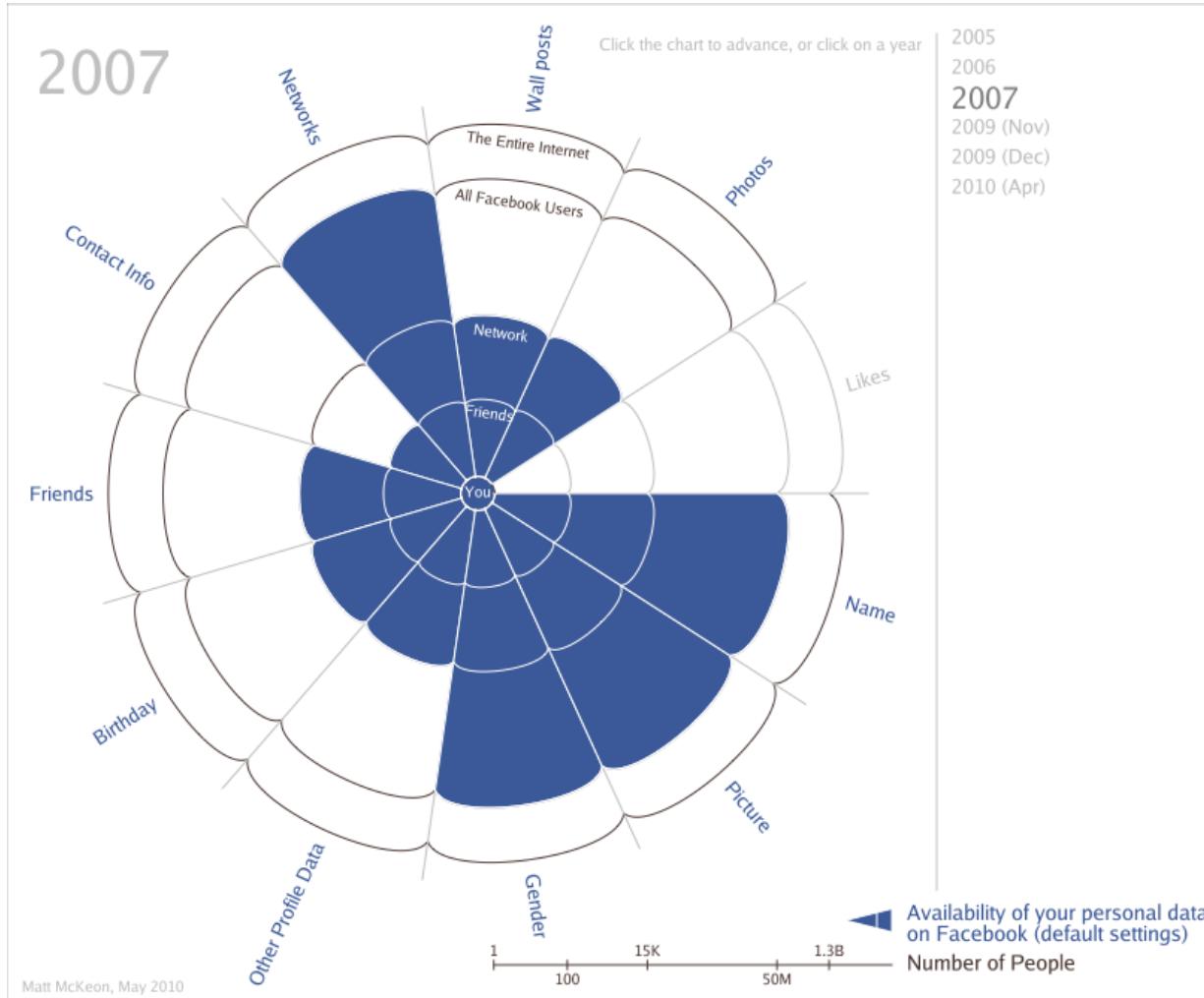


Source: McKeon

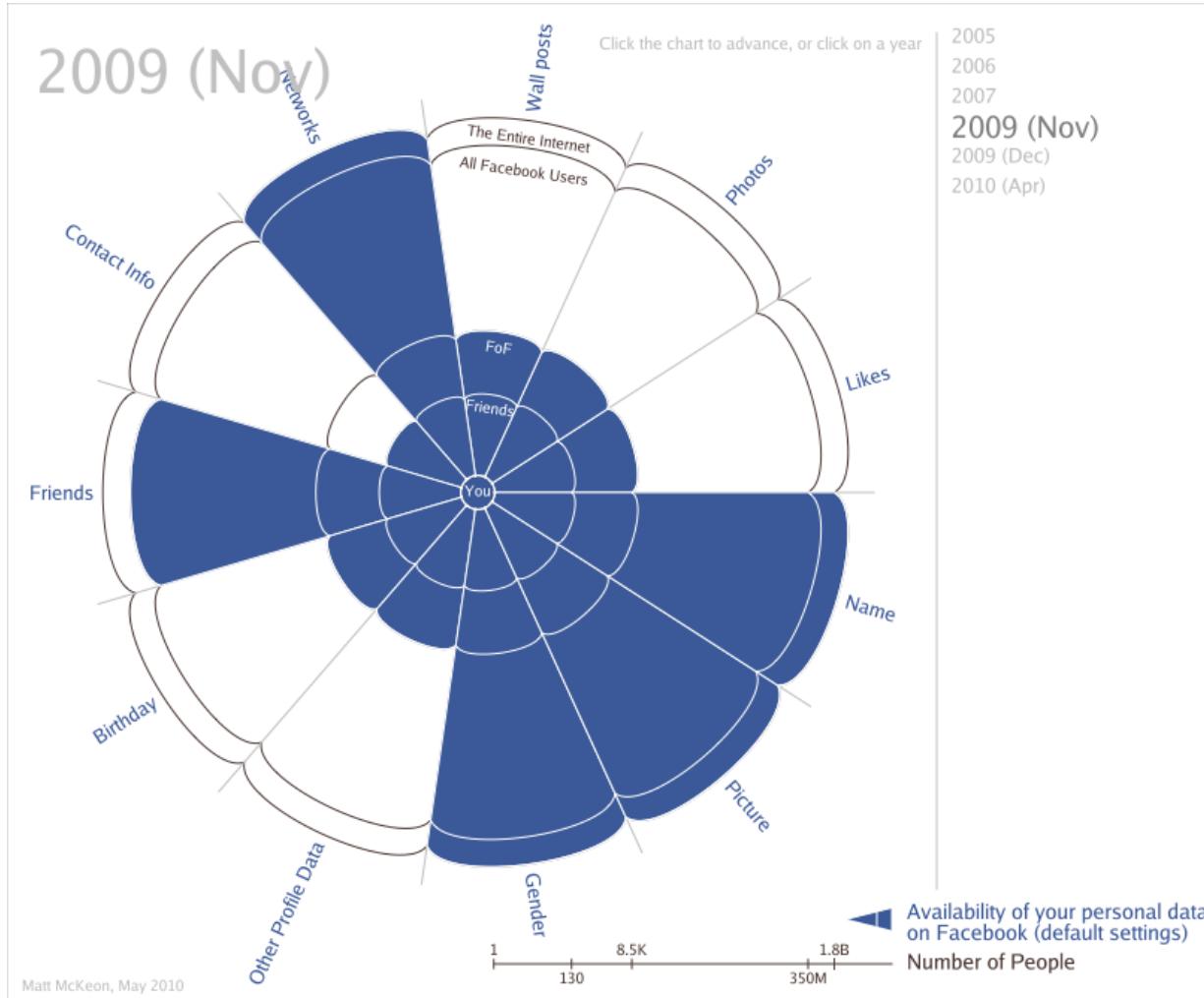
Facebooks Privacy-Evolution: 2006



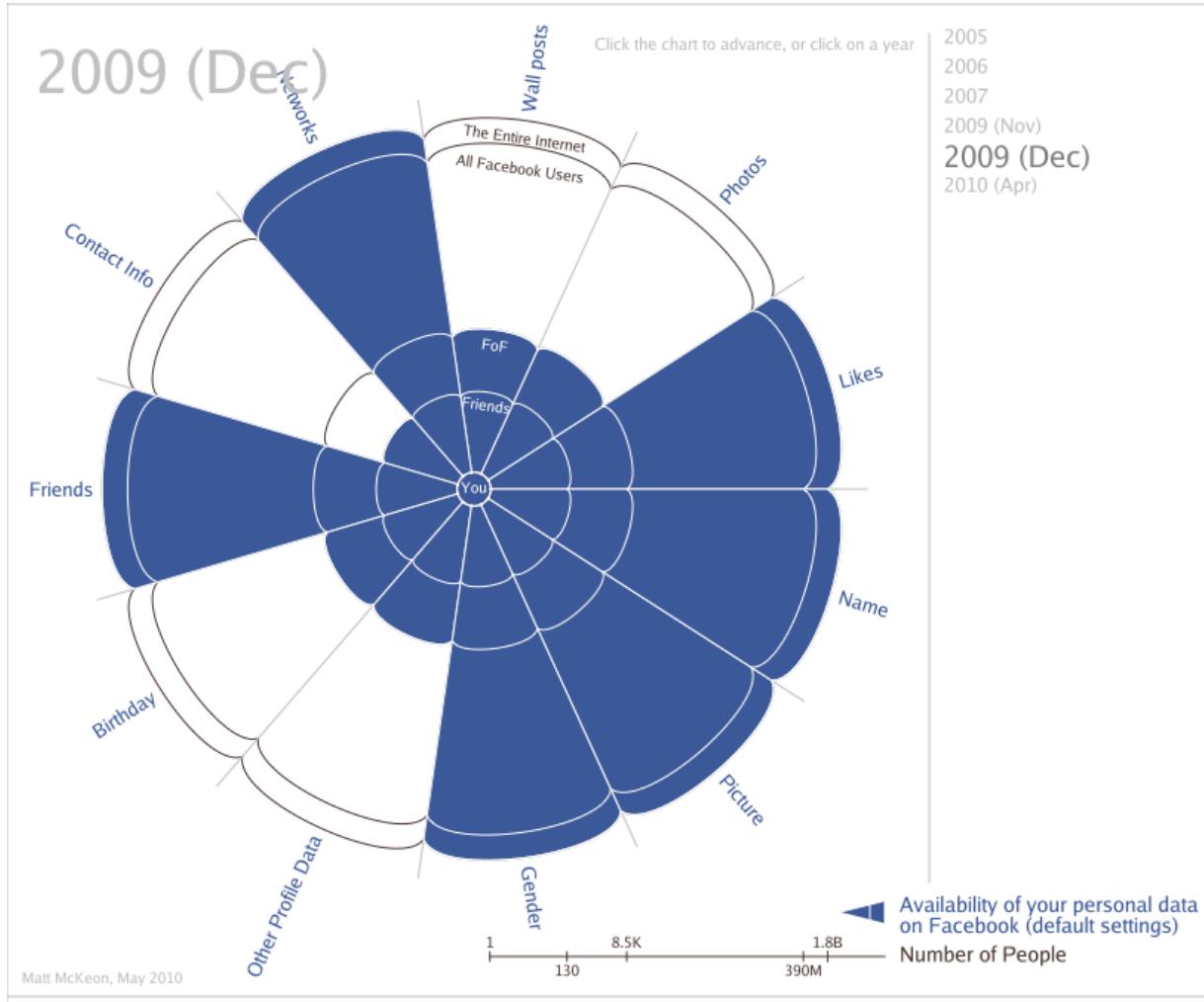
Facebooks Privacy-Evolution: 2007



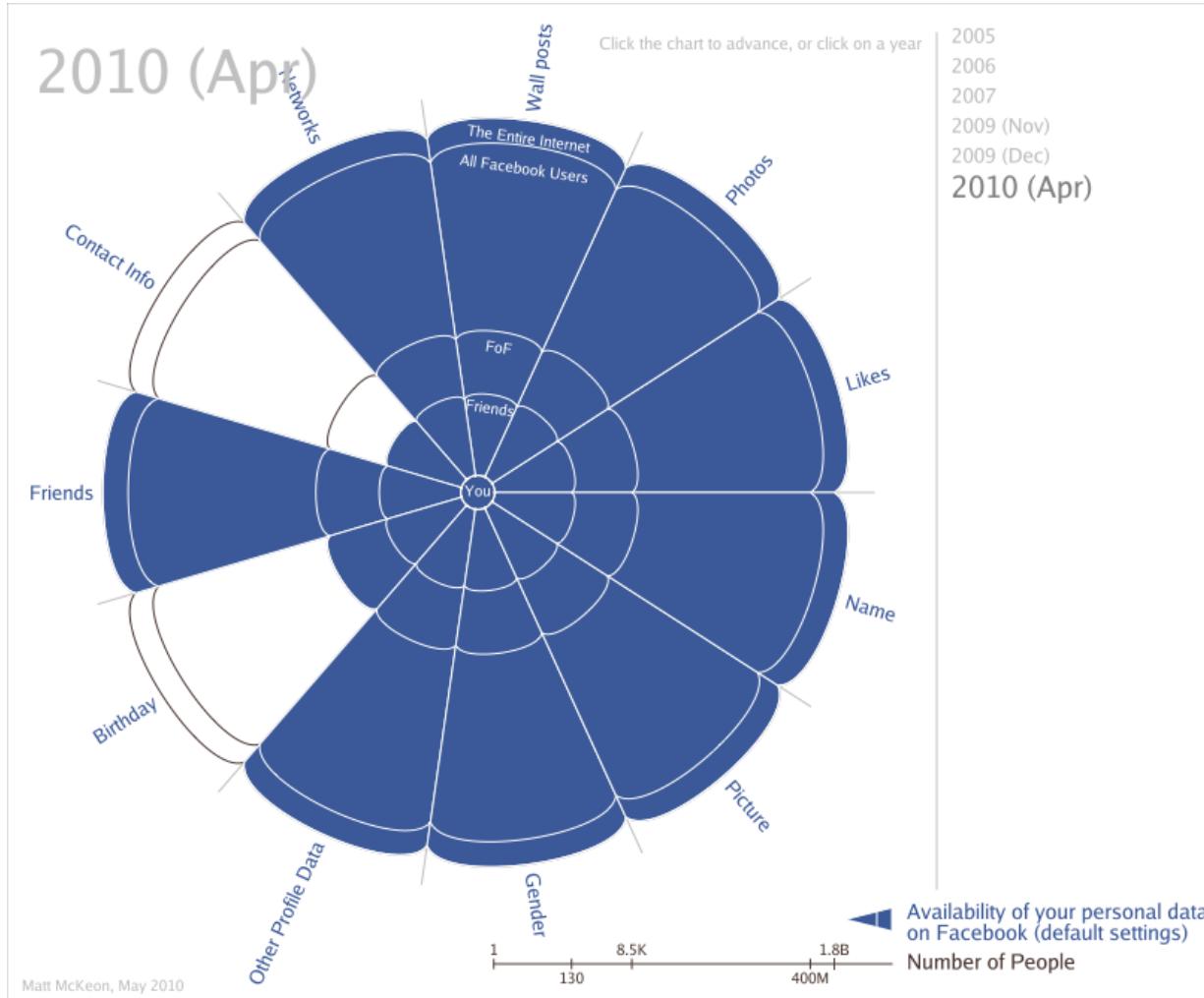
Facebooks Privacy-Evolution: 2009



Facebooks Privacy-Evolution: 2009



Facebook's Privacy-Evolution: 2010



Social Networking Sites as Targets



- Information on social networks is sensitive and attractive by nature:
 - E-mail addresses
 - Postal addresses
 - Educational background
 - Gender/Relationship status/partner
 - Spoken languages
- This information is valuable for attackers
 - Spear phishers
 - Spammers (implicit trust between users)
 - Malware authors (effective infection medium)

Cloning Attacks on Social Networks



- Target: get on the friend list of real users to get access to their personal information and their circle of trust
- Two Cloning Attacks
 - Clone the account of an existing user inside the same network and send friend requests to her contacts
 - Clone the victim profile into a different social network where she is not registered and contact her friends

1. INTRODUCTION

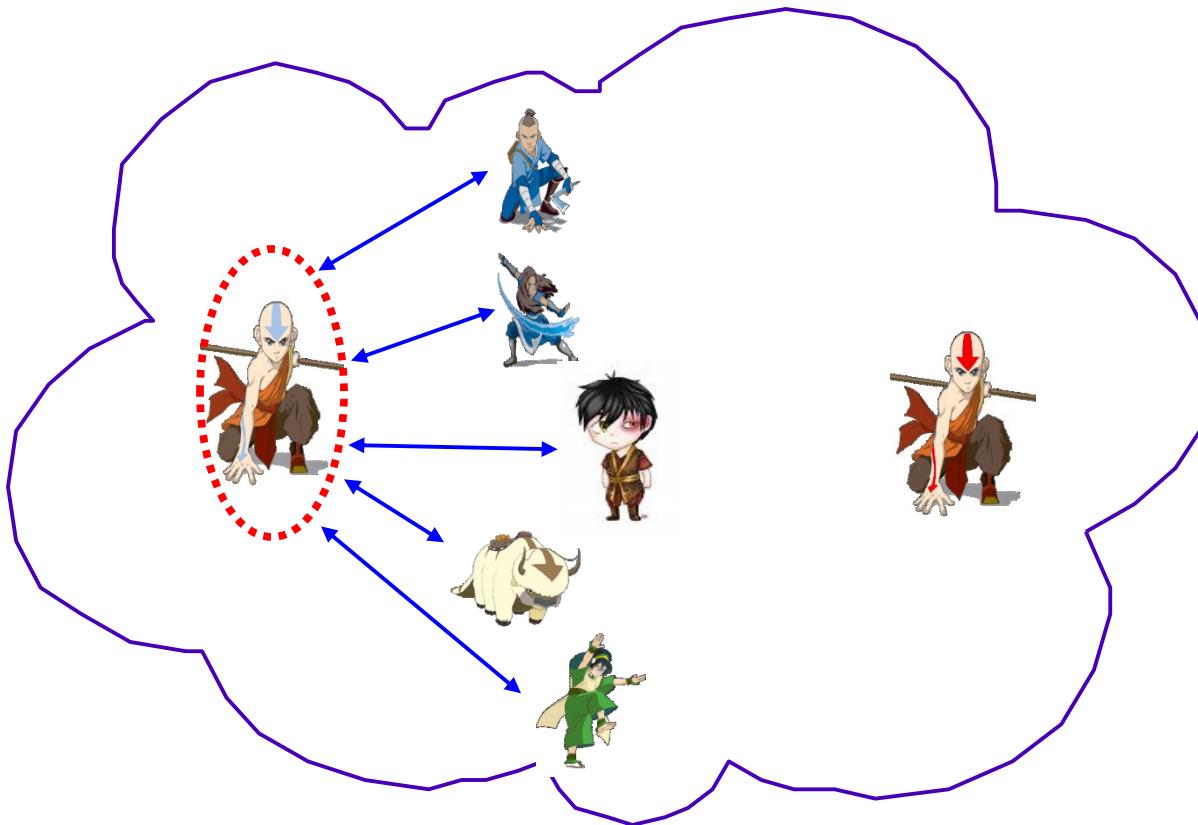
A social network is a social structure that is made up of nodes representing individuals or organizations. These nodes may be tied to each other by properties such as friendship, common values, visions, ideas, business relationships and general interests. Although the idea of social networks has been around for a long time (e.g., see [14]), social networking web sites and services are a relatively new phenomenon on the Internet. Business relationship-focused social networking sites such as XING [13] (previously known as OpenBC) and LinkedIn [6], as well as friendship-focused social networking sites such as Facebook [4], MySpace [8],

Aim of Attacks and Experiments

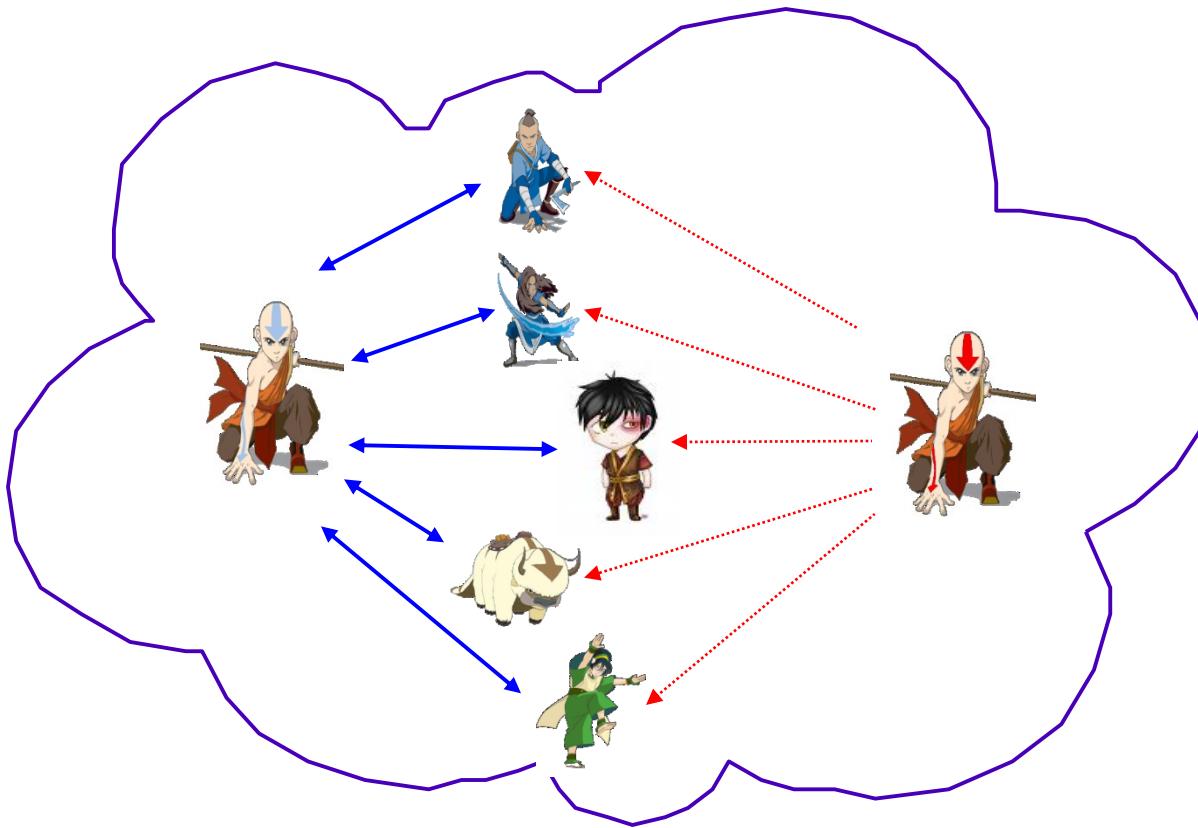


- Is it possible for an attacker to launch impersonation attacks on a large scale against a number of popular social networking sites?
 - Facebook (international)
 - XING (international)
 - LinkedIn (international)
 - MeinVZ (popular in Germany, Austria, Switzerland)
 - StudiVZ (popular in Germany, Austria, Switzerland)

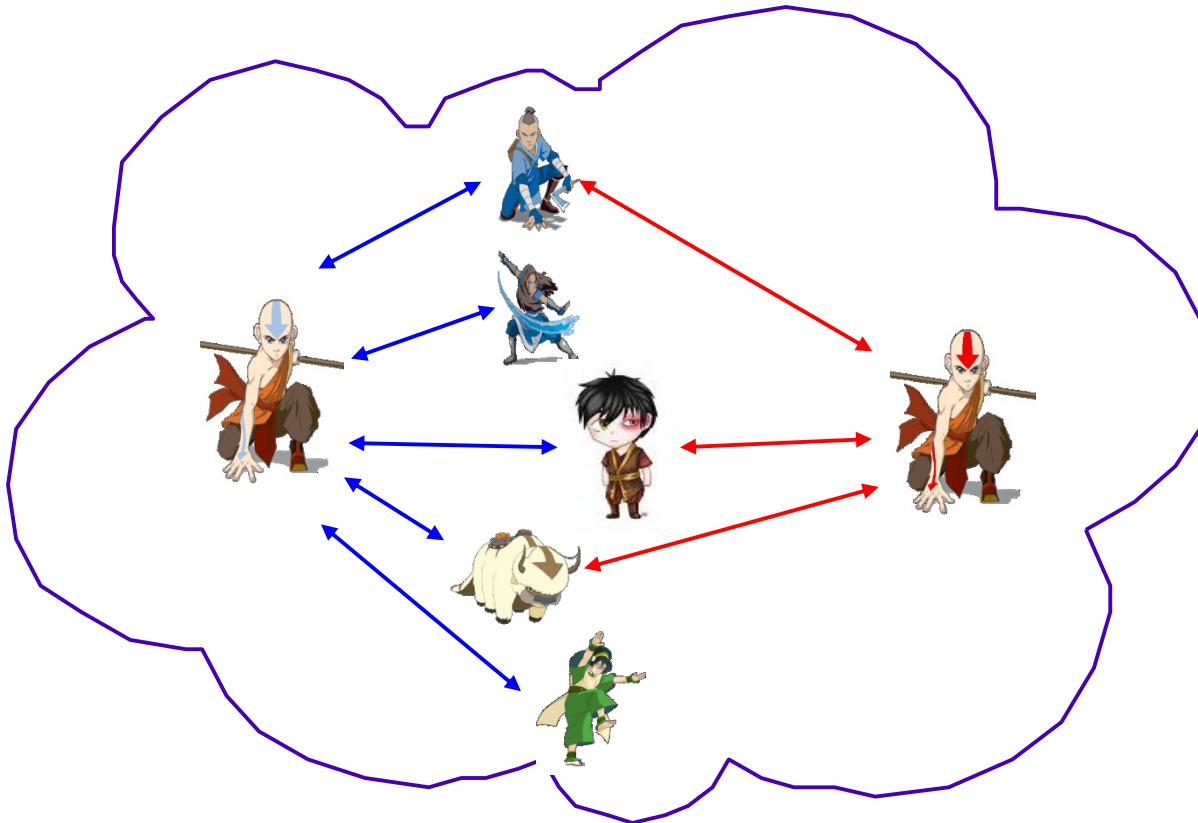
Automated Profile Cloning



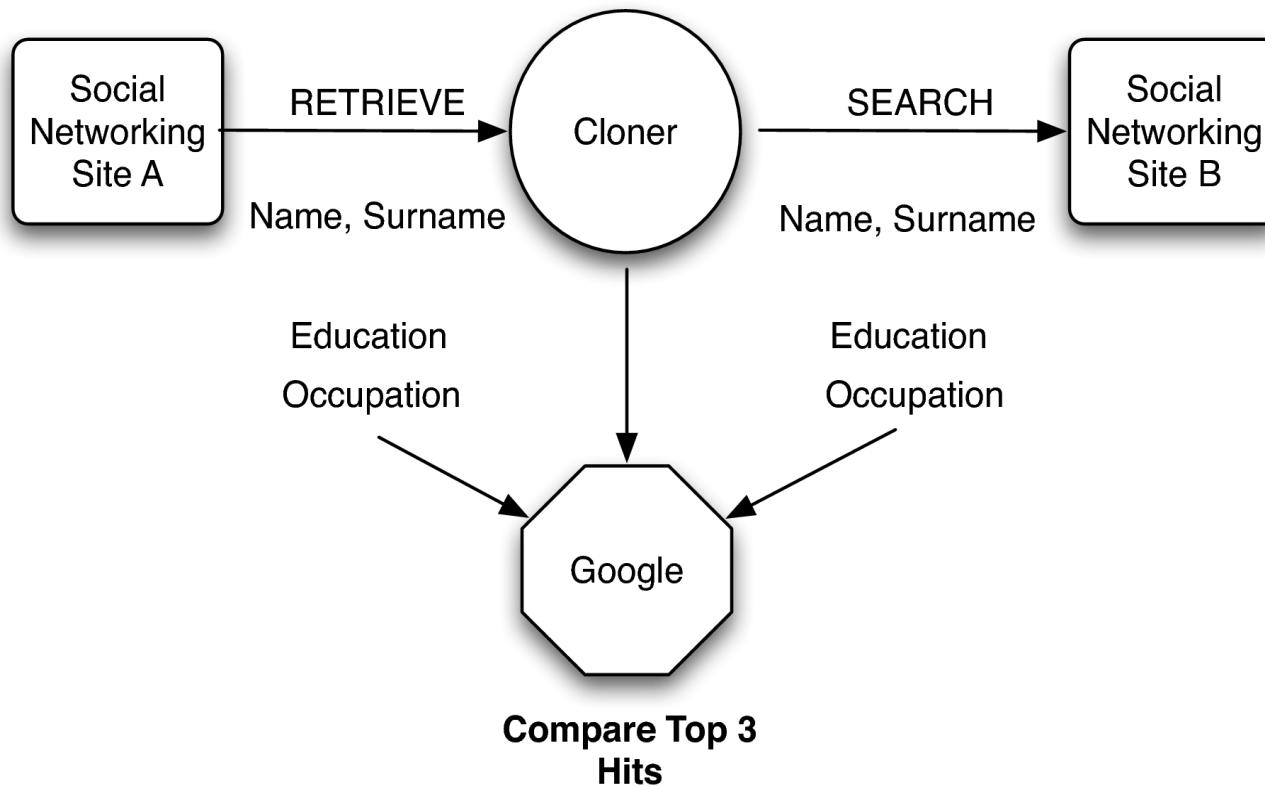
Automated Profile Cloning



Automated Profile Cloning



Cross-Site Profile Cloning



Obstacle: CAPTCHAs



- CAPTCHA: Completely Automated Public Turing test to tell Computers and Humans Apart
- CAPTCHAs are employed to prevent automated programs from accessing and abusing the services
- In order to automate the attacks, a number of CAPTCHA breaking techniques were developed
 - “Quick and dirty”, techniques are not perfect
 - The aim is to break the CAPTCHAs efficiently enough to make automated attacks against several social networking sites possible



- GD Library (PHP) CAPTCHAs
- CAPTCHAs always contain 5 letters
- Each letter is written in
 - Different font
 - Different background and foreground color
- Often tilted, scaled or blurred
- A simple grid-base noise is added to the image
- Quick script* with success rate of 88.7%



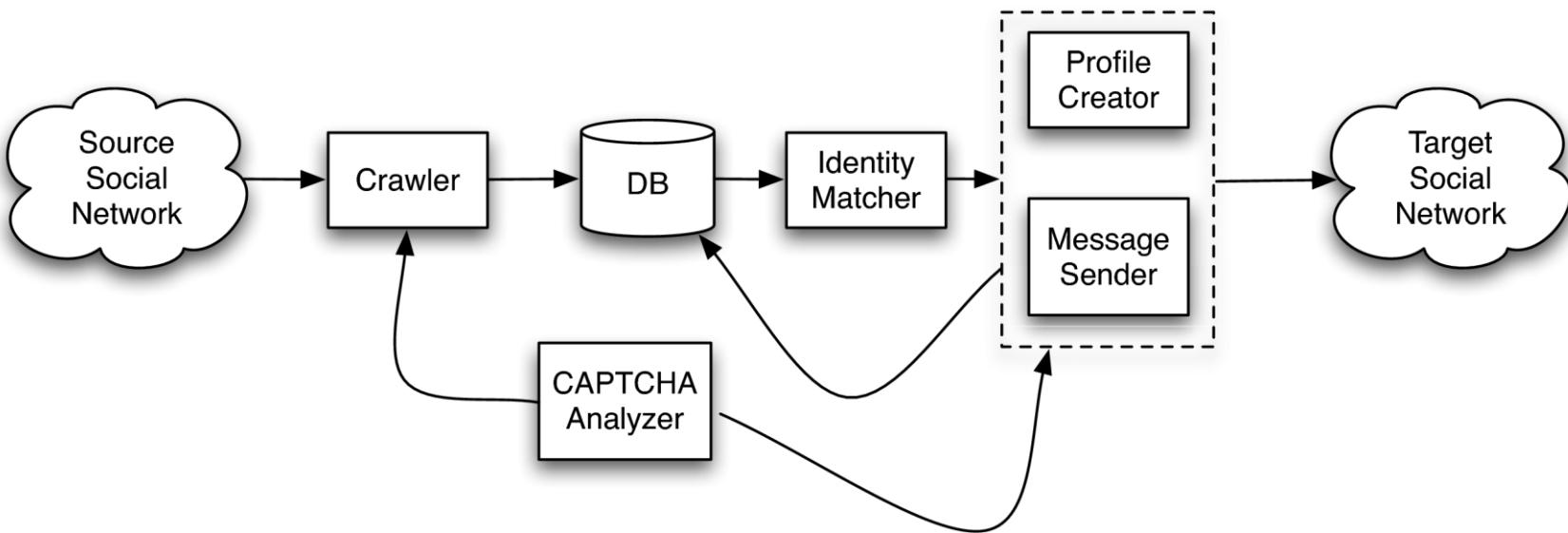
**Cracking the CAPTCHAs was done with serious amounts of help from Michael Roßberg/TU-Ilmenau*



- Adopts ReCAPTCHA technique
 - Asks words that are encountered while digitizing books that cannot be correctly recognized by the OCR program
 - By solving the CAPTCHAs, the user contributes to the effort to increase the accuracy of the text of the digitized book
- ReCAPTCHA asks meaningful words. Therefore, after solution is found, the word is sought in a dictionary
 - Result additionally submitted to Google as check
- Script with success rate of 7%
- Might seem small, but...
 - If every bot is capable of solving 7 CAPTCHAs per day, a botnet that consists of 10.000 can send 70.000 friend requests per day
 - Attack against Microsoft Live Hotmail had similar success rate



Prototype Implementation: iCloner



Real-World Experiments



- Is it feasible to perform cloning attacks in the real-world?
- Questions:
 - Can an attacker launch large-scale attacks?
 - How willing are users to accept friendship requests from forged profiles of people who are already in their friendship lists?
 - Is it possible to efficiently find two identical accounts in two different social networks?

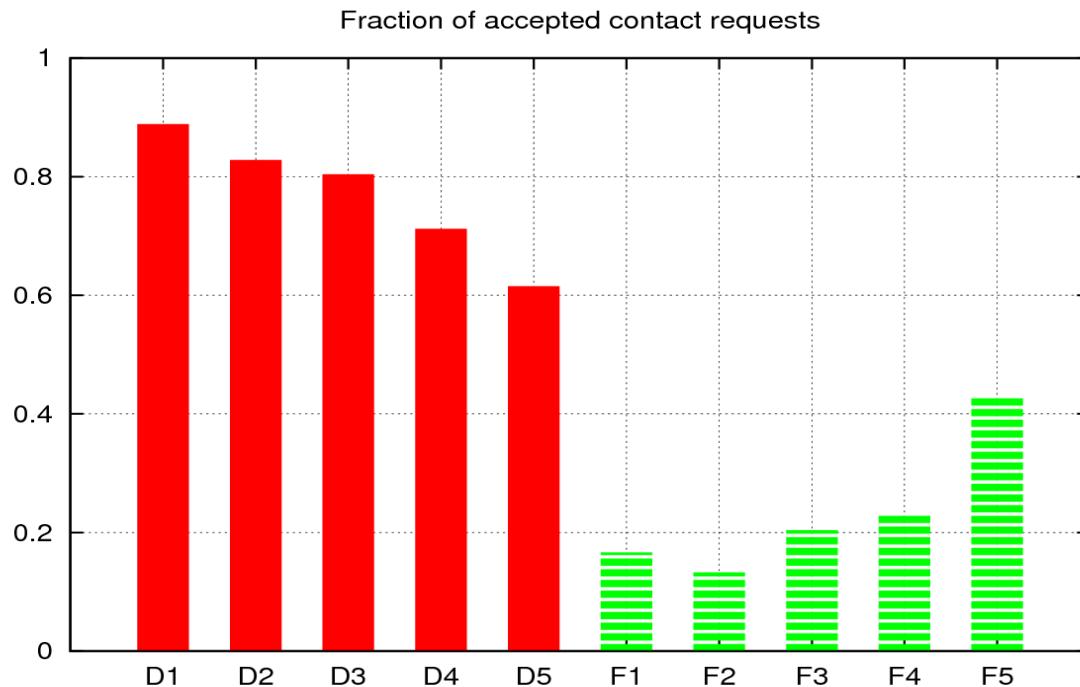


- StudiVZ and MeinVZ
 - Displays CAPTCHA if large number of requests come from one account
 - To collect as much information as possible, without being noticed, 16 accounts were created, and separately used for crawling
 - Collected 5M profiles with contact information, and 1.2M complete user profiles
- XING
 - Does not display CAPTCHA, but disables the account if the account requests around 2000 pages consecutively
 - 118,000 accounts were crawled

Profile Cloning Experiments



- Attack: duplicate the profiles of five users (D1,...,D5) and create fictitious profiles (F1,...,F5 as control group)



Profile Cloning Experiments



- Do the users really trust their friends in their friend list?
- Would they click the link seen in the message below?

Hey, I put some more pictures online. Check them here!:

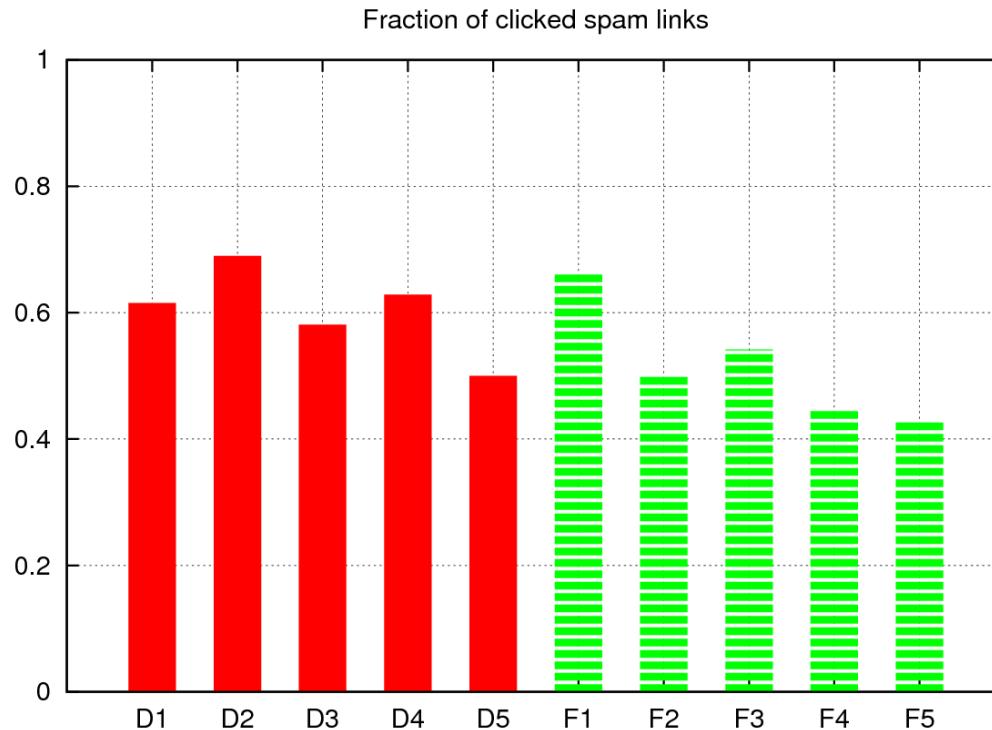
`http://193.55.112.123/userspace/pix?user=<account>&guest=<contact>&cred=3252kj5kj25kj325hk}`

Ciao, <account first-name>

Profile Cloning Experiments



- Click through rate for messages from duplicate / fictitious profiles



Cross-Site Profile Cloning Experiments



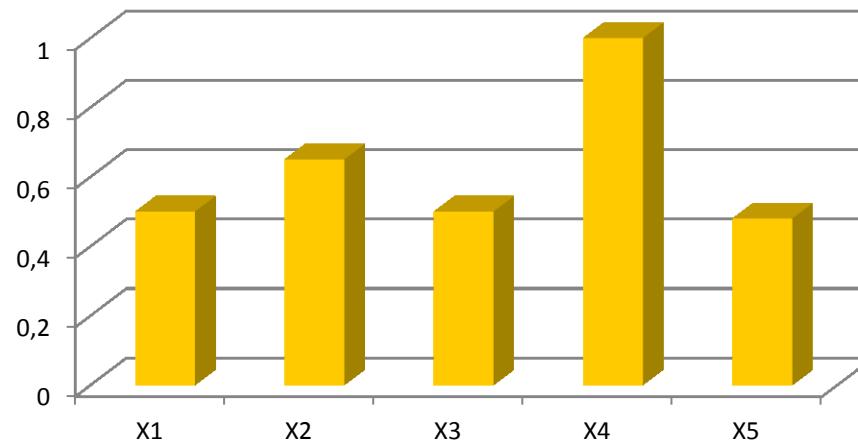
- Cloning profiles that exist on XING, but not on LinkedIn
- The success of the cross-site profile cloning depends on the number of users that have a profile in both of the networks
- From around 30.000 crawled profiles in XING, 3.700 were also registered in LinkedIn

Cross-Site Profile Cloning Experiments



- Clone 5 users from XING to LinkedIn
- iCloner identified 78 out of 443 XING friend contacts that were also registered in LinkedIn
- Fraction that has actually accepted the contact requests:

Fraction of accepted contact requests



How Can Social Network Security be Improved?

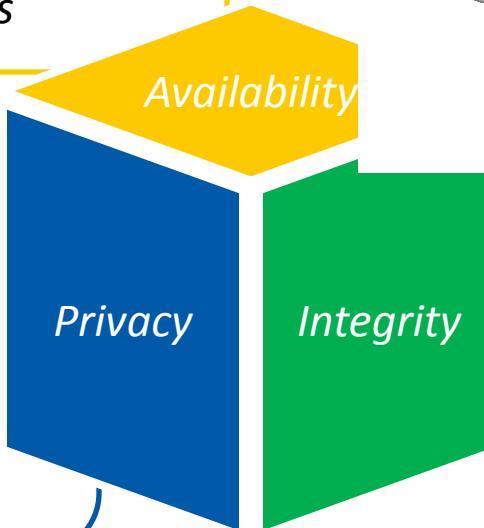


- Make the users aware of the risk
- Make privacy controls easier to use (and show consequences!)
- Provide a stronger way to authenticate users
- Increase defense (rate ctrl against massive automatic harvesting):
 - Make CAPTCHAs more difficult to break
 - Reduce response rates
 - Limit number of accessible profiles
- Apply anomaly detection techniques to detect:
 - Crawling
 - Massive amount of friend requests or messages
 - Profiles with identical/similar information
 - Attempts to automatically solve CAPTCHAs
- ...decentralize, but that comes a bit later ;-)

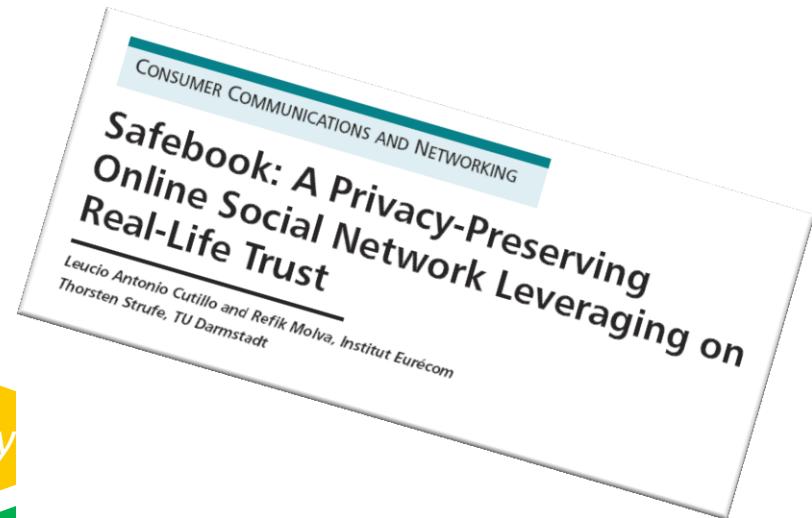
Security Objectives in OSNs



1. *Of data and services*
2. *Robustness against censorship*
3. *Prevent from seizure or hijacking of identities*

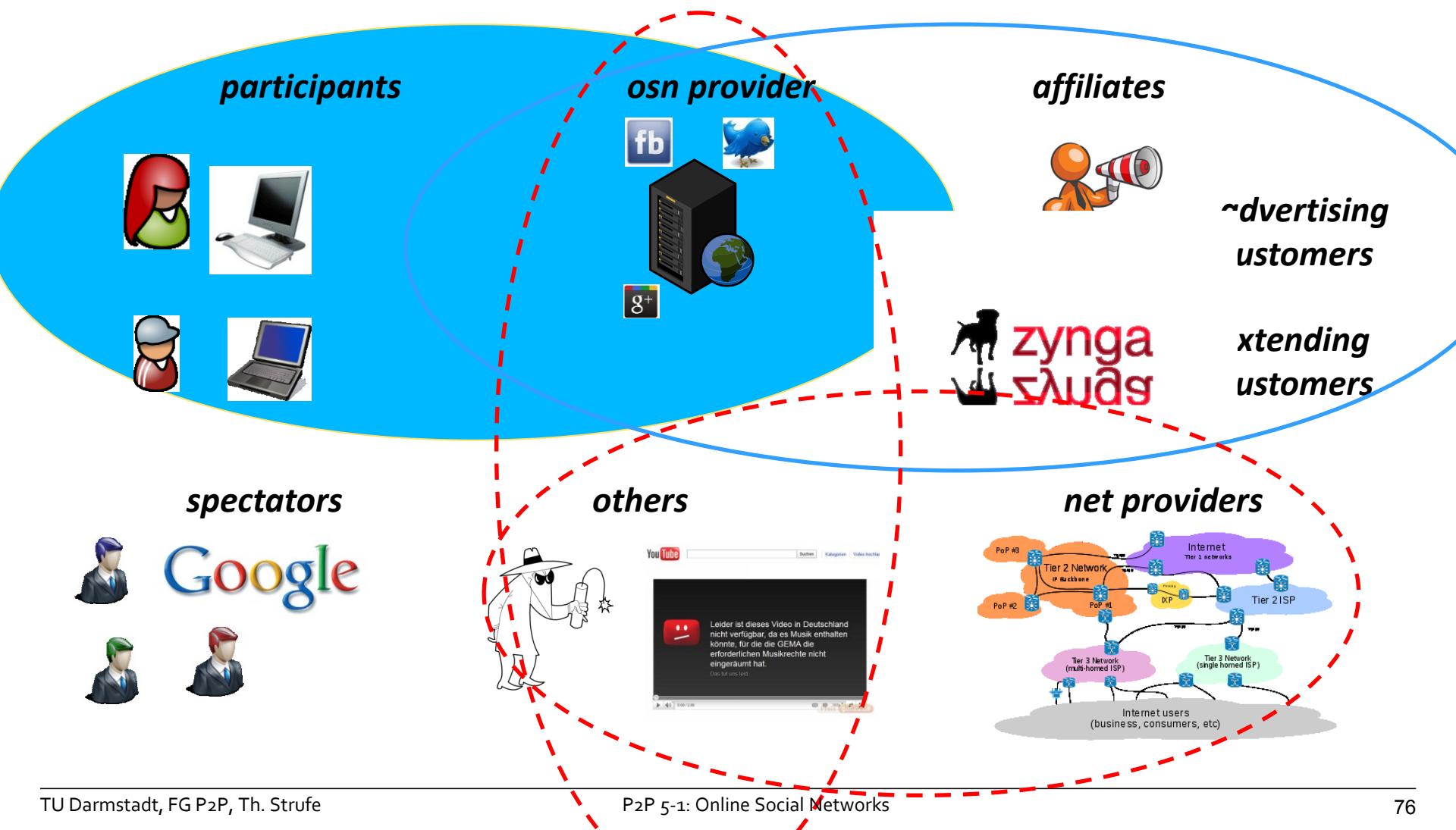


1. *Only trusted parties can access:*
 - *Personal Information*
 - *Communication*
2. *Possibility to hide any data about any user*



1. *Protect user data from unauthorized modification and tampering*
2. *Ensure the link between OSN accounts and people*

Stakeholders in OSN



Types of Information



■ Explicit

- created content (+self descriptive)
- annotations/comments
- preferences/structural interaction (contacts, +1, etc)



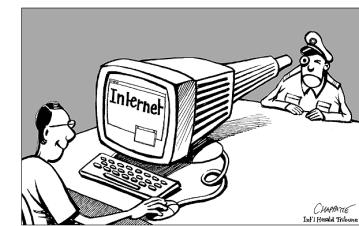
■ External

- interest/preferences (clickstreams through adnetworks, fb-connect)



■ Implicit

- inferable from environment (homophily)
- observable
 - **session artifacts** (time of actions), **interest** (retrieved profiles; membership in groups/ participation in discussions), **influence** (users)
 - **clickstreams**, ad preferences, exact sessions, **communication** (end points, type, intensity, frequency, extent), **location** (IP; shared; gps coordinates), **udid**



■ Processed/derived

- preference models
- **image recognition models**
- aggregates (characteristic properties/ descriptors of user groups)



- **Storage:**

- clear text at the provider

- **Authorization:**

- audience selection by owner
through SNP functions (per
object, per subject/group)

- **Authentication:**

- login to account at SNP

- **Access Control:**

- through SNP upon content
request/presentation to
authorized subjects only

- **User/Identity Mgmt:**

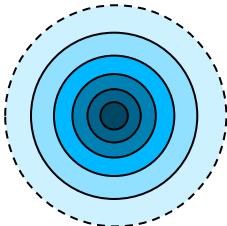
- Account creation (email),
SNP functions for contacting
user, group/list management

De-facto Access



Explicit

- user
- Grantable
 - specific contact(s)
 - “friends”
 - friends of friends
 - service subscribers
 - The public



Implicit

- SNP



Everything the installing user can see

- Affiliates
 - Extended
 - Advertising



*Not much (aggregates)
Unless they pay really well*

- ISP



*Everything their subscribers see/write
(until Nov 21st '12)*

Ignoring downstream abuse...

The OSN Security Model



*Assuming that users can
and do apply authorization
and user mgmt*

participants



affiliates



*advertising
customers*



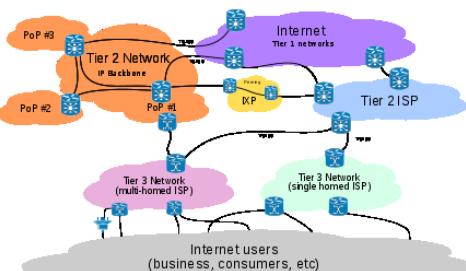
*extending
customers*

sf

G



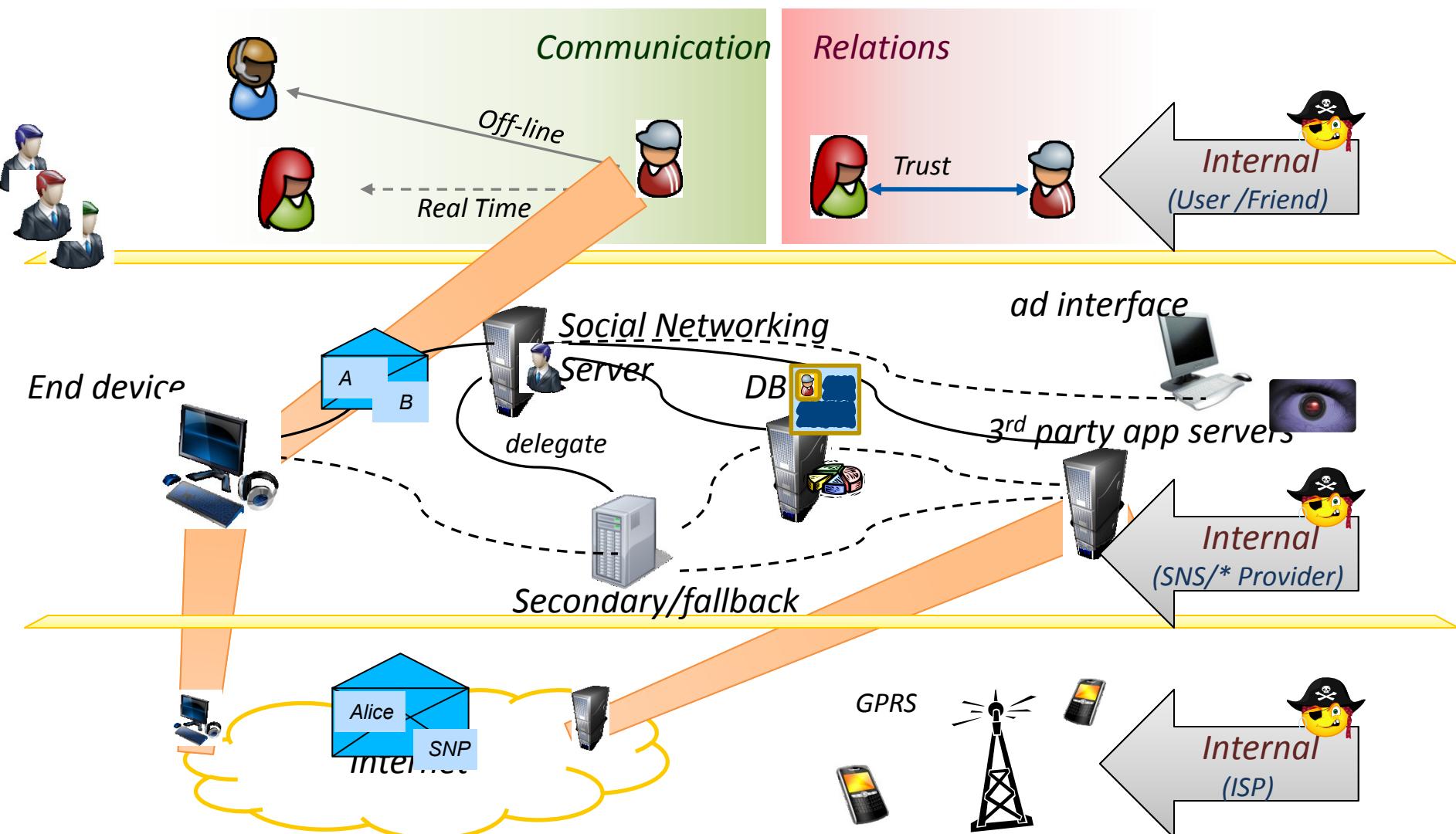
net providers



others

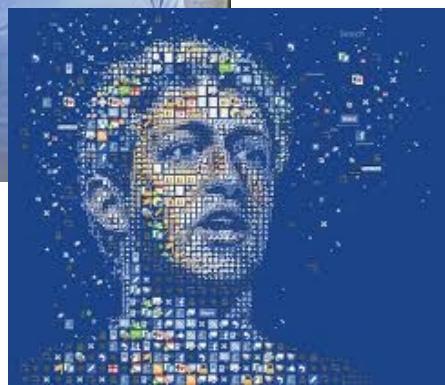
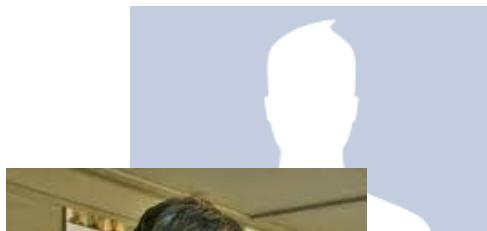


The Model Revisited: Potential Adversaries



Privacy Concerns

Which disclosure are users concerned about?



Because They Don't Realize...



- Extent of collected information (data)
- Increasing ubiquity of monitoring
 - Facebook.com
 - -> places/foursquare
 - -> google maps
 - -> browser location sharing

Loss of benefits

*Illegal access,
stalking*

*Cooperation with
Intel services*

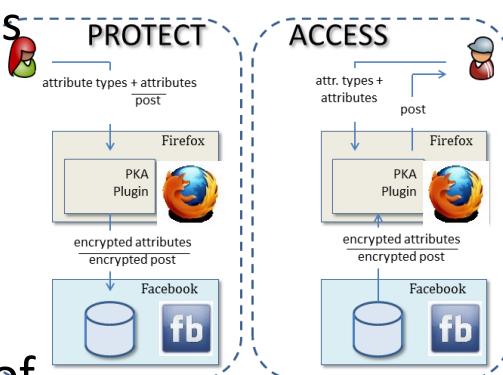
Price discrimination





Crawler Defense [1]

- Access trade-off (oversharing):
 - *Crawlers/Adversaries: deny!*
 - *Friends: grant* w/ low overhead
- *Partial-knowledge based Access*
- Use attribute sets as credentials
- Shamir shares stored in profile
- Encrypt posts



- Security proof
- Analysis on real world data:
 - Security comparable to passwords
- Firefox extension for download

FB Privacy Settings [2,3,4]

- Default settings too weak and **complicated** to grasp
- Comprehensible controls:
 - Ease of use, direct feedback



- Implementation: Firefox-Extension
- Comprehensive user studies (TK)
- Download: „Privacy Ampel“ (50k)

[1] WWW '13 [3] HCI WS '11
 [2] SocInfo '12 [4] WWW '12

Immediate Defenses – Crypto Schemes



PMS [1,2]

- Plain content stored at provider
- Comprehensive AC scheme using crypto
- Simple group management
- Encrypt on attribute level
- Shared key vs. Broadcast Enc.
- Proof of confidentiality, unlinkability/anonymity

BroadMask [3]

- Implementation of PMS
- Chrome/FF extension
 - PBC, Crypto++, GnuPG
 - Picasa plugin
 - Bitmap wrapping

The diagram illustrates the process of decrypting and unwrapping content from a BroadMask Facebook post. It shows a Facebook wall post with a message, images, and an audio file. A callout box points to the message with the text "Decryption, BMP fetching & unwrapping".

Original Facebook Wall

Bob BroadMask
==== BEGIN BM DATA ====
eyJnaWQ0IjZkNjYmVuY2htYXJrX3NlbnRlcilsmRhdGEiOnsidGV4dfF9jZXNzYWdlIjotTUNyBeIEISXhnQ0F4TWpns01dQUrFZDE3U0d853h0akZwdjIzMGpmUmRPNWhmVzueJkcdhwY1NhcmfpFdIkSta2IMK1l5VGtw2zczMEVxc0z0R1Z4YVkbhqa1dSbUIMVhNIektGVGxhWDVlMU0JryeDhHWURSW4vcndN
by9xbvdWY2k2WUtUnftU1lvaik4yaFlheTY4RHluSMvc09ROUfyc1ltOXNUWU9mWnVLZC9p3RHNnpLREZZWWhOY2
VaZDlvt2dM3g2V1hERy9MaUvStOo4czJP23FPY2l4c0xmW...
[Continue Reading](#)

Wall with Injected Content

This is an exemplary private message shared with a subset of this user's friends. It consists of this text message, some images, and an audio file, which are appended at the bottom of this post.

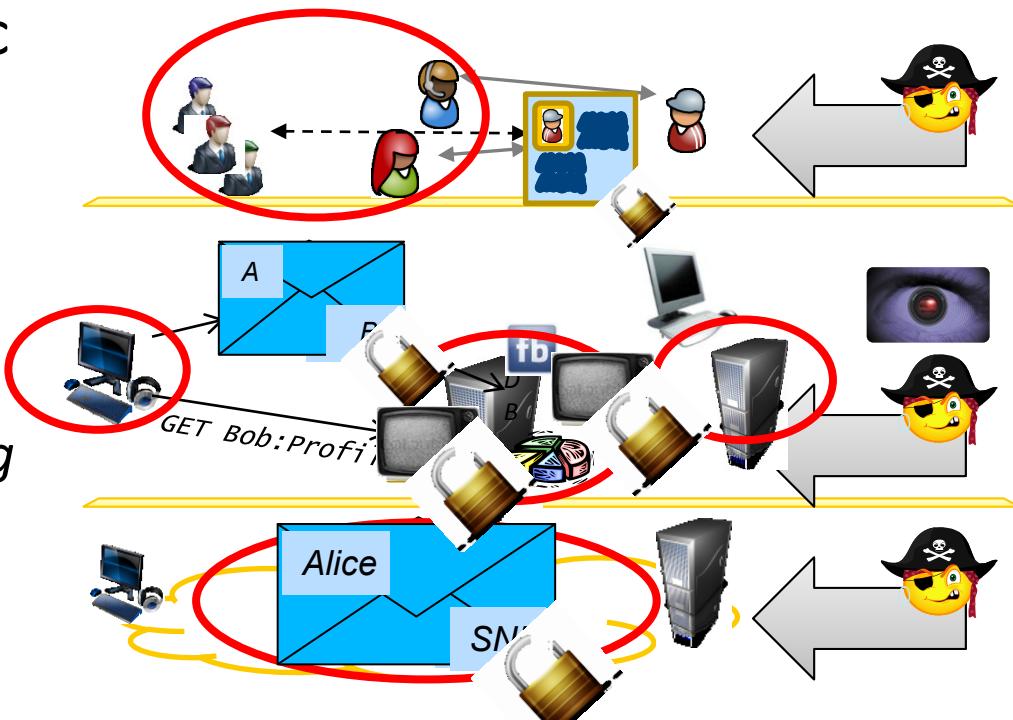
[Click to view content \(Mimetype: data:audio/mp3\).](#)

Decryption, BMP fetching & unwrapping

Potential Solution Classes



- Trust „everybody“
- Suspect Network
 - Transport Layer Security
- Suspect subscribers, public
 - Trust provider (& affiliates)
 - Apply OSN Access Control
- Suspect affiliates/browser
 - Access abuse, *unsolicited msg*
 - *Web security, Sandboxing..*



What if: you don't want to be observed?

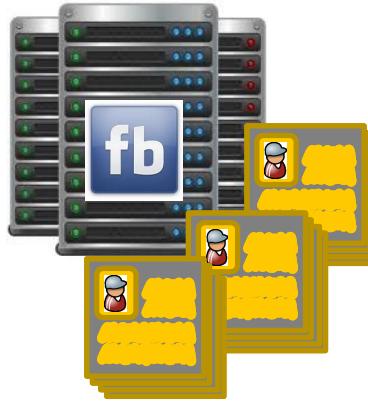


- Provider still has the power to
 - Intercept and drop messages
 - Observe who is active
 - What else the profiles are interested in (cookies)
 - Who is communicating with whom
 - OSN identities
 - Network layer addresses (actual individuals with post addresses)
 - ...

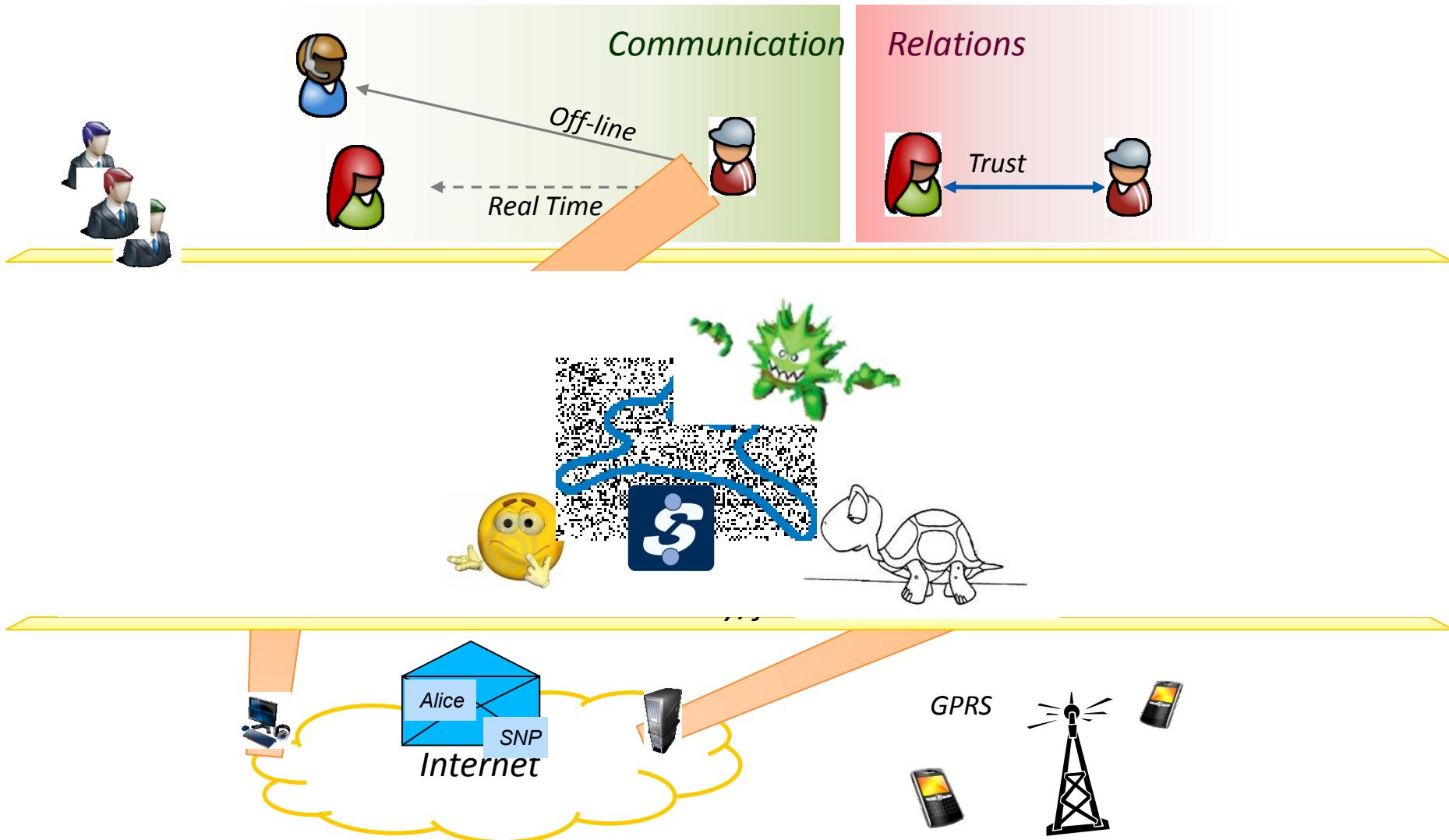
Potential Adversaries - 2 (and solution classes)



- Suspect provider and affiliates
 - Threat: anonymity, behavior
 - Decentralization
 - Threat: o
 - Darknets



The Model Revisited: Potential Adversaries



Solution Classes Summarized



- Web Security
- Transport Layer Security (https)



- OSN Access Control
 - (plus Usability)
- Web Security 2
 - (Sandboxing)

- *Crypto Schemes*
 - (with different properties)



- *Decentralization*



- *Darknets*

- *Awareness!*

- *Then again...*



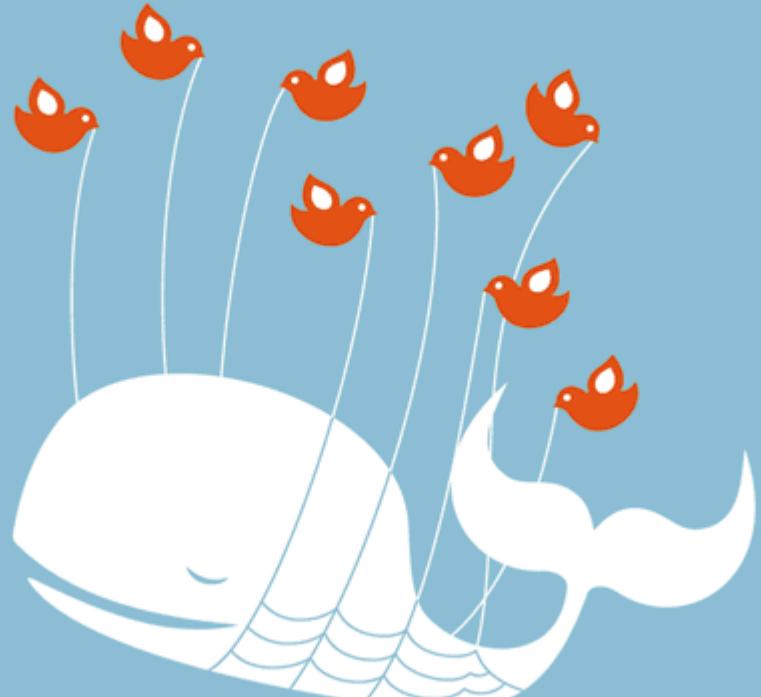
Quest to Decentralize



[Home ›](#)

Twitter is over capacity.

Too many tweets! Please wait a moment and try again.



© 2010 Twitter [About Us](#) [Contact](#) [Blog](#) [Status](#) [API](#) [Help](#) [Jobs](#) [TOS](#) [Privacy](#)



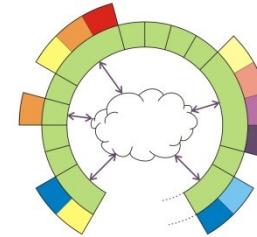
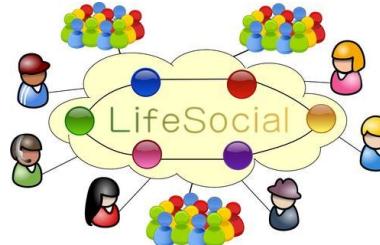
- FOSS
 - BuddyPress, CrabGrass, Cobs, DaisyChain, Diki, Elgg, FETHR, GNUNet, Gossple, Jappix, Lorea, Mycella, Movim, PeerScape, Pinax, StatusNet
- Commercial Approaches
 - diaspora, wuala, LifeSocial
- Academia
 - Friend-of-a-Friend, FriendStore, HelloWorld, LifeSocial, LotusNet (Likir), PeerSon, Safebook, SocialCircle, Tribler, Vis-a-Vis
- Focus on systems that
 - implement social networking and publication functions
 - Provide running software or a comprehensive protocol / system description
 - Actually are decentralized

Selected Systems and Proposals



- diaspora
- Friend-of-a-Friend
- LifeSocial
- LotusNet / Likir
- PeerSon
- Safebook
- Vis-A-Vis

diaspora*



*PeerS*ocial \wedge \mathcal{N}



Classifying Decentralized OSN



- Type of storage / service provision
 - Infrastructure-based
 - Dedicated Web-Servers
 - Deployed in the cloud
 - Peer-to-Peer-based
 - Hybrid
- Granularity of service provision
 - Replicating whole service (profile and provision)
 - Distributed storage of attributes
- Level of integration
 - Stand alone system
 - Extension of existing systems
- Resource sharing incentives
 - None
 - Social cooperation
 - Payed premium services



<i>Approach</i>	<i>Type of Service</i>	<i>Service Granularity</i>	<i>Integration</i>	<i>Incentives</i>
<i>diaspora</i>	web-based	complete	external	Premium
<i>FoaF</i>	web-based	complete	external	
<i>vis-à-vis</i>	Cloud	complete	external	
<i>LotusNet</i>	P2P	split	stand alone	
<i>PeerSon</i>	P2P	split	external	
<i>Safebook</i>	P2P	complete	stand alone	Social Coop
<i>LifeSocial</i>	hybrid	split	stand alone	Premium

Challenges for P2P OSN



- Let's use P2P to increase privacy in OSN!
- *How could we proceed?*
- *Which challenges are we about to encounter?*

- Performance
 - Properties: P2P → filesharing! OSN: small, short lived objects
 - Extreme heterogeneity (friends, uploaded content, resources)
 - User behaviour / sessions (diurnal patterns, session- and inter session times)
 - Incentives
 - Mobile users vs. p2p / recursive routing, etc...

- Security
 - How can we „guarantee“ availability?
 - How do you search for somebody?
 - Key management!?
 - Identification services (we don't want fake accounts!)

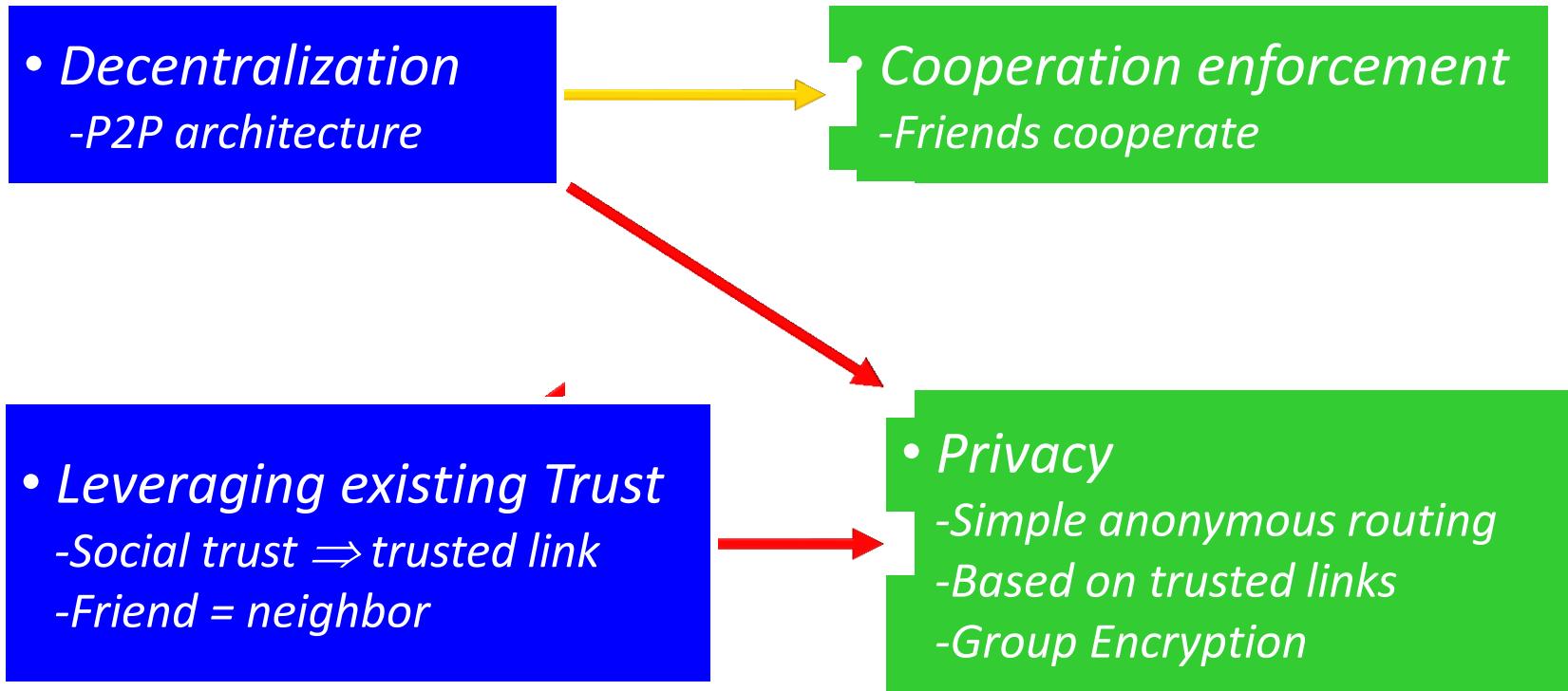


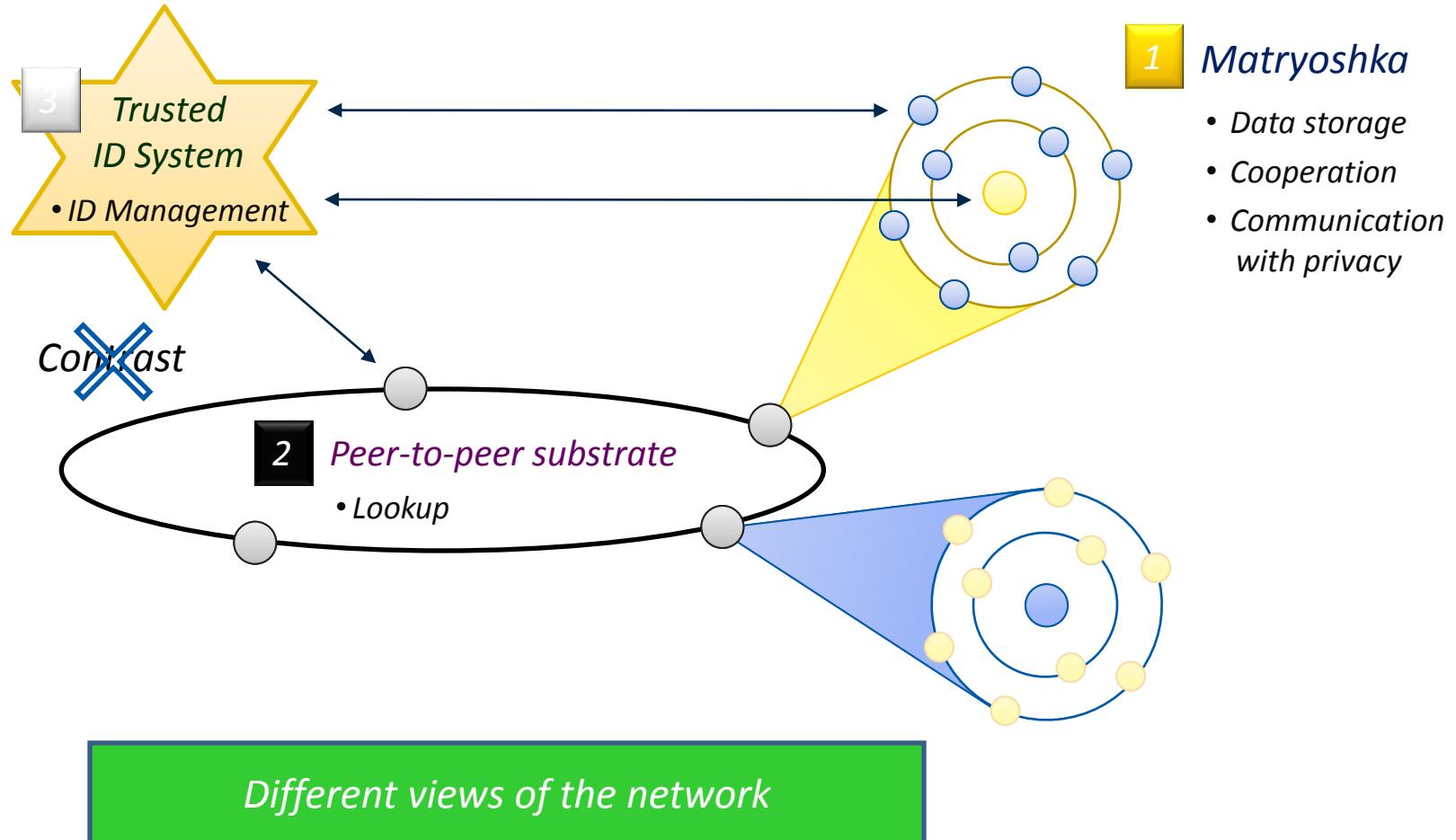
Safebook : Privacy-Preserving Online Social Networking

<http://www.safebook.us>

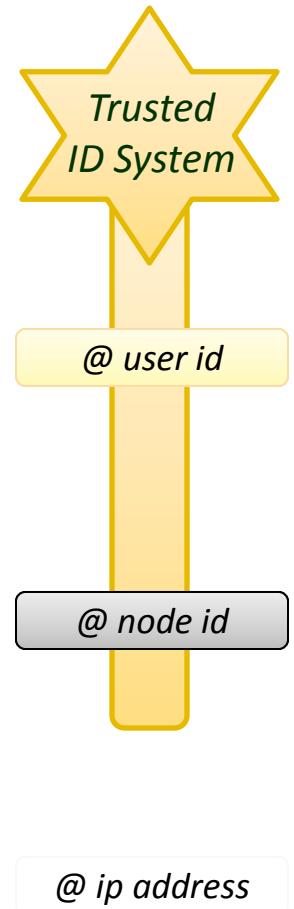
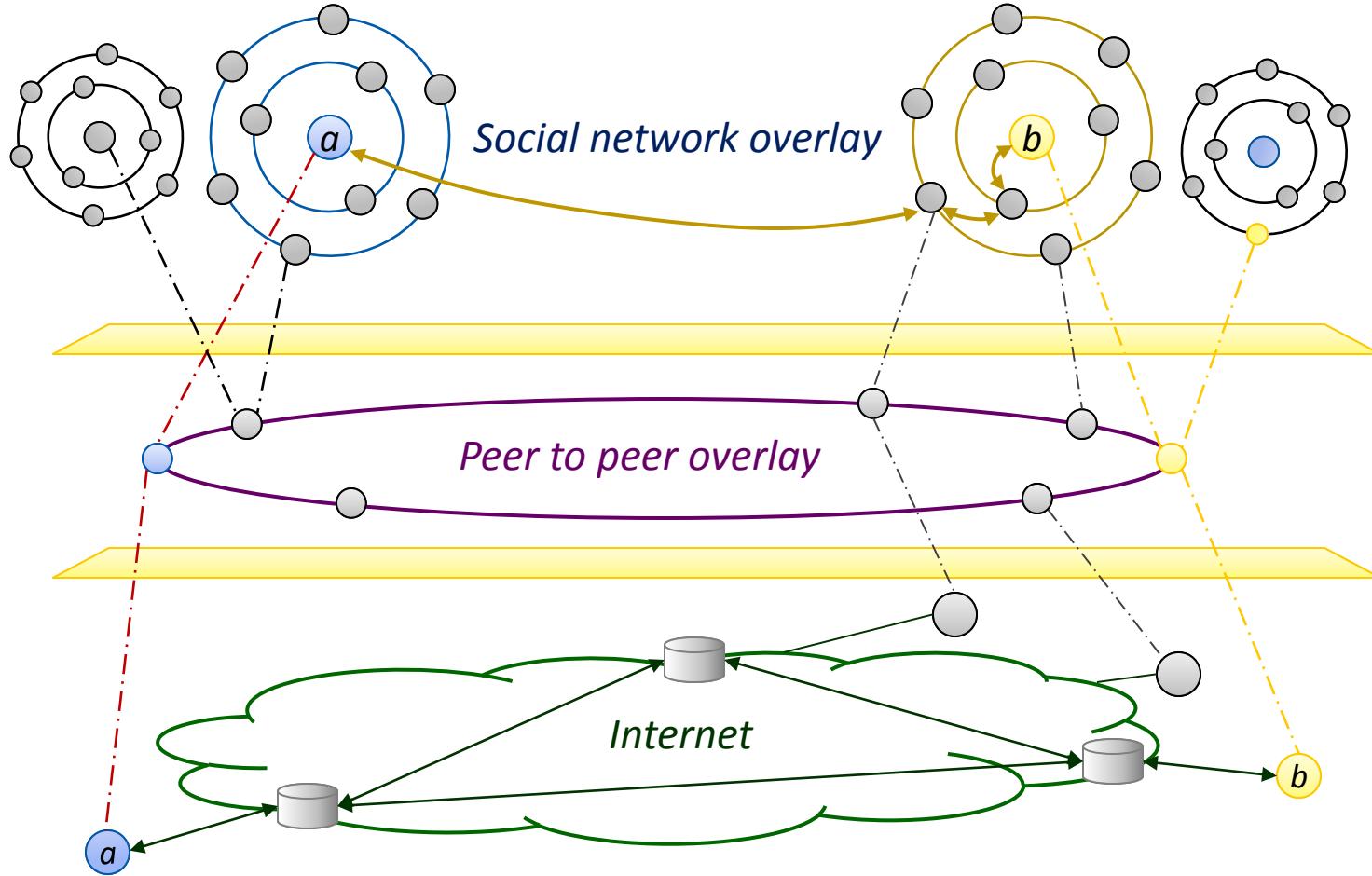
L. A. Cutillo, R. Molva, M. Önen, T. Strufe

Design Principles

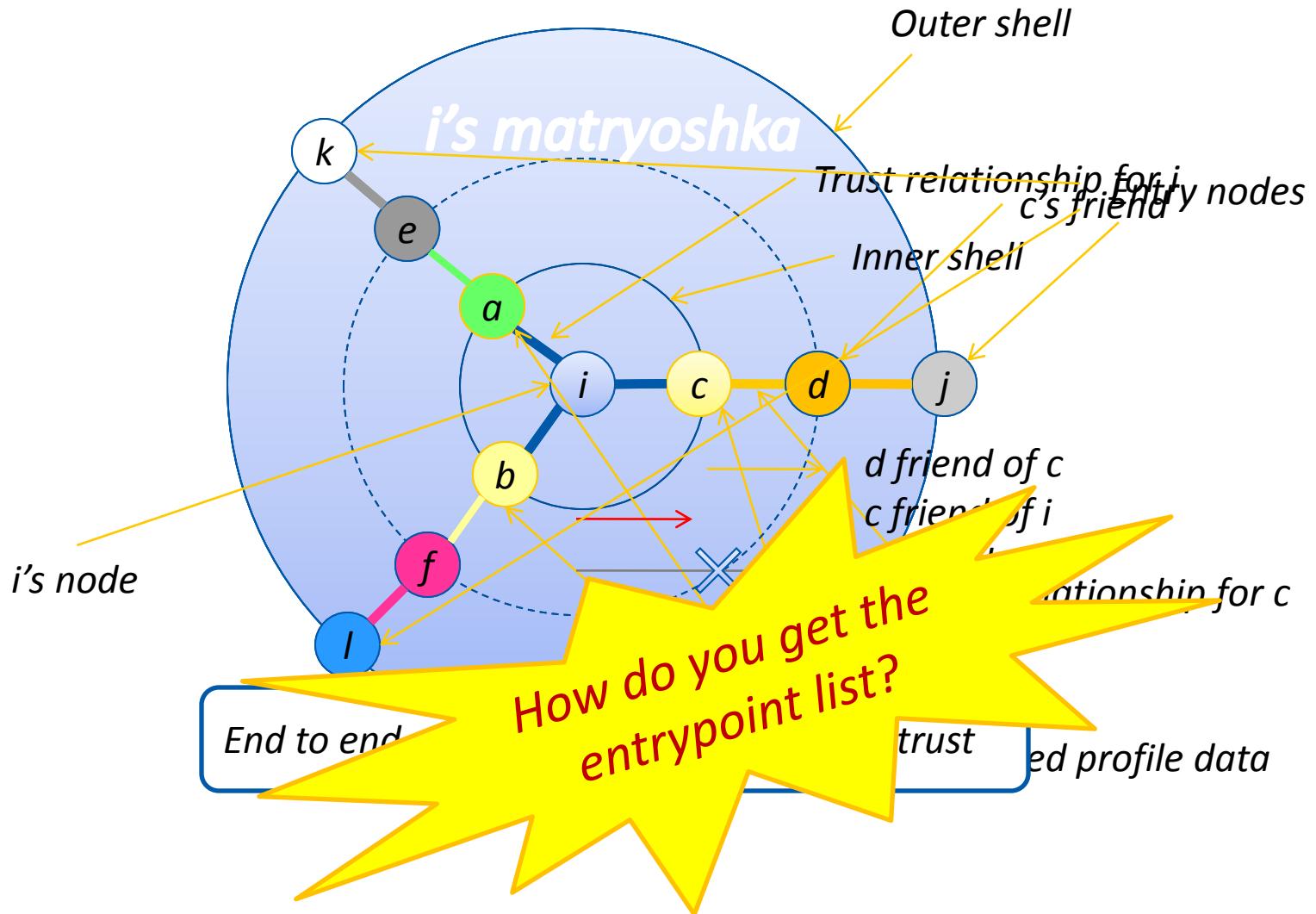




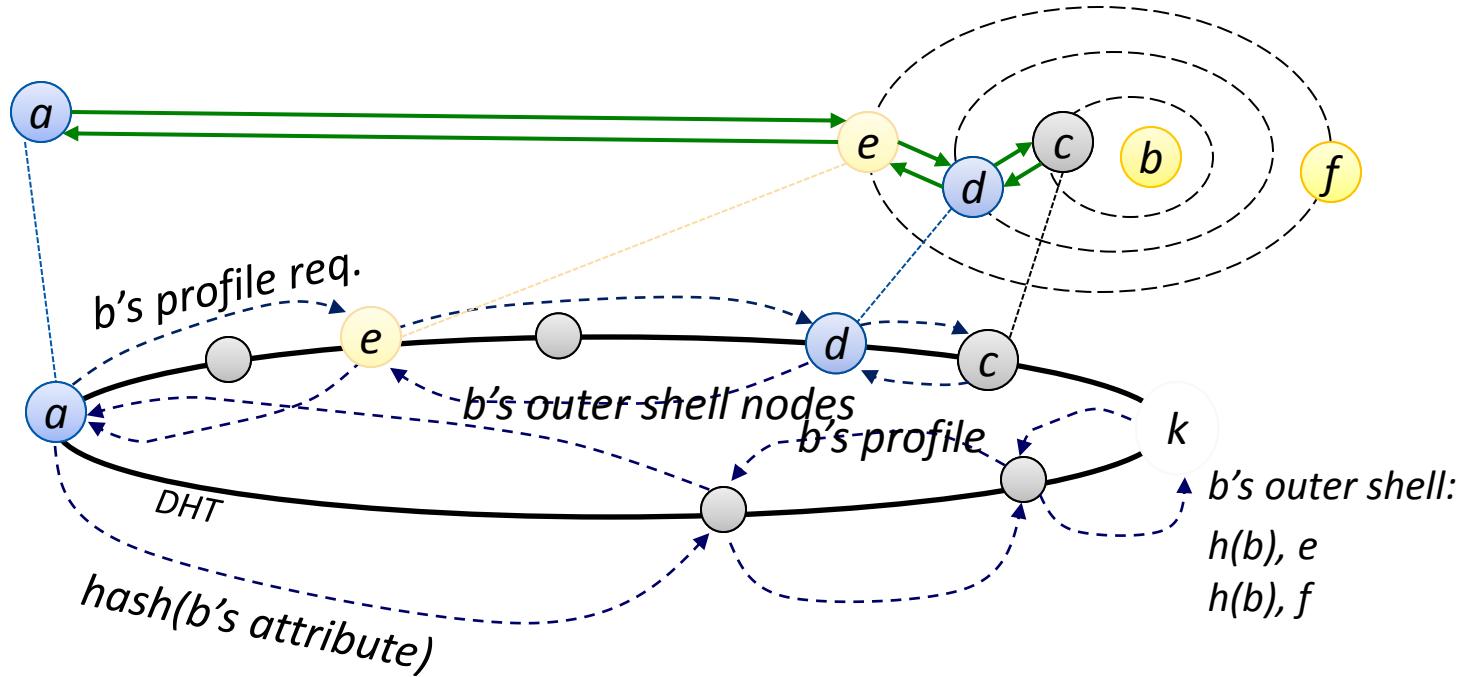
Network view



User i's Matryoshka



Finding it, using P2P: a looks for b



lookup

- a looks for b 's entry nodes
- k provides b 's outer shell nodes

data request

- a sends profile data request to a b 's entry node

Data reply

- One of b 's inner shell nodes answers

Evaluation of the scheme (1)



Privacy



Friendship relations hidden through Matryoshkas



Untraceability through pseudonymity and anonymous routing

Cloning prevention



ID management

Dos prevention



Access control



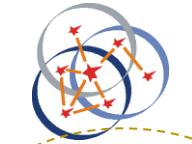
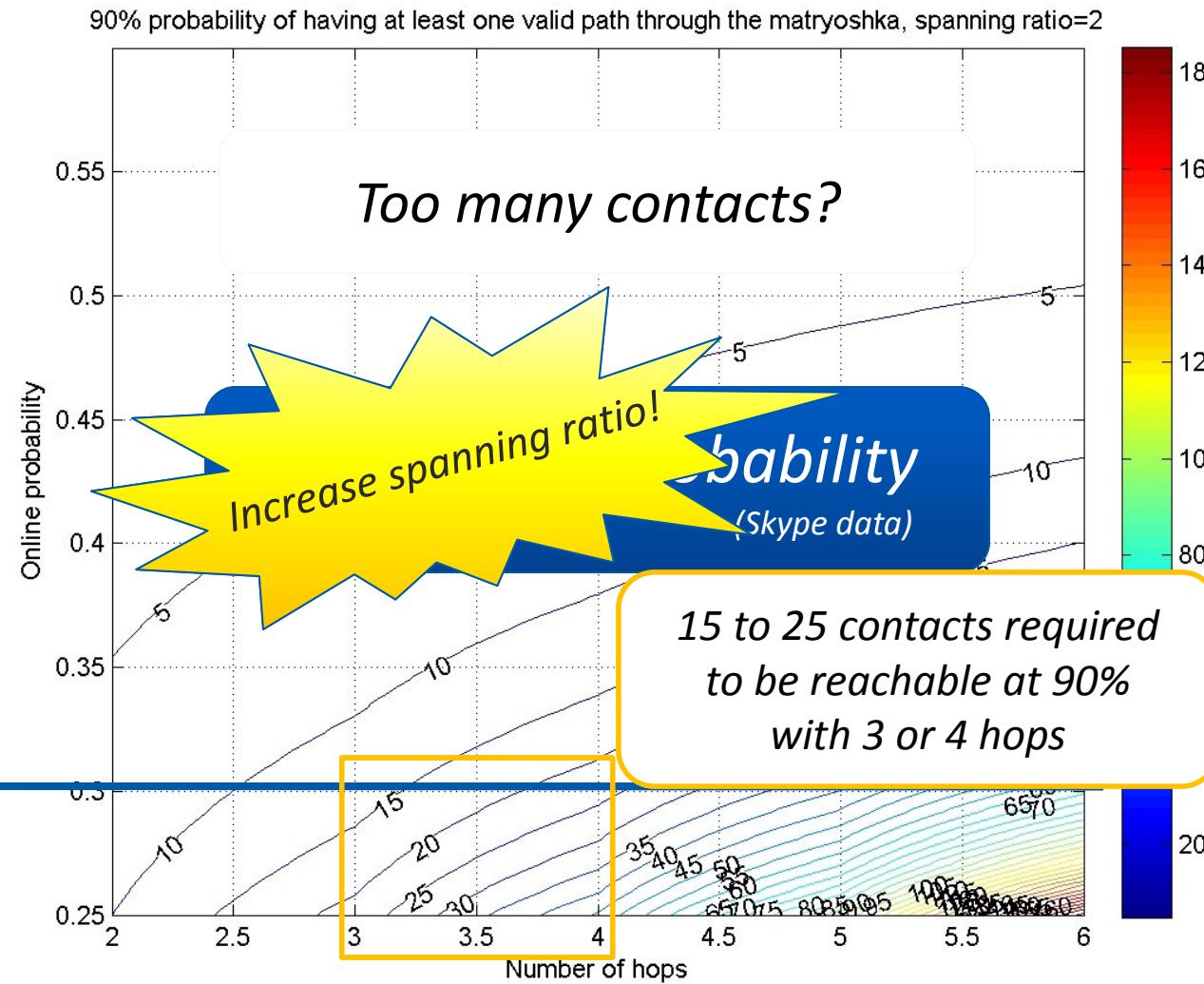
Key management

Availability

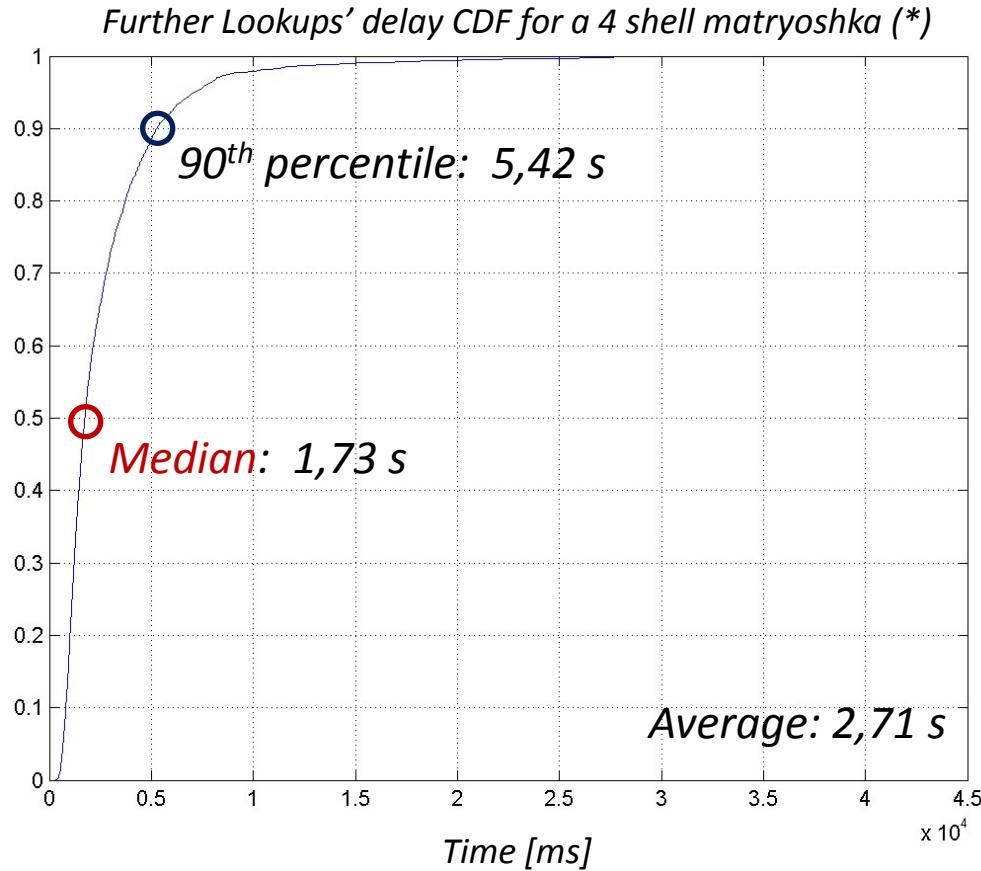


Data replication at friends' nodes

Performance - Reachability



Performance - Delay



Total data lookup time:

$$T_{dl} = T_{DHT} + T_{Mat}$$

- *Further lookups: $T_{DHT}=0$ thanks to caching*

(*) Data computed by applying the Monte Carlo sampling technique on single hop delay measurements and on delay measurement for a successful DHT key lookup in KAD (Biersack..)

Conclusions for Safebook



■ Safebook

- Decentralized OSN, based on mutual trust
- Modelled, analysed, simulated and prototype

■ Open Challenges

- ***Performance*** is insufficient
- ***Availability*** questionable (correlated churn)
- ***Concealed participation*** impossible

The screenshot shows the Safebook interface. At the top, there are navigation links: 'Settings ▾', 'Square', 'Podium', and 'Log in'. Below the header, a banner reads 'Secure Online Social Network'. The main area features a user profile for 'Leucio Antonio Cutillo' with a status update: 'Leucio Antonio Cutillo is going on working on Safebook, and hopes to see you all soon on this new Social Network!'. Below the profile is a 'Find Friend' search bar and a list of friends: Paolo Viotti, Etienne Peron, Luca Boasso, and JB Barrau. To the right, there is a discussion thread with three messages:

- Leucio Antonio Cutillo: 'In Safebook, you can decide who has access to your personal information (contact list, pictures, status, posts..)'
- Refik Molva: 'Right, and there's no central entity storing everybody's data..'
- Thorsten Strufe: 'Moreover, Safebook addresses severe concerns, such as the impersonation of users'

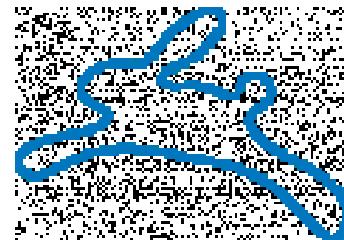
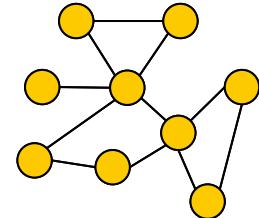
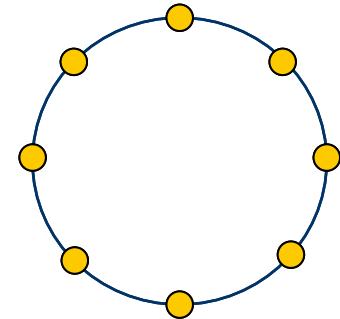
At the bottom of the thread, there is a text input field for a reply and a 'Reply' button.

- [1] WONS 2008
- [2] WWW 2009
- [3] WoWMoM 2009
- [4] IEEE CommMag 2010
- [5] Eurosyst/SNS 2010
- [6] FC/RLCPs 2011
- [7] WoWMoM 2011

Bullet Proof Privacy – Social Overlays



- Threat: Observability of behavior and identity
- Aim: Conceal participation and prevent (untrusted) observation
- Communication substrate from trusted links:
 - Social graph defines overlay
 - Restricted connectivity
 - ***Adapt addresses/identifiers*** (without data loss)
 - Estimate routing structure
 - Route on imprecise embedding with (Poly)log path length

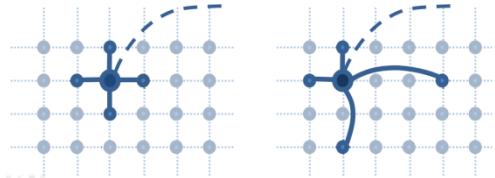


Enhancing Routing and Embeddings

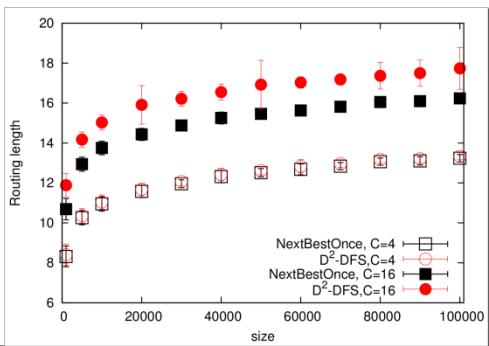


Darknet Routing [1]

- Darknet model $K(n, d, C, L)$



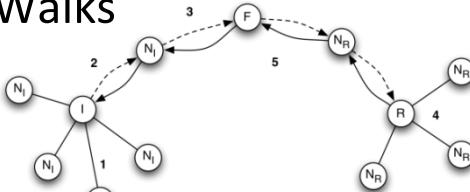
- Freenet routing **not** polylog
- Design of novel routing **NBO**
- Proof of polylog routing length
- Simulate and Implementation



Attacking Freenet [2]

- Cooperative ID-Adaptation:

- Random Walks
- Swap ID



- Single Adversary Attacks



- Resistant LMC embedding
- Analysis and simulation show superior resistance, performance

- WiP: *Novel Tree-based Embeddings* [3]

[1] INFOCOM mini '13

[2] SRDS/WNR '11

[3] NetSys '13

Summary



- Online Social Networks exhibit sustained, rapid growth
- Decreasing *expertise*, increased *abuse, malfunctions, leaks*
- *Protecting privacy* must be at the focus of CS research
- Important to know user behavior
- Decentralization one possibility
- Large number of projects that leverage the “social” term
- Good (*challenging!*) example for an application that gains from the properties of P2P
- A few real approaches to decentralize
 - Privacy, availability, cost (load balancing)
- Case study: Safebook and its properties