TECHNISCHE UNIVERSITÄT DRESDEN

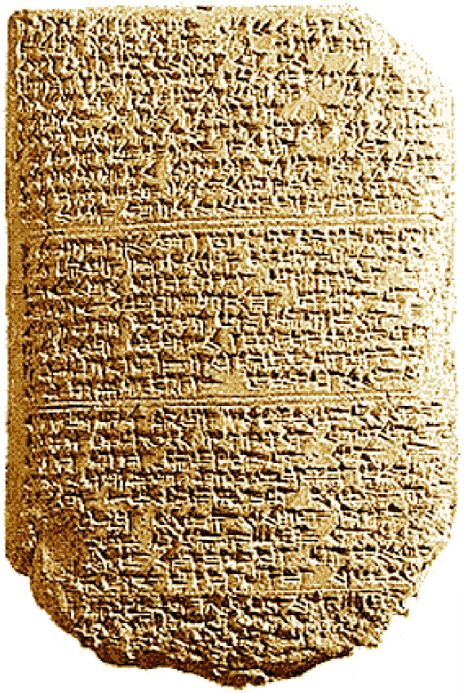
DRESDEN concept

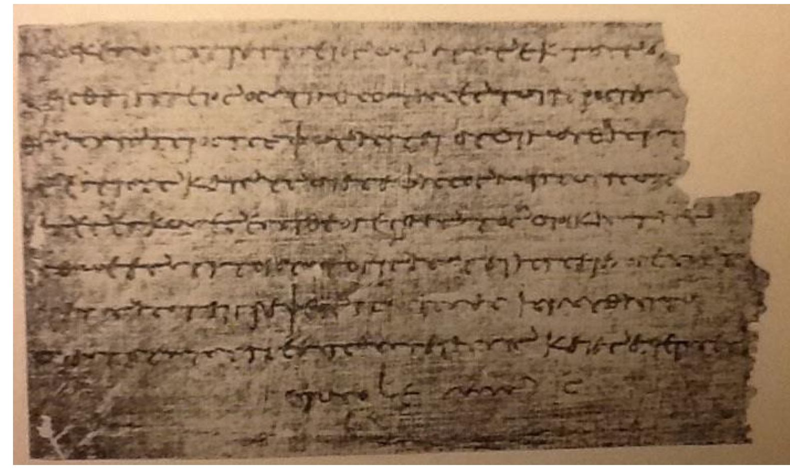
Privacy in Social-X

Thorsten Strufe

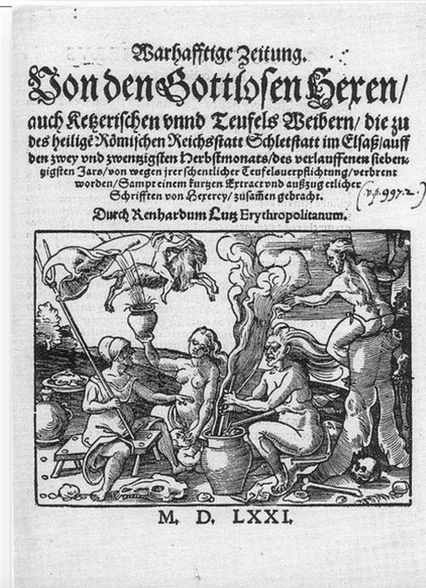
Nijmegen, 22.06.2017

Mail and „Telecommunication“





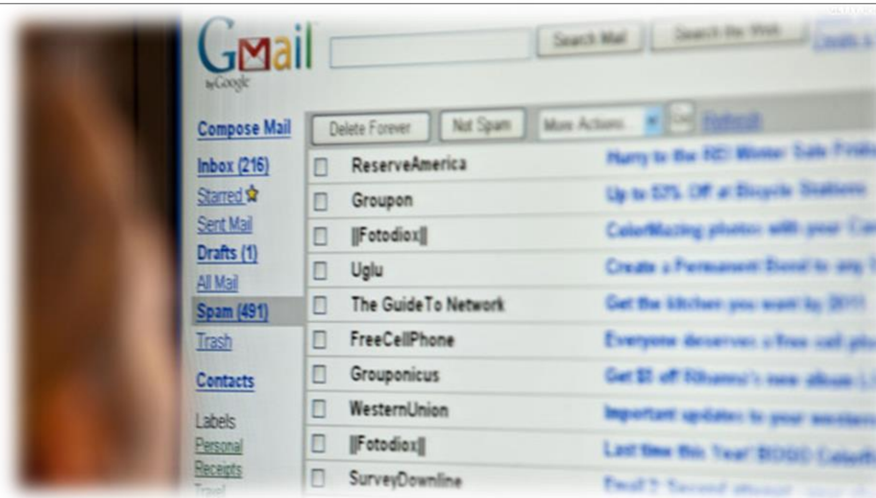
Information Acquisition


Warhafftige Zeitung.
Von den Gottlosen Hexen/
auch Ketzerischen vnnd Teufels Weibern/ die zu
des heiligẽ Römischen Reichsstatt Schletstatt im Elsaß/auff
den zwey vnd zwentzigsten Herbstmonats/des verlauffenen sieben-
zigsten Jars/von wegen jrer schentlicher Teufelsuerpflichtung/verbrent
worden/Sampt einem kurtzen Extract vnd außzug etlicher
Schrifften von Hexerey/ zůsamen gebracht.
Durch Renhardum Lutz Erythropolitanum.
M. D. LXXI.

Trade and Payments

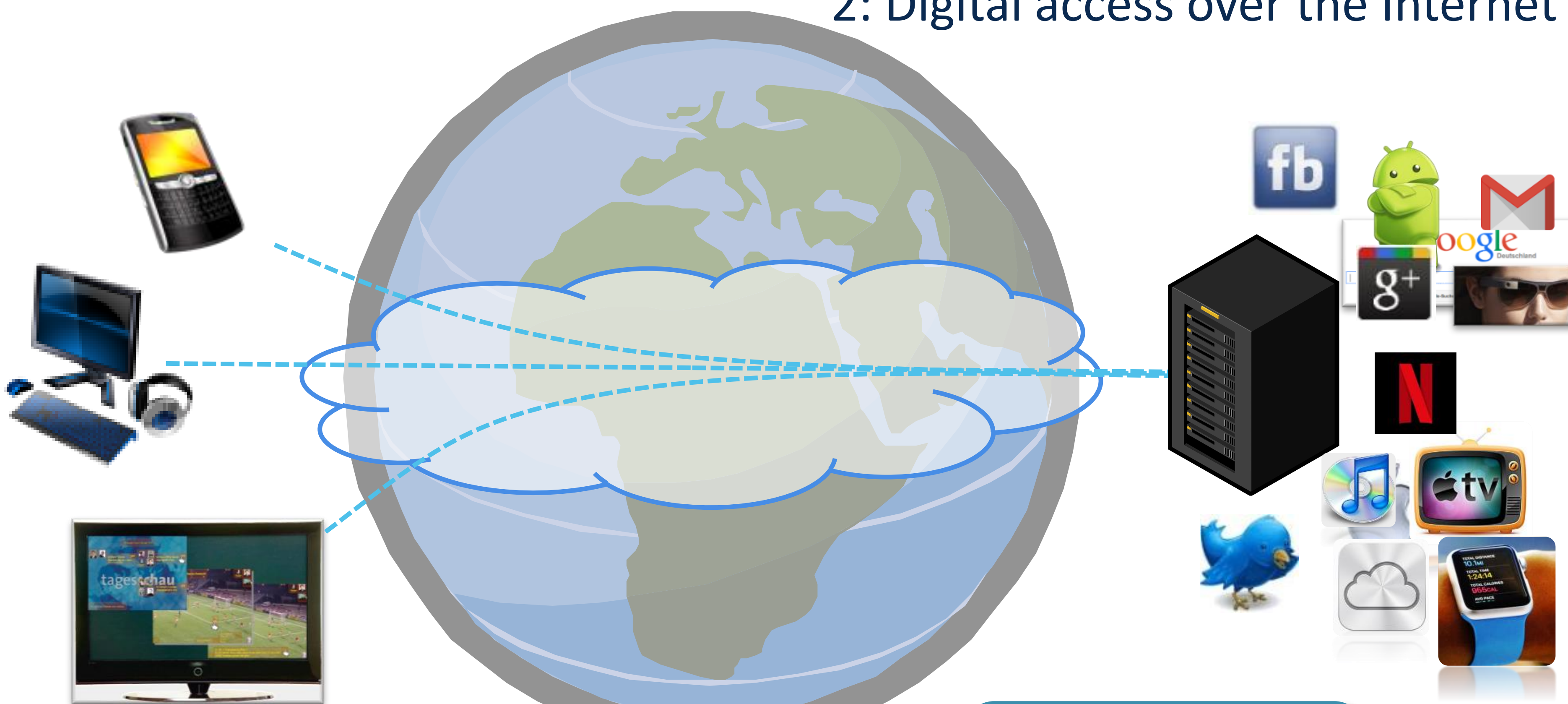
Access: Type and Scope

Our (brave?) New World...



Access: Type and Scope, Today

1: Central service providers
2: Digital access over the Internet



Secrets are lies.
Sharing is caring.
Privacy is theft.

The Stakeholders

Partner

Advertisers

zynga

Extending Partner

Cloud/CDN Provider

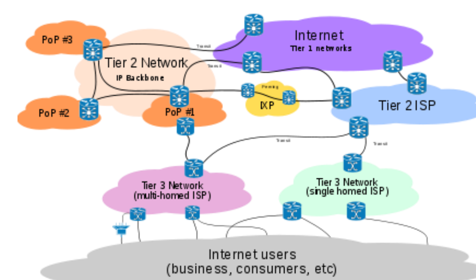
CLOUDFLARE

Akamai

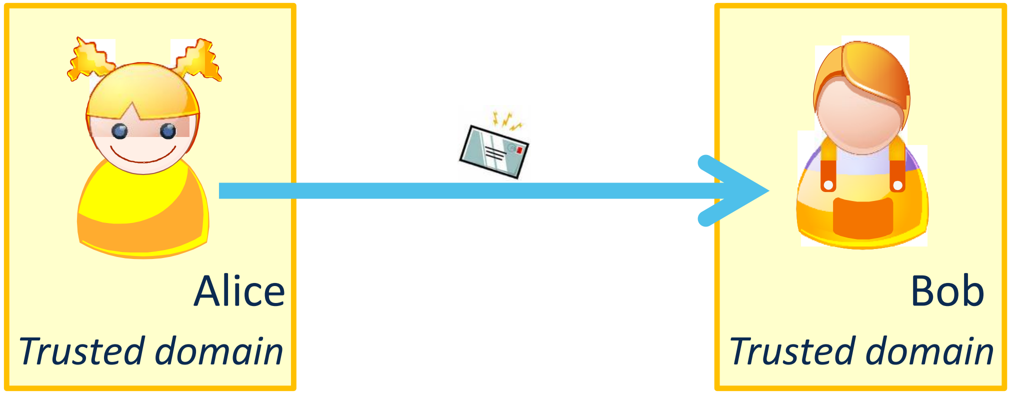
Public

Institutions

Network Provider

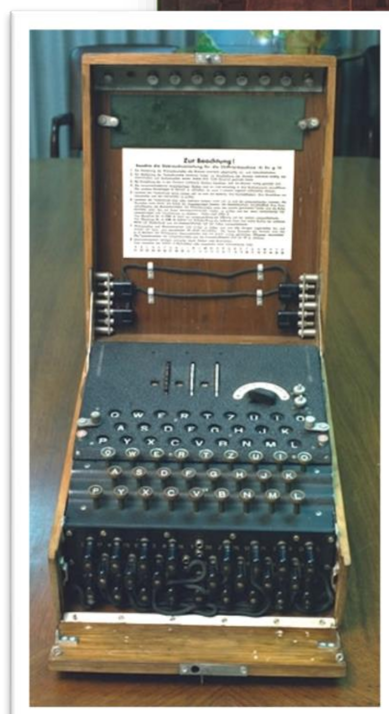


The Traditional Security View



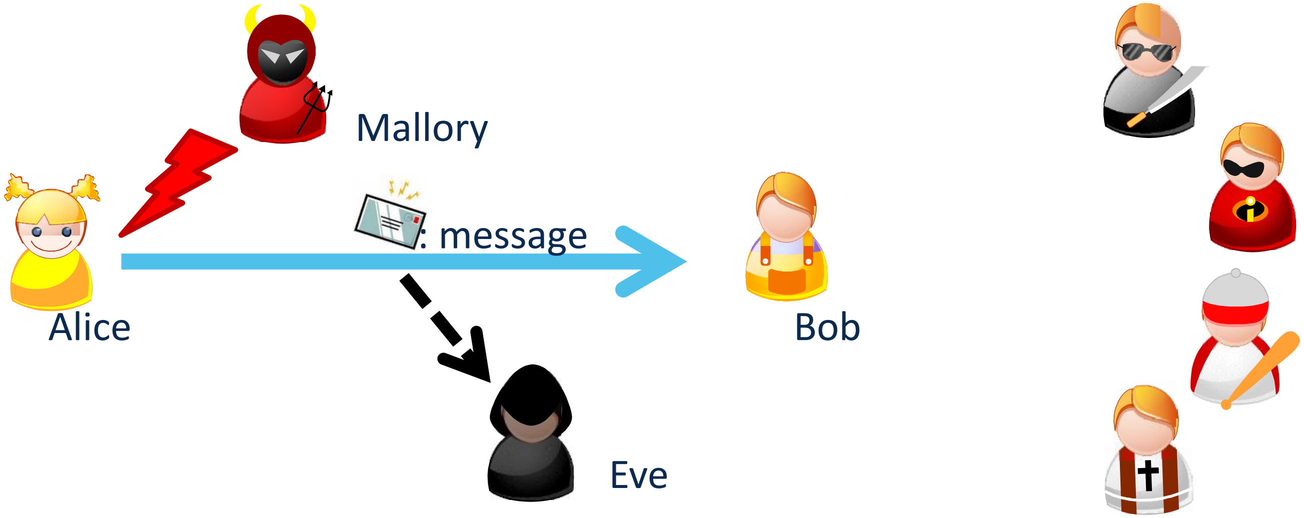
Threats!!

- Data loss
  - Data accessible to unintended parties
- Manipulation and forgery
  - Tampered, spoofed data



Classic Security Goals and Adversaries

- Confidentiality
  - Data transmitted or stored should only be revealed to the intended audience
- Integrity
  - Modification of data is detected (identify source, first!)
- Availability
  - Services should function correctly upon request



Privacy

- So what is this thing, anyways?

TECHNISCHE UNIVERSITÄT DRESDEN

Privacy (according to google et al.)

Which disclosures are people concerned about? (study from ’10)

Privacy

- So what is this thing, anyways?

Privacy „in the US": Right to be let alone

- Samuel Warren, Louis Brandeis: "**The Right to Privacy**", Harvard Law Review, Vol. IV, No. 5, 15th December 1890

- **Reason**: "snapshot photography" (recent innovation at that time)
  - allowed newspapers to publish photographs of individuals without obtaining their consent.
  - private individuals were being continually injured
  - this practice weakened the "moral standards of society as a whole"

- **Consideration**:
  - basic principle of common law: individual shall have full protection in person and in property
  - "it has been found necessary from time to time to define anew the exact nature and extent of such protection"
  - "Political, social, and economic changes entail the recognition of new rights"

- **Conclusion**:
  - "***right to be let alone***"

Privacy „in Europe“: Data Protection

- Principles
  - collect and process personal data **fairly and lawfully**
  - **purpose binding**
    - keep it only for one or more specified, explicit and lawful purposes
    - use and disclose it only in ways compatible with these purposes
  - **data minimization**
    - adequate, relevant and not excessive wrt. the purpose
    - retained no longer than necessary
  - **transparency**
    - inform who collects which data for which purposes
    - inform how the data is processed, stored, forwarded etc.
  - **user rights**
    - access to the data, correction, deletion
  - **keep the data safe and secure**

Now, what about Privacy?

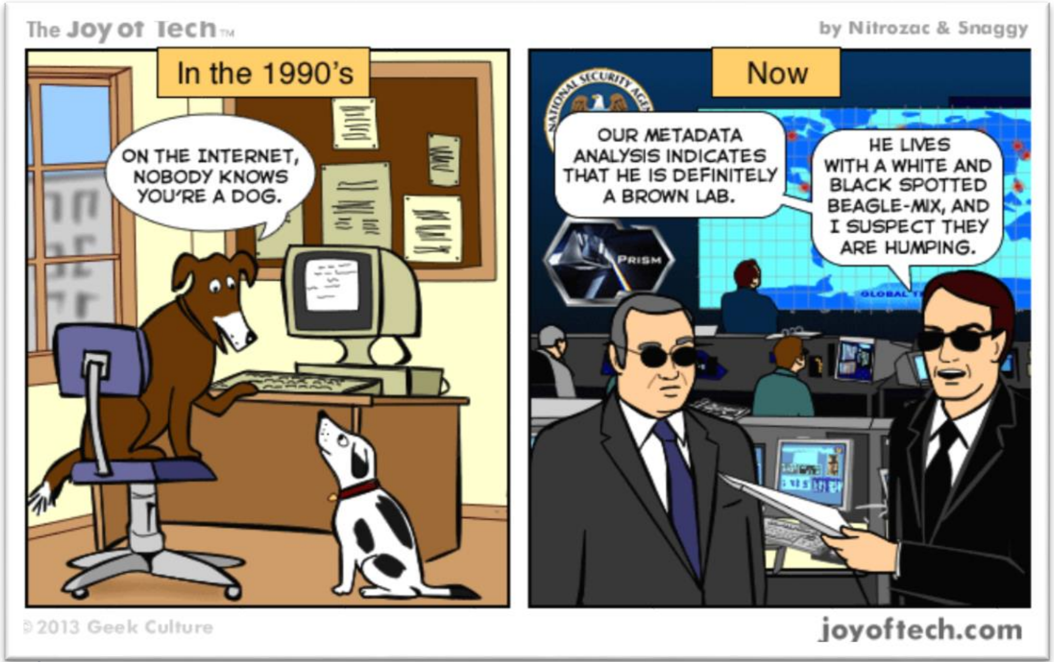
- Protect data?
- Rather: Protect integrity of individuals
- Hence: Protect individuals FROM data

- Hang on! What‘s all this „data“ about?

Types of Data

- Data without any relation to individuals
  - Simulation data
  - Measurements from experiments

- Data with (obvious) relation to individuals
  - Types
    - Content
    - Meta data
  - Revelation
    - Consciously
    - Unconsciously



Types of Disclosure

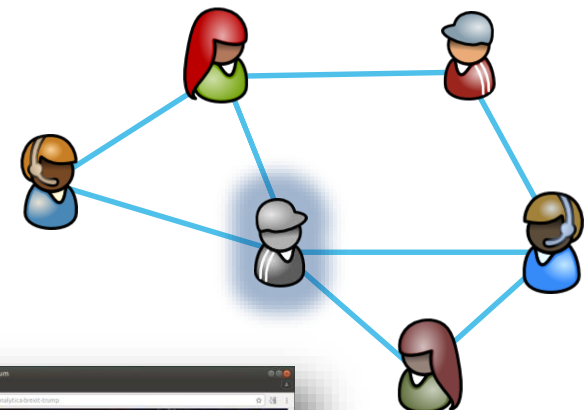
- What can be disclosed?
- Disclosure of attributes
  - Infer a (hidden) attribute of an individual

- Disclosure of identity
  - Identify an individual in a dataset

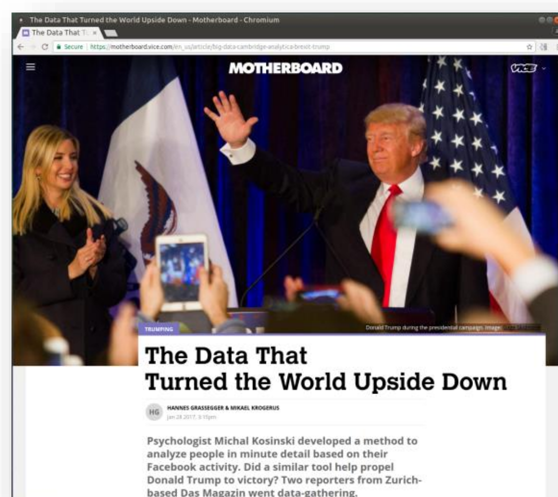
  - Both must be prevented!

„It‵s only Meta Data"

- „Facebook Mining" attacks
- Single term lecture (students without any prior knowledge on ML)
- Information (ab)used:
  - Partial profiles
  - Neighborhood



- Inferred, with high accuracy:
  - Gender
  - Age
  - Education level
  - Expected tenure with employer
  - Sexual preferences
  - *Religious beliefs*
  - *Political preferences*

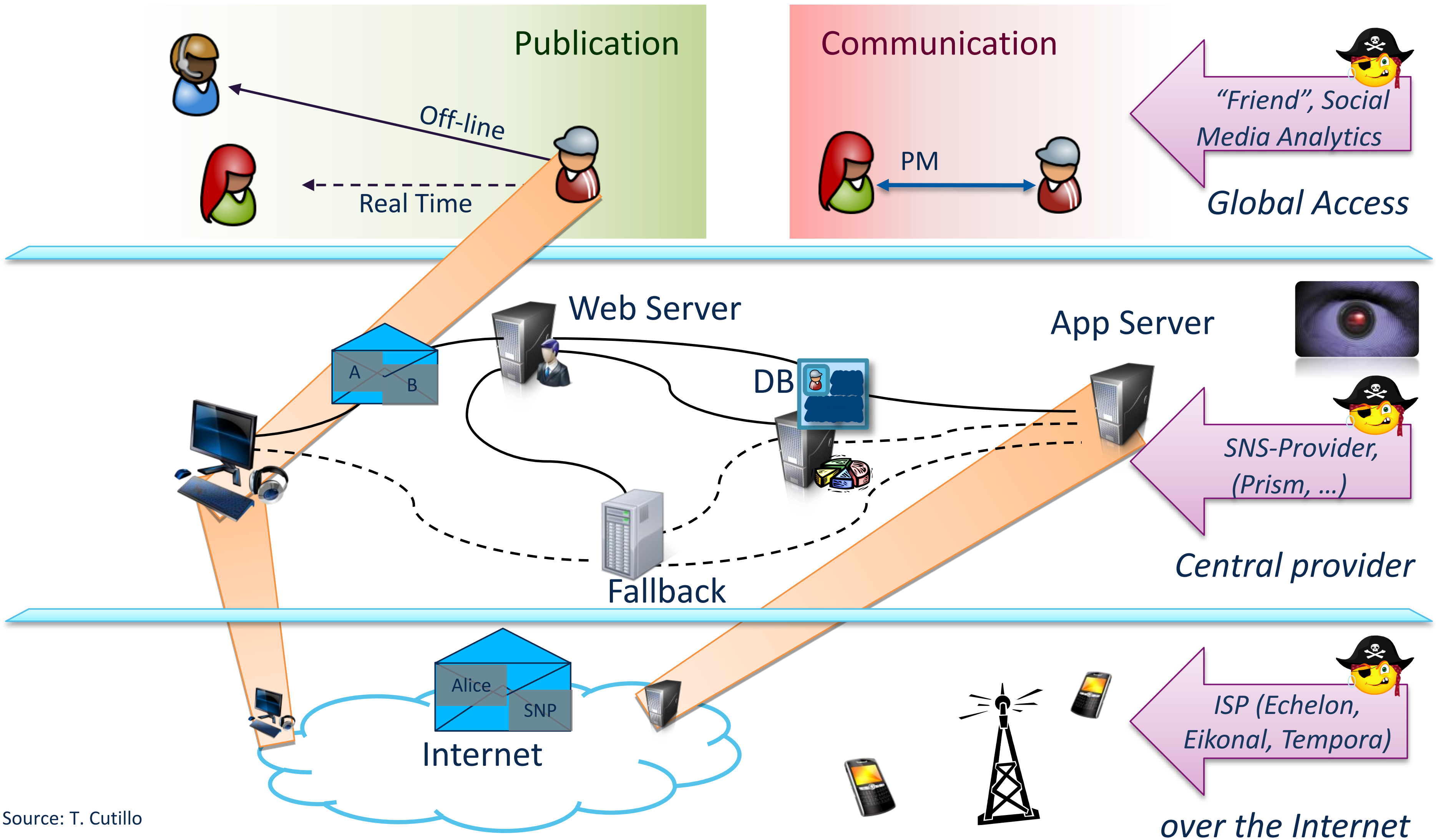

MOTHERBOARD

The Data That Turned the World Upside Down

HANNES GRASSEGGER & MIKAEL KROGERUS

Psychologist Michal Kosinski developed a method to analyze people in minute detail based on their Facebook activity. Did a similar tool help propel Donald Trump to victory? Two reporters from Zurich-based Das Magazin went data-gathering.

Model, Access, and Adversaries



Source: T. Cutillo

- Explicit
  - Created content
  - Comments
  - Structural interaction (contacts, likes)

But I‘ve got nothing to hide…?

- „Meta data“
  - ***Session artifacts*** (time of actions)
  - ***interest*** (retrieved profiles; membership in groups/participation in discussions)
  - ***influence***
  - Clickstreams, ad preferences
  - ***communication*** (end points, type, intensity, frequency, extent)
  - ***location*** (IP; shared; gps coordinates)

- Inferred
  - Preference– and
  - Image recognition models
  - Personal details

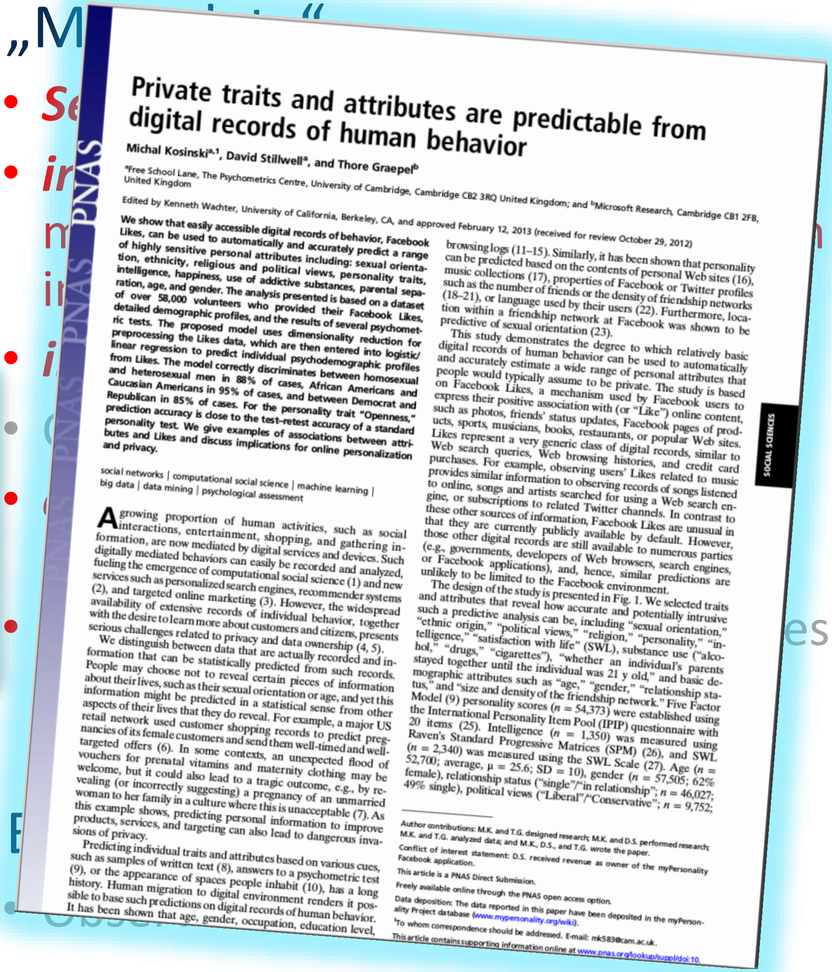
- Externally correlated
  - Observation in ad networks

- Explicit
  - Created content
  - Comments
  - Structural interaction (contacts, likes)

- Inferred
  - Preference– and
  - Image recognition models
  - Personal details (location, health…)

But I‘ve got nothing to hide…?

- „M
  - Se
  - ir
  - m
  - ir
  - i
  - es)
- E
  - Obser



Private traits and attributes are predictable from digital records of human behavior

Michal Kosinski[a,1], David Stillwell[a], and Thore Graepel[b]

[a]Free School Lane, The Psychometrics Centre, University of Cambridge, Cambridge CB2 3RQ United Kingdom; and [b]Microsoft Research, Cambridge CB1 2FB, United Kingdom

Edited by Kenneth Wachter, University of California, Berkeley, CA, and approved February 12, 2013 (received for review October 29, 2012)

We show that easily accessible digital records of behavior, Facebook Likes, can be used to automatically and accurately predict a range of highly sensitive personal attributes including: sexual orientation, ethnicity, religious and political views, personality traits, intelligence, happiness, use of addictive substances, parental separation, age, and gender. The analysis presented is based on a dataset of over 58,000 volunteers who provided their Facebook Likes, detailed demographic profiles, and the results of several psychometric tests. The proposed model uses dimensionality reduction for preprocessing the Likes data, which are then entered into logistic/linear regression to predict individual psychodemographic profiles from Likes. The model correctly discriminates between homosexual and heterosexual men in 88% of cases, African Americans and Caucasian Americans in 95% of cases, and between Democrat and Republican in 85% of cases. For the personality trait "Openness," prediction accuracy is close to the test–retest accuracy of a standard personality test. We give examples of associations between attributes and Likes and discuss implications for online personalization and privacy.

social networks | computational social science | machine learning | big data | data mining | psychological assessment

A growing proportion of human activities, such as social interactions, entertainment, shopping, and gathering information, are now mediated by digital services and devices. Such digitally mediated behaviors can easily be recorded and analyzed, fueling the emergence of computational social science (1) and new services such as personalized search engines, recommender systems (2), and targeted online marketing (3). However, the widespread availability of extensive records of individual behavior, together with the desire to learn more about customers and citizens, presents serious challenges related to privacy and data ownership (4, 5).

We distinguish between data that are actually recorded and information that can be statistically predicted from such records. People may choose not to reveal certain pieces of information about their lives, such as their sexual orientation or age, and yet this information might be predicted in a statistical sense from other aspects of their lives that they do reveal. For example, a major US retail network used customer shopping records to predict pregnancies of its female customers and send them well-timed and well-targeted offers (6). In some contexts, an unexpected flood of vouchers for prenatal vitamins and maternity clothing may be welcome, but it could also lead to a tragic outcome, e.g., by revealing (or incorrectly suggesting) a pregnancy of an unmarried woman to her family in a culture where this is unacceptable (7). As this example shows, predicting personal information to improve products, services, and targeting can also lead to dangerous invasions of privacy.

Predicting individual traits and attributes based on various cues, such as samples of written text (8), answers to a psychometric test (9), or the appearance of spaces people inhabit (10), has a long history. Human migration to digital environment renders it possible to base such predictions on digital records of human behavior. It has been shown that age, gender, occupation, education level,

browsing logs (11–15). Similarly, it has been shown that personality can be predicted based on the contents of personal Web sites (16), music collections (17), properties of Facebook or Twitter profiles such as the number of friends or the density of friendship networks (18–21), or language used by their users (22). Furthermore, location within a friendship network at Facebook was shown to be predictive of sexual orientation (23).

This study demonstrates the degree to which relatively basic digital records of human behavior can be used to automatically and accurately estimate a wide range of personal attributes that people would typically assume to be private. The study is based on Facebook Likes, a mechanism used by Facebook users to express their positive association with (or "Like") online content, such as photos, friends' status updates, Facebook pages of products, sports, musicians, books, restaurants, or popular Web sites. Likes represent a very generic class of digital records, similar to Web search queries, Web browsing histories, and credit card purchases. For example, observing users' Likes related to music provides similar information to observing records of songs listened to online, songs and artists searched for using a Web search engine, or subscriptions to related Twitter channels. In contrast to these other sources of information, Facebook Likes are unusual in that they are currently publicly available by default. However, those other digital records are still available to numerous parties (e.g., governments, developers of Web browsers, search engines, or Facebook applications), and, hence, similar predictions are unlikely to be limited to the Facebook environment.

The design of the study is presented in Fig. 1. We selected traits and attributes that reveal how accurate and potentially intrusive such a predictive analysis can be, including "sexual orientation," "ethnic origin," "political views," "religion," "personality," "intelligence," "satisfaction with life" (SWL), substance use ("alcohol," "drugs," "cigarettes"), "whether an individual's parents stayed together until the individual was 21 y old," and basic demographic attributes such as "age," "gender," "relationship status," and "size and density of the friendship network." Five Factor Model (9) personality scores (n = 54,373) were established using the International Personality Item Pool (IPIP) questionnaire with 20 items (25). Intelligence (n = 1,350) was measured using Raven's Standard Progressive Matrices (SPM) (26), and SWL (n = 2,340) was measured using the SWL Scale (27). Age (n = 52,700; average, μ = 25.6; SD = 10), gender (n = 57,505; 62% female), relationship status ("single"/"in relationship"; n = 46,027; 49% single), political views ("Liberal"/"Conservative"; n = 9,752;

Author contributions: M.K. and T.G. designed research; M.K. and D.S. performed research; M.K. and T.G. analyzed data; and M.K., D.S., and T.G. wrote the paper.

Conflict of interest statement: D.S. received revenue as owner of the myPersonality Facebook application.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

Data deposition: The data reported in this paper have been deposited in the myPersonality Project database (www.mypersonality.org/wiki).

[1]To whom correspondence should be addressed. E-mail: mk583@cam.ac.uk.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.

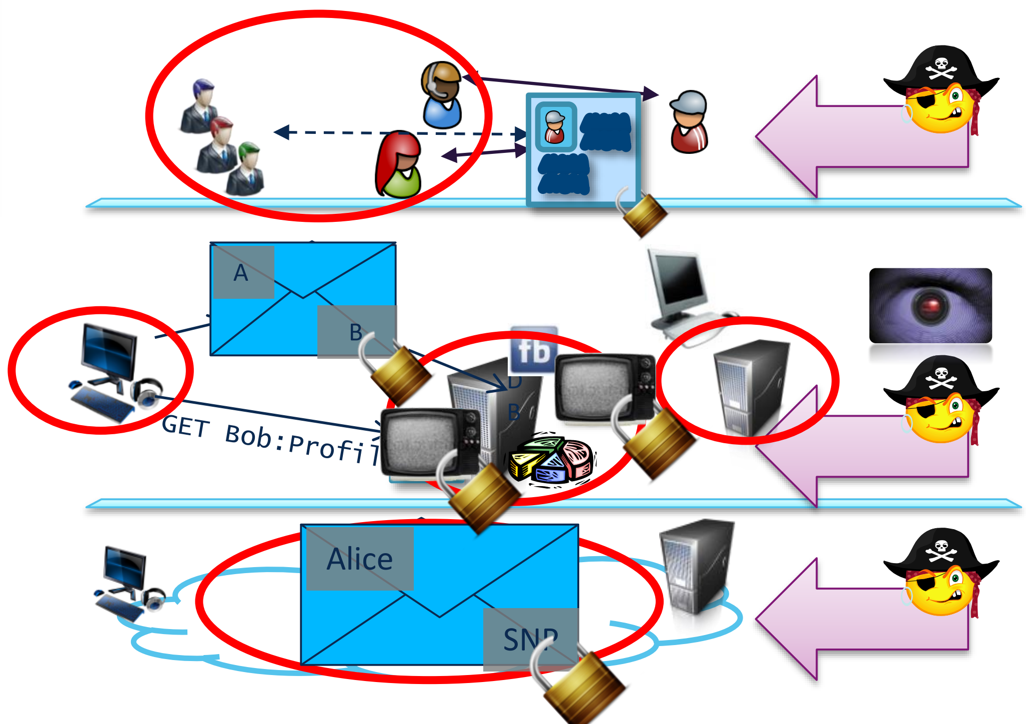
PNAS

SOCIAL SCIENCES

TECHNISCHE UNIVERSITÄT DRESDEN

Solution Classes and our Contributions

- Resilient Networking
  - Confidential transmission
  - Defending the network
- PETs
  - Anonymous communication
  - Service decentralisation
- System security
  - Protocol/service partitioning
  - Hardware extensions (SGX)
- User Understanding
  - Privacy assessment, metrics
  - Intention recognition
  - User support



PETs: Entire Distribution of Data and Control

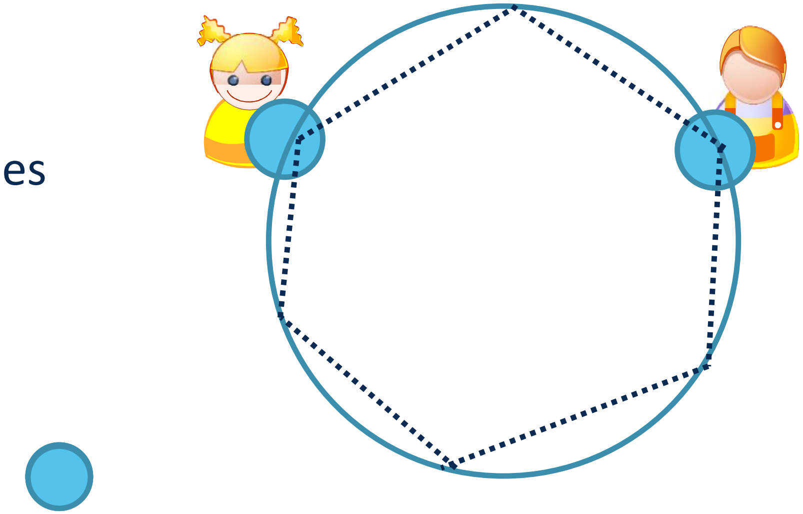
- TOR allows you to hide your IP, but what about the service itself…
- Decentralize the services
- Federated SNS
- DOSN
- Social overlays

Resilient Social Overlays

- Prevent identification, censorship and retribution.

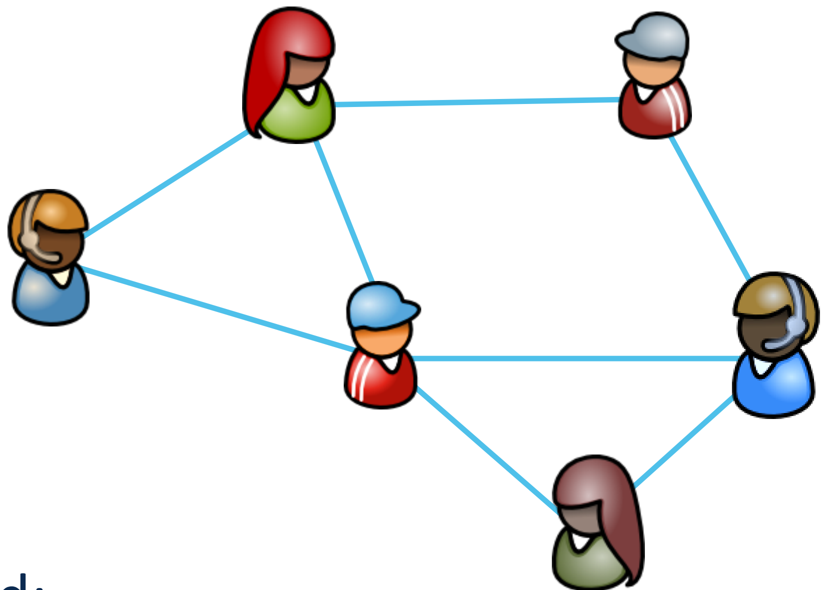
- From DOSN to darknets: Tightening requirements
  - Concealed participation
  - Unobserveability
  - Metadata privacy (sender-, receiver-, relationship anonymity)

- So where‘s the problem?
- Classic overlays:
  - Two degrees of freedom: ID, links
  - Eclipse, *-hole attacks
  - Disclosure of IP address to unknown parties



Preventing Censorship and Surveillance

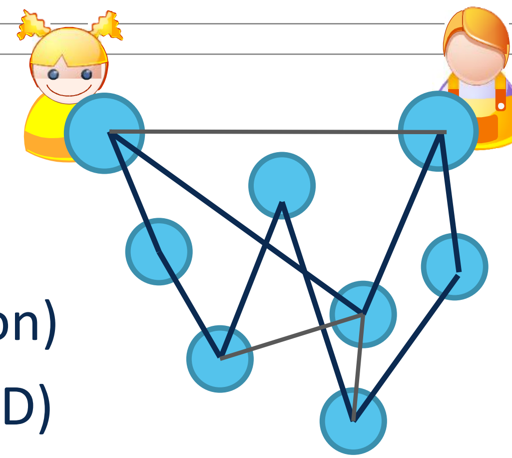
- Let‘s go „dark“!



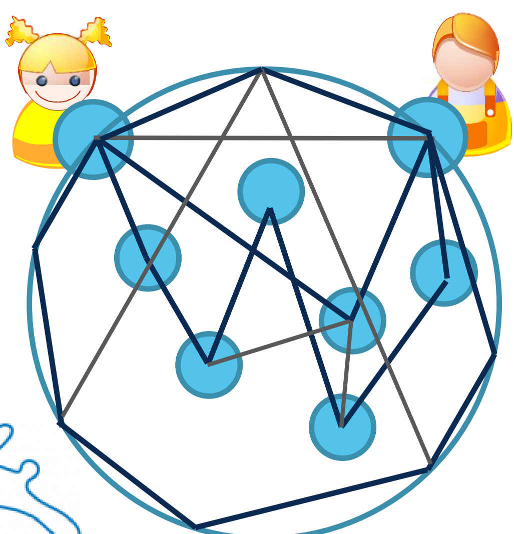
- Friend-to-Friend:
  - Membership concealing
  - Freedom from observation
  - Resilient to censorship and sabotage

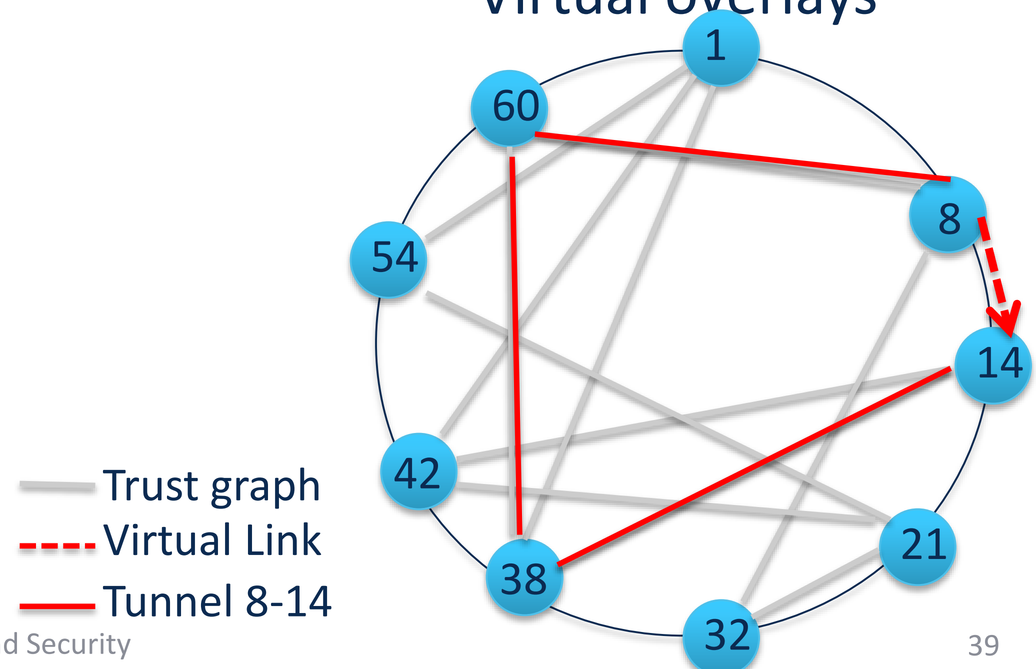
Social Overlays: Embedding/Virtual Overlays

- Concepts of social overlays:
  - Constrain connectivity to social links
  - Constrain information (hop-by-hop anonymization)
  - *Attempt* to route messages (degree of freedom: ID)



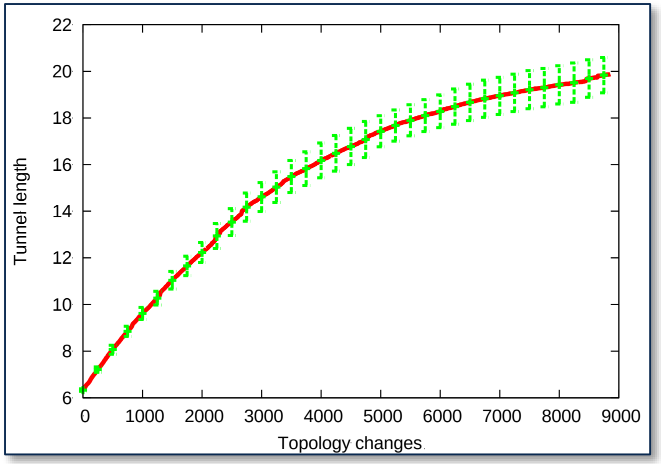
- Embeddings   Virtual overlays

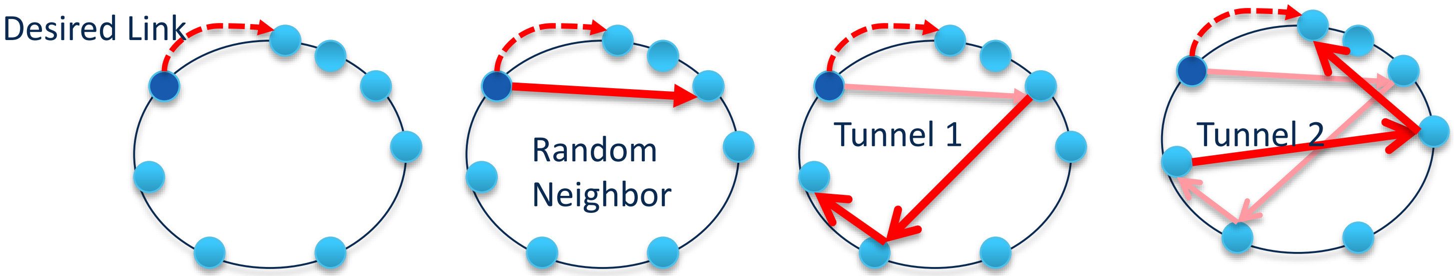




Virtual Overlays – Tunnel Maintenance

- Establishment & maintenance of „trails“
  - Flooding
    - Finds shortest paths, is excessively expensive
  - Routing
    - Leverage overlay routing to trail endpoint
    - Concatenate existing tunnels





- e.g. WSN, X-Vine

- Efficiency: *Can tunnels remain polylog over time – at polylog cost?*
- Proof by contradiction: *Concatenation of trails diverges beyond polylog length over time*

[1] Roos and Strufe: INFOCOM 2015
[2] Roos et al.: PETS 2014

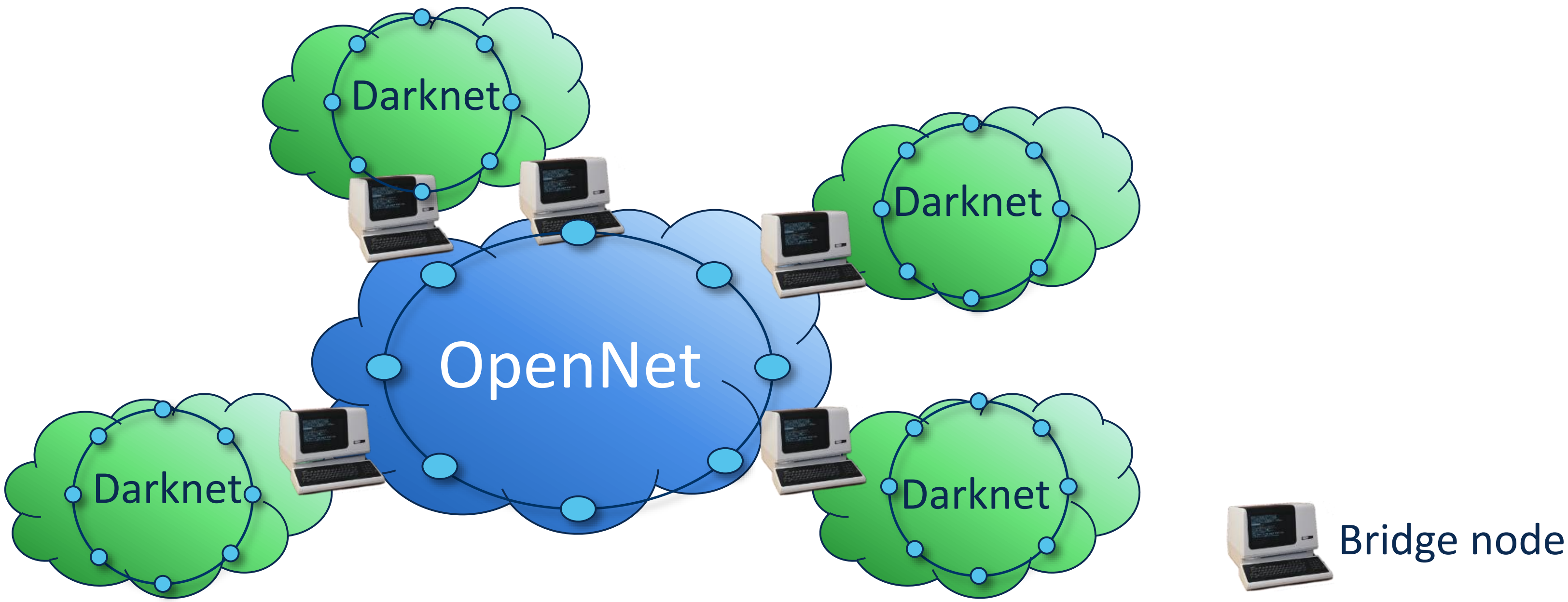
Freenet – Protecting Free Expression Online

- „Censorship resistance requires anonymous communication“
  - [Clarke 2000], [Clarke, Miller, Hong, Sandberg, Wiley 2002]

- Basic concepts
  - ***Push-based P2P data store*** with probabilistic on-path caching
  - ***Create overlay***
    - Random ID selection („location“)
    - Unidimensional lattice (unit circle)
    - Approximation of Kleinberg (see below)
  - ***Routing***
    - Information containment: Recursive routing with source rewriting
    - Greedy: distance-directed depth first search („steepest-ascent hill climbing“)

- ***Publishing, storing, and requesting nodes can‘t be identified***

Integrating Opennet and Darknets



- Each darknet exists on its own
- Nodes participating in darknet ***and*** opennet act as „bridges“

Understanding it – A Measurement Campaign

- *How does Freenet work in the first place?*
- *Does Freenet routing work (what does the topology look like)?*
- *How many people are using Freenet, and where?*
- *What usage / behavior is to be expected?*
- *What is the popularity of content?*
- *Do Darknets exist and can we find them?*
- *How resistant is Freenet to sabotage?*

TECHNISCHE UNIVERSITÄT DRESDEN

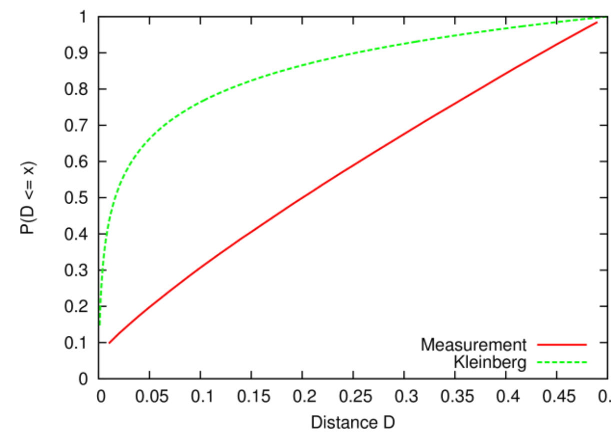
General Methodology and Setup

- How can we find out?
  - Code analyses (papers, online/“code“ documentation are not reliable)
  - Instrumentation of client software
  - Passive measurements (logging all messages)
  - Active probing (active node discovery and tracking)
  - *Campaigns: Summer/autumn `12 (1407/1410), spring `13 (1442/1457), summer/fall ‘16*

- Hardware Setup
  - 4 older machines from the lab for long term measurements:
    - 2 barebones, 1.5GHz, 2GB RAM
    - 2 sun solaris workstations
  - Our „monster“ for specific probing campaigns:
    - 4 x 16 cores, 2.8GHz, 512GB RAM
  - ***Side note**: main limiting factor is memory, each barebone hosts max. 11 nodes*

Measuring the Routing

- Methodology:
  - Log topology updates (upon changes to neighborhood)
  - Trace forwarded requests
  - *Additionally*: create Darknet of 10 nodes, and connect through own bridge
  - Simulate routing with measured, corrected DD



- Corrected distance distribution
  - Many neighbors with d < 0.05
  - Uniform distribution for d > 0.05
  - Simulated average 37 vs 13 hops (Kleinberg)

- Measured routing success
  - Opennet (92.5% of requests) yields 22.5% success
  - Darknet (7.5% of requests) yields 0.4% success

Adapt neighbor selection
Ignore Darknets, or
Skew Darknets to ID of bridge

Measuring the Population

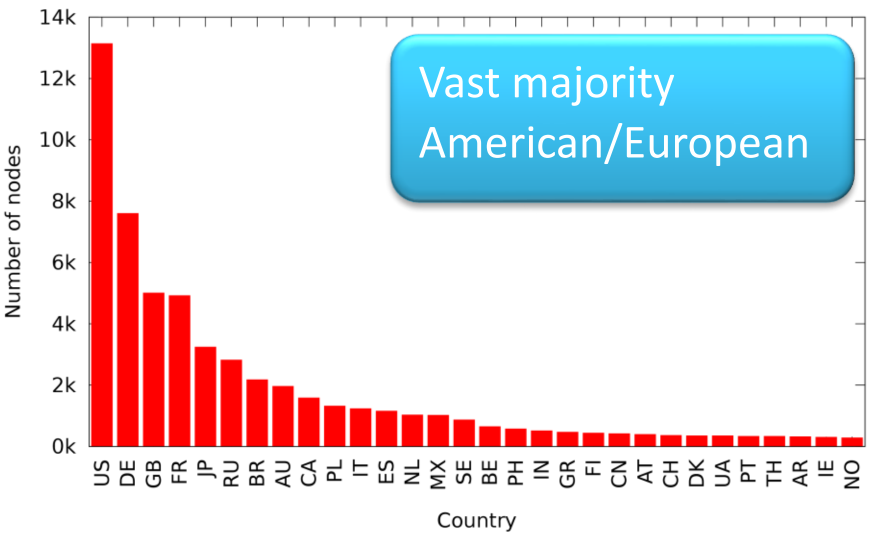
- **[1407] *FNPRoutedPing*:** Ping/Pongs of specific locations
  - Discover nodes, track selection *(55 clients, 680h)*
  - Routing success well below 100%:
    - Place *M* monitors on ID space
    - Ping monitors periodically to assess current success rates
    - Ping target and report success to server
    - On failure, ping from next monitor, until *k=5* attempts for 99.9% certainty

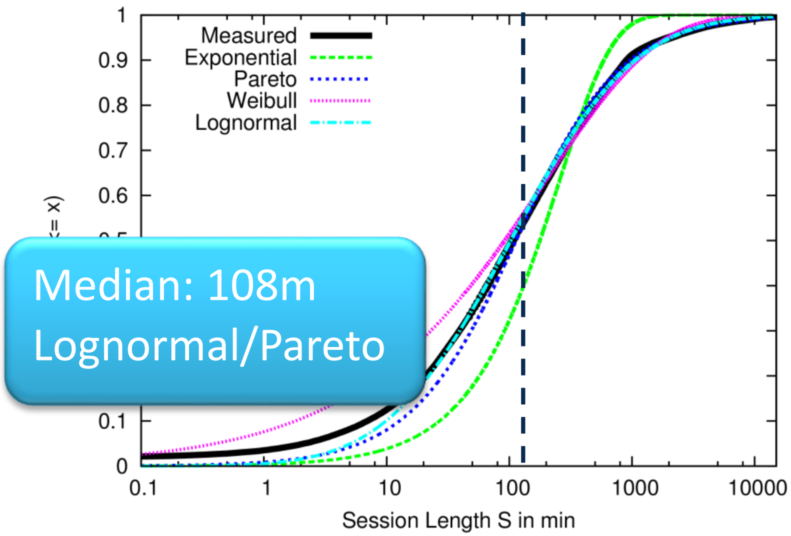
Really nice tool to track users!

- **[1410,…] *FNPRHProbeRequest*:** Random Probe for [location|uptime]
  - Probe is forwarded along 10 hops unweighted random walk
    - Estimate probability to detect node within specific interval
    - Flood FNPRHP_R_ for locations (2.4 mio/h)
    - Collect responses with timestamps
    - Extract sessions for each discovered location
  - *(150 clients, 216h)*

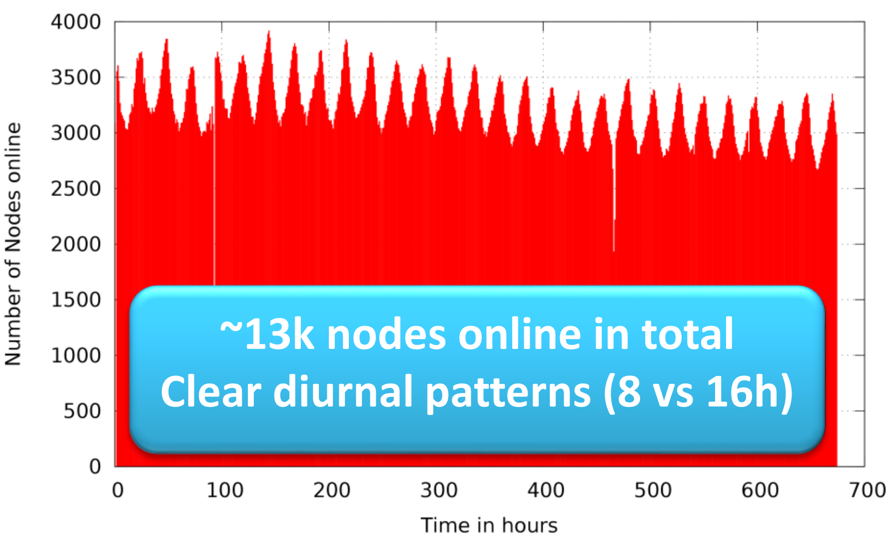
Still quite convenient tool to track users!

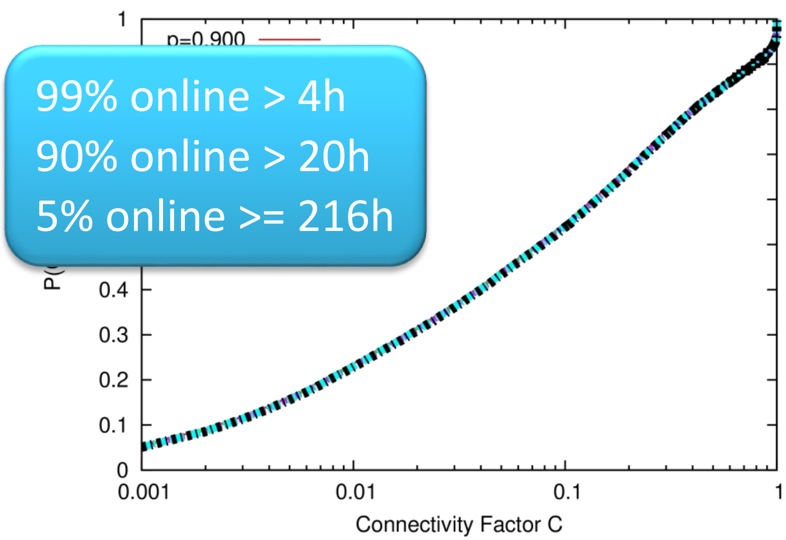
The Freenet Population (58.571 locations)



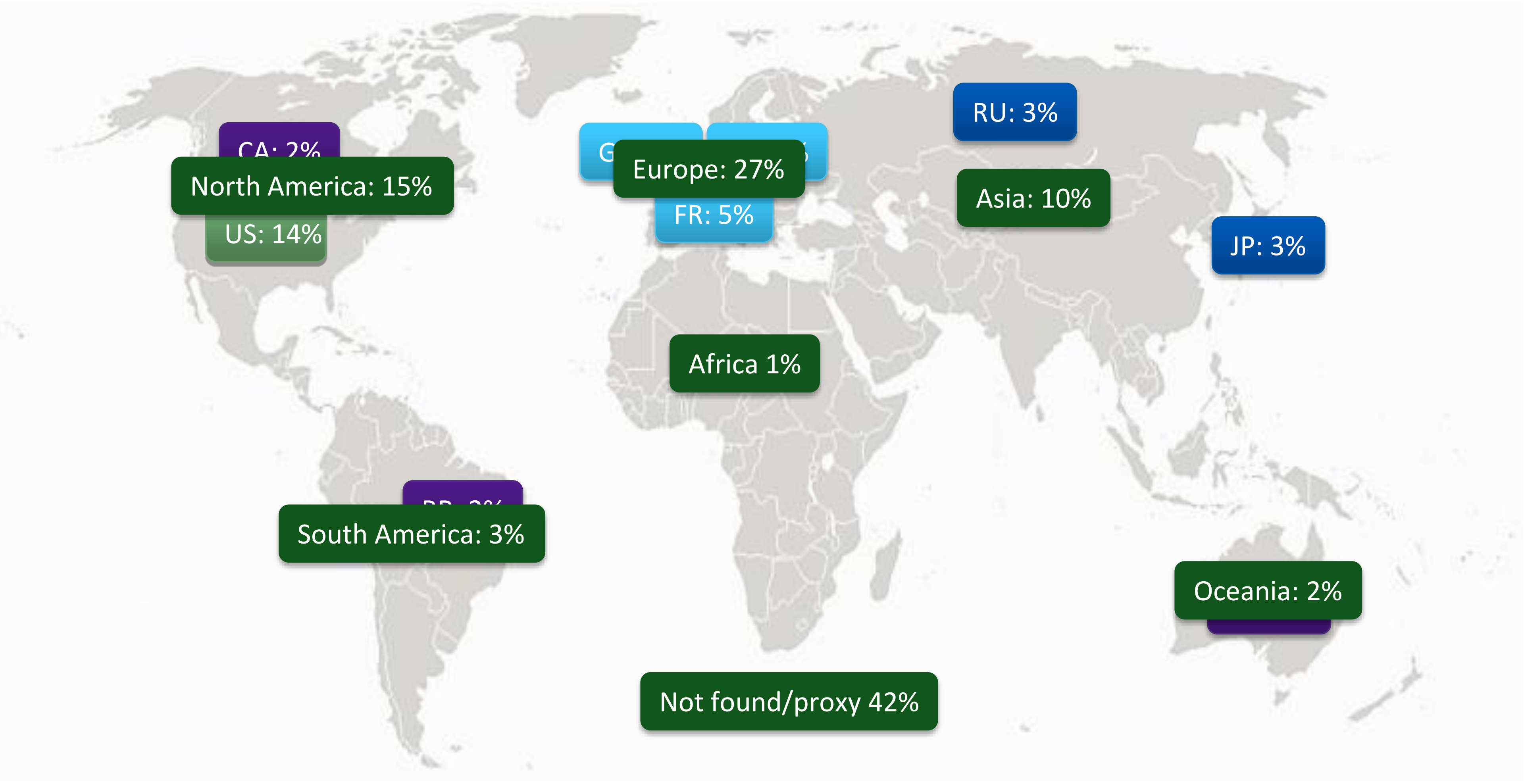


Tracking 15.503 random nodes

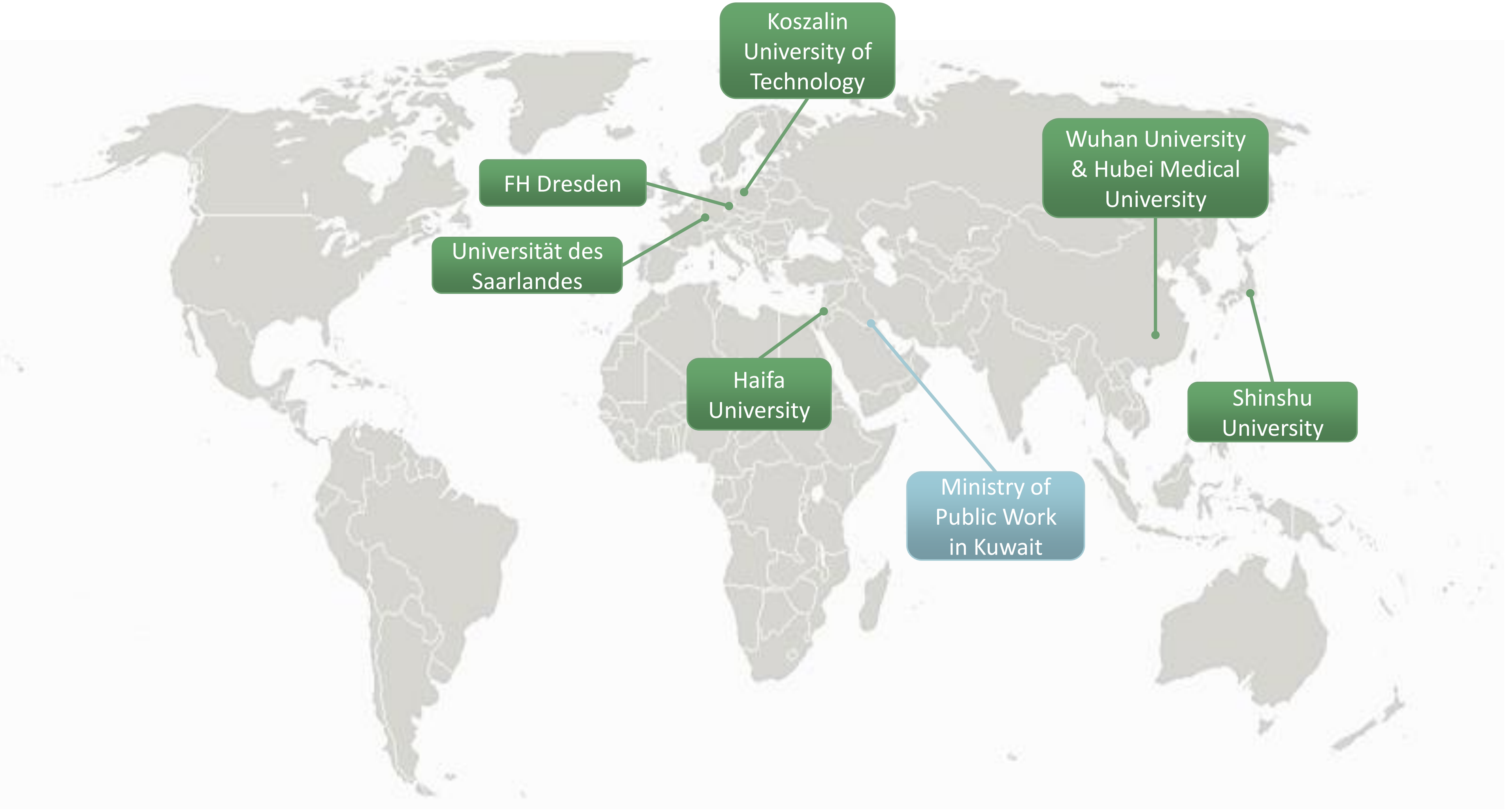




Freenet Population – Geolocation



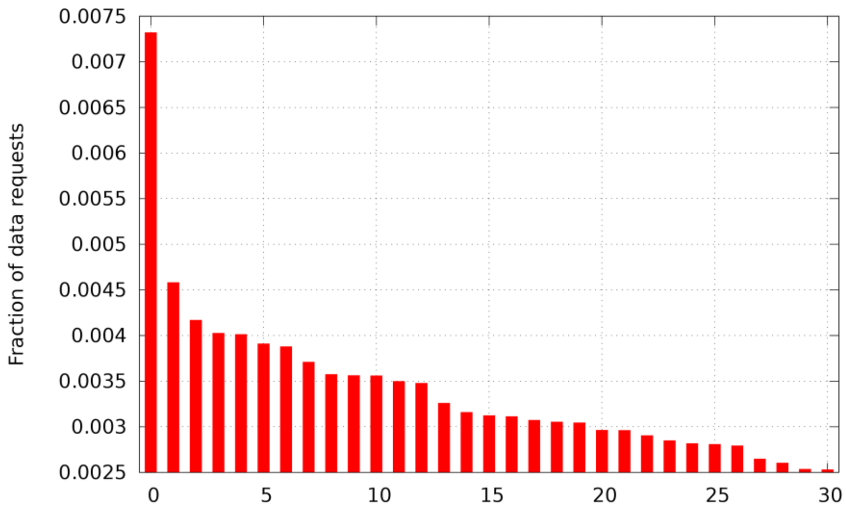
Freenet Population – Institutions



TECHNISCHE UNIVERSITÄT DRESDEN

Measuring Popularity

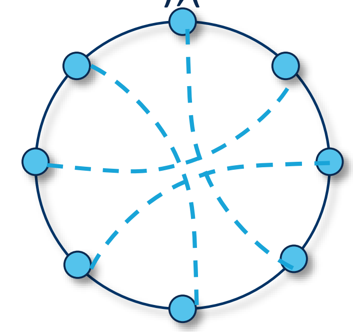
- Methodology
  - Collect routing keys from forwarded requests
  - Extract publisher‘s keys (SSK/USK)
  - Estimate content
- Measured Popularity of keys



- Order of content types (top 5)
  - Freenet updates
  - Developer blogs
  - Freesite indices
  - Freenet documentation
  - Freemail content

Freenet isn‘t about terrorism, rebellions, and organized crime… (Goto BlackMarket reloaded for that ;)

Freenet / Opennet Mode



- Aim at recreating unidimensional Kleinberg:

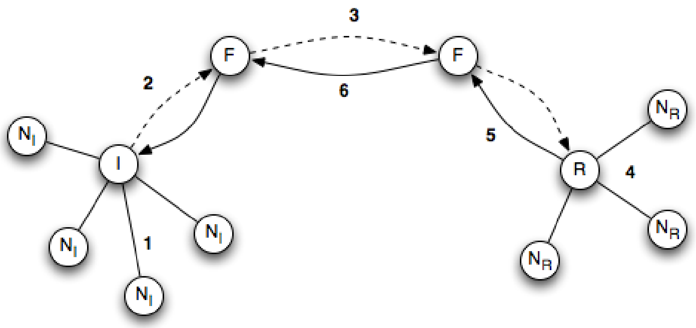
- Bootstrapping
  - Bootstrap at seed node
  - Seed node replicates and routes request according to location
  - Termini of routes establish connections

- Topology control
  - Allow neighbors depending on bandwidth
  - Establish additional connections if necessary (nodes discovered in operation)
  - Additionally: Connect to further discovered nodes (content discovery)

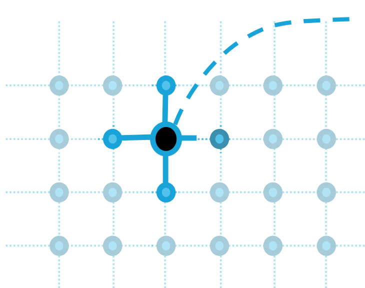
- ***Sender/storage/receiver „anonymity“, participation disclosed***

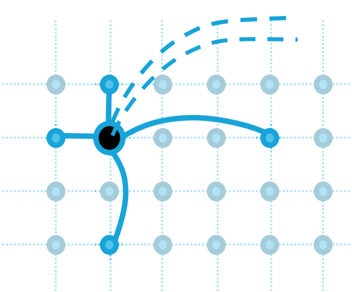
The Dark Freenet

- Only deployed (used) darknet



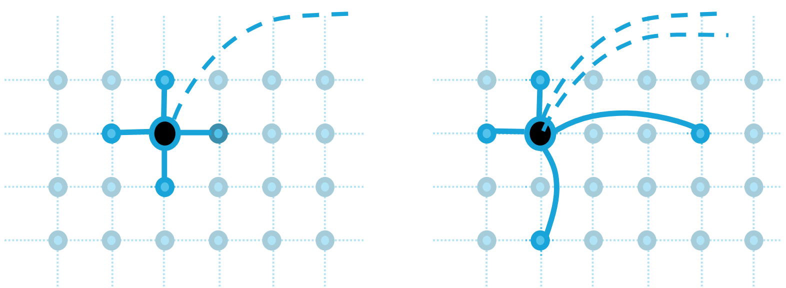
- Assumptions:
  - Social graphs are small world, power law
  - Kleinberg
- Approach:
  - Embed nodes into *Kleinberg-like topology* (namespace: [0,1))
  - Simulated annealing to *approximate lattice* with additional long-range neighbor L_u for each node *u:* $P(L_u = v) \propto \frac{1}{d(u,v)^d}$
    - Periodic random sampling of node pairs
    - Comparison of neighborhoods: $c(u,v) = \frac{\prod_{i \in N(u)} d(ID(u),ID(i)) \ \prod_{i \in N(v)} d(ID(v),ID(j))}{\prod_{i \in N(u)} d(ID(v),ID(i)) \ \prod_{i \in N(v)} d(ID(u),ID(j))}$
    - ID swap with probability: $min\{1,c(u,v)\}$
  - Embedding not greedy, adapted routing (DDFS)





Routing: Extending Kleinberg's Model

- Observe: *Perfect lattice not achieved*
- Extend Kleinberg:
  - Max. distance to closest neighbor ≠ 1
  - Multitude of long range neighbors

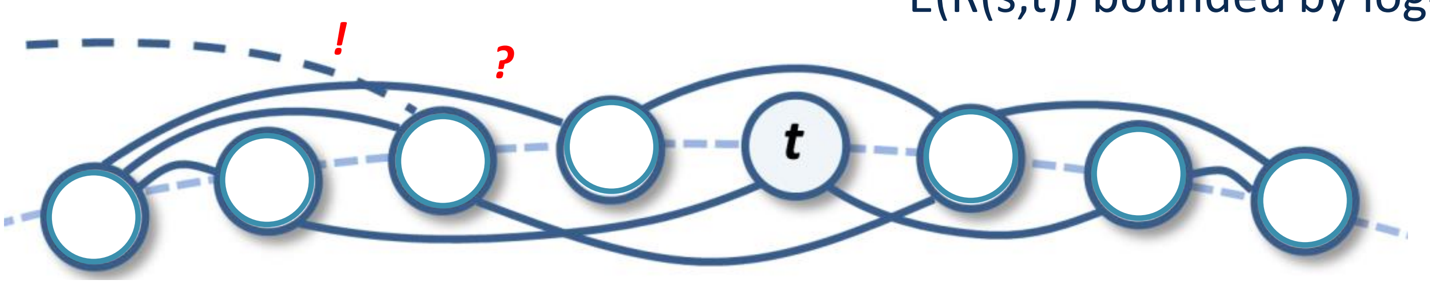


  - $K'(n,d,C,L)$
    - n^d nodes in d dimensional lattice
    - $C \in \mathbb{N}$: max of distance to closest neighbor over all nodes
    - *L:* distribution of long-range links

Routing: Freenet not polylog

- Routing: *Distance-directed depth first search*
  - Forward to neighbor closest to t *that has not received the message before*
  - Backtrack when no neighbor left
  - „On backtrack“: *next closest* neighbor
- *„Try best node that has not received the message before…“*

- *Proof idea* ($C>2$, bounded L):
  1. Adverse scenario: local routing unsuccessful, long range link taken
  2. Success only on backtrack or other long-range link
  3. P_1 linear, P_2 in polylog steps negligible
- Result:
  - E(R(s,t)) bounded by $\log^{\rho} n$

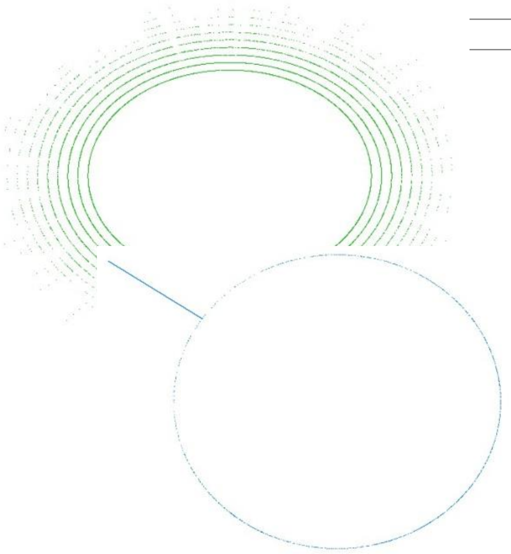


Embedding: Attacking Freenet

- Vulnerabilities: Unattested
  - Request period, source of random walk, TTL
  - ID, neighborhood (arbitrarily bad)

- Ad-hoc attacks:
  - Randomize (all IDs constantly)
    - Pretend having random ID, distant neighbors
  - Contract (all to target ID)
    - Pretend having target ID, distant neighbors

- Simulate
  - 10k users
  - 1% adversaries



single adversary

- Results:
  - Hit Ratio

Attack Type	Immediate attack		Attack after convergence	
	R	H	R	H
Randomize	24%	21%	32%	22%
Contract	27%	22%	32%	31%

No adversary: 60%

random embedding: 21%

Network Embedding

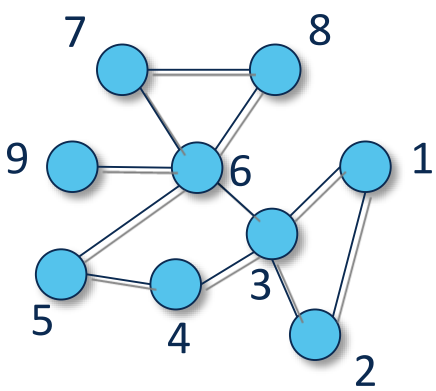
A ***network embedding*** on an undirected graph $G = (V, E)$ is a function

$$ID : V \rightarrow M$$

to a metric space M equipped with a distance

$$d : M \times M \rightarrow \mathbb{R}+ .$$

For a node u ∈ V, *ID(u)* is the identifier of *u*.



- ***Greedy embeddings***

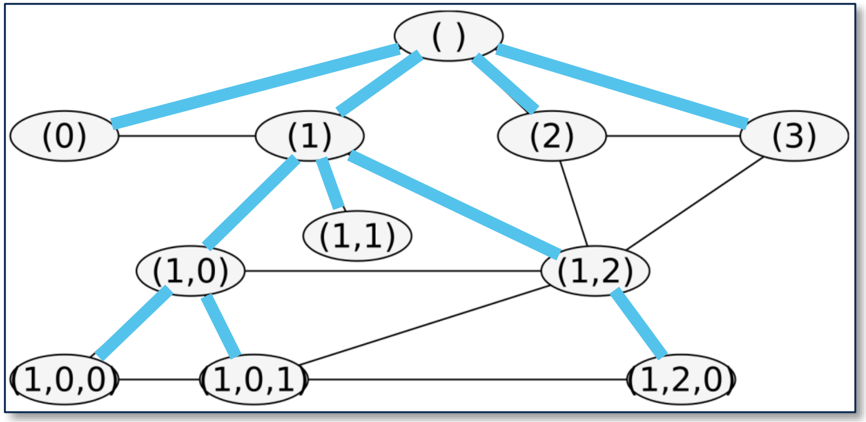
  guarantee greedy routing success (for every distinct node pair *s,t*: *s* is connected to or has a neighbor that is closer to *t*).

- ***Goal:***

  find a decentralized algorithm that approximates a greedy network embedding

Enhancing Freenet`s Embedding

- Distortion extends paths
- Aim: greedy embedding
- Trees can be embedded
- PIE tree embedding
- Find spanning tree
- Enumerate children
- Distance metric:
- d(s,t) := |s| + |t| – 2cpl(s,t)



- Challenges:
  - Tree addresses
    - Leak neighborhood
    - Addresses leak receiver
  - Attacks on tree construction

Recovering Sender and Receiver Anonymity

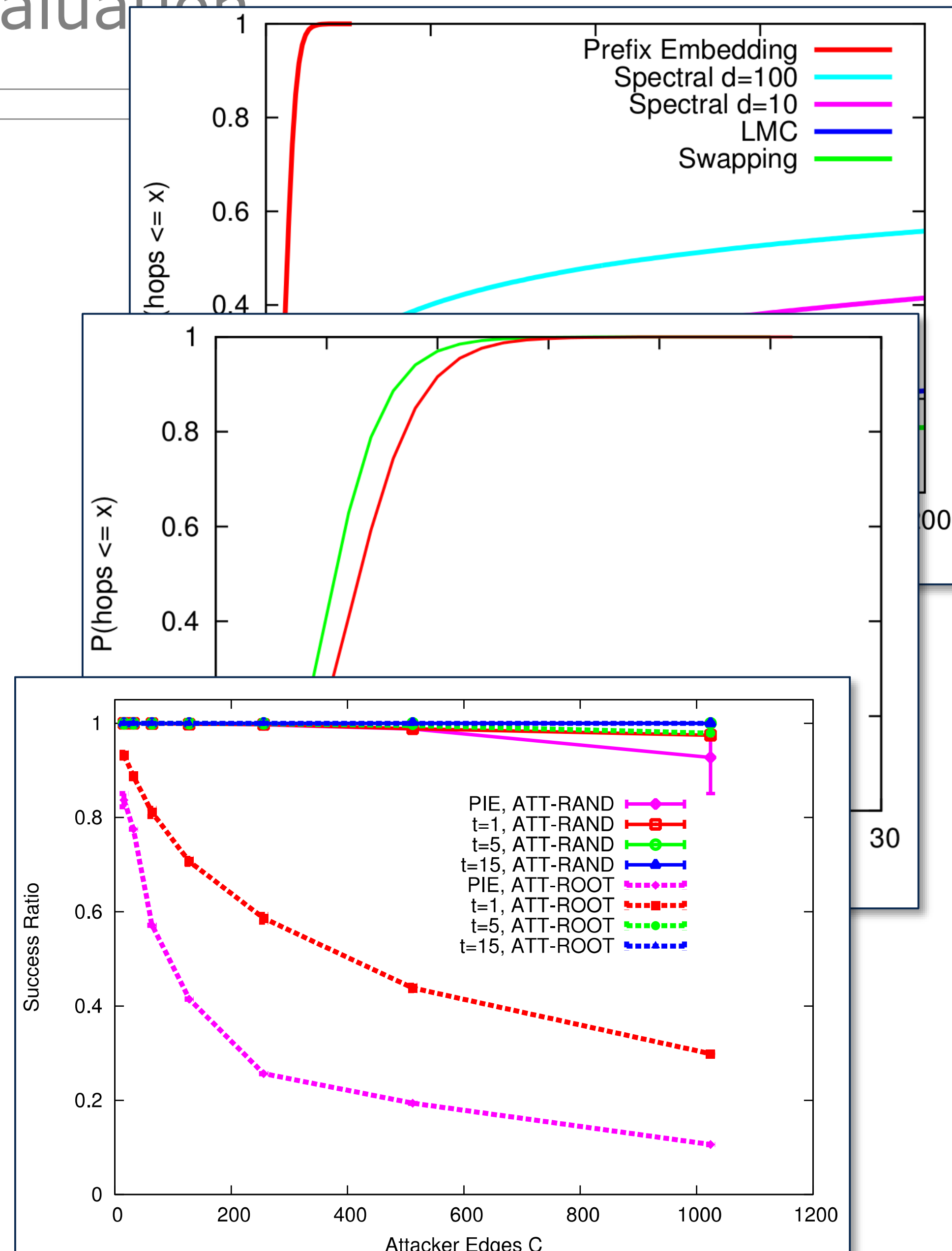
- Receiver anonymity
  - (Return) address needed
  - Distance: longest prefix match
  - Blinded addresses:
  - Randomize:
    - [1,2,0] -> [r1,r2,r3]
  - Padding
    - [r1,r2,r3] -> [r1,r2,r3, rk+1, … ,rL]
  - Blinding
    - k, [r1,..,rL] -> (k,[h(r1⊕k),h(r2⊕h(r1⊕k)…])
  - Distance metrics:
- d1 (s,t) := |s|+|t|−2cpl(s,t)
- d2 (s,t) := L - cpl(s,t) - δ

- Theoretical analysis
- Performance bounds
  - Tree routing O(log n)
  - Tree maintenance O(log n)
  - per join/leave
- Security analysis
  - Plausible deniability: Receiver cannot uniquely be identified
  - Minimal information loss to allow for routing

[1] **Roos, Beck, and Strufe: INFOCOM 2016**

Performance Evaluation

- TE is a greedy embedding
- Simulation experiment
  - Topology: PGP Web of Trust
  - Embeddings: Freenet/RW
  - Routing: DDFS/Greedy
- Are we there yet?
- Summary:
  - It‘s robust and fast!
  - Integration under construction
  - Load balanced???



Summary

- What is privacy
- How is it threatened (directly and indirectly)
- What are potential effects
- What can we do about it

References

Cutillo, Leucio Antonio, et al. "Security and privacy in online social networks." Social Network Technologies and Applications. Springer US, 2010.

Günther, Felix, et al. "Cryptographic Treatment of Private User Profiles." In Financial Cryptography and Data Security, RLCPS, 2011

Gürses, Seda and Diaz, Claudia. "Two tales of privacy in online social networks" IEEE Security & Privacy, 2013

Lauber-Rönsberg, Anne: "Research Ethics and Data Protection Laws". Online

Nissenbaum, Helen. "Privacy as Contextual Integrity", Washington Law Review, 2004

Paul, Thomas et al. "Improving the Usability of Privacy Settings in Facebook.", arXiv:1109.6046 [cs.CR]

Paul, Thomas, et al. "C4PS – Helping Facebookers Manage their Privacy Settings.", In SocInfo, 2012

Pfitzmann, Andreas, and Hansen, Marit: "A terminology for talking about privacy by data minimization." Online: https://dud.inf.tu-dresden.de/literatur/Anon_Terminology_v0.34.pdf

Roos, Stefanie, et al. "Anonymous Addresses for Efficient and Resilient Routing in F2F Overlays." In IEEE INFOCOM, 2016

Roos, Stefanie and Strufe, Thorsten. "On the impossibility of efficient self-stabilization in virtual overlays with churn." In IEEE INFOCOM, 2015

Roos, Stefanie, and Strufe, Thorsten. "Dealing with Dead Ends: Efficient Routing in Darknets", In ACM Trans. Model. Perform. Eval. Comput. Syst., Vol. 1, No. 1, 2016.

Scharloth, Joachim. "Research Ethics: Principles and New Challenges". Online: http://scharloth.com/slides/research_ethics/folie_19.html

Schulz, Stephan, and Thorsten Strufe. "d² Deleting Diaspora: Practical attacks for profile discovery and deletion." 2013 IEEE International Conference on Communications (ICC). IEEE, 2013.

Warren, Samuel, Brandeis, Louis. "The Right to Privacy", Harvard Law Review, Vol. IV, No. 5, 1890

All pictures credit wikimedia, unless stated differently