TECHNISCHE UNIVERSITÄT DRESDEN

DRESDEN concept

Professur Privacy and IT Security

Forschungslinie 2017

Thorsten Strufe

Dresden, 22.05.2017

TECHNISCHE UNIVERSITÄT DRESDEN

A little Motivation: The analog World…

...turns digital...

Note, in the analog world of yesterday there's:

- No law to report TV consumption
- No law to subscribe to journals and papers by name
- No law to specify sender on/in a letter
- No law to be identified upon cash payments

The Situation on the World Wide Web

Web traffic is converging to sites of 6 corporations

- Success due to integration and strong personalization
- Data minimization and avoidance in conflict to business modell

Convergence of communication and expression

- Facebook evolves to integrated communication platform with 1.3 Bn users
- Google, g+: 500 Mio User
- Clear name: perfectly identifiable

Increasingly mobile utilization

- Perfect location, easy tracking
- Configuration more tedious

TOP 10 WEB BRANDS BY UNIQUE AUDIENCE (U.S. TOTAL)

Rank	Brand	Unique Audience	Time Per
1	Google	170,629,000	2:05:30
2	Facebook	145,297,000	6:41:44
3	Yahoo!	135,100,000	2:32:52
4	YouTube	124,073,000	1:57:28
5	MSN/WindowsLive/Bing	123,133,000	1:15:40
6	Microsoft	86,986,000	0:47:26
7	Amazon	84,735,000	0:38:14
8	AOL Media Network	83,826,000	2:09:36
9	Wikipedia	76,310,000	0:24:25
10	Ask Network	69,447,000	0:12:30

[Nielsen]

The Providers and the Data at their Hands

Explicit

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

Extracted

- Profiling
- preference models
- image recognition models

Incidental / „metadata“

- Observed:
  - ***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***
- Inferred
  - derived from observations
  - homophily

Externally correlated

- interest/preferences (external clickstreams)

The Stakeholders

Subscribers

Provider

Partner

Advertisers

zynga Extending Partner

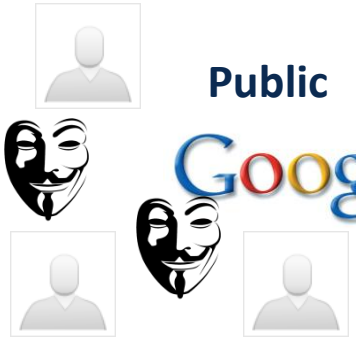
Cloud/CDN Provider

CLOUDFLARE

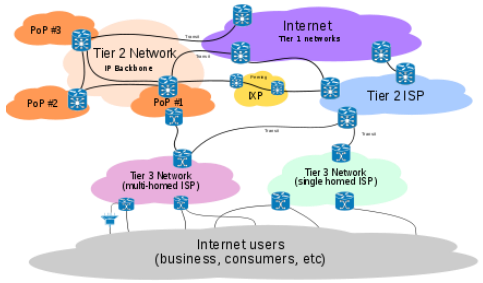
Akamai

Institutions

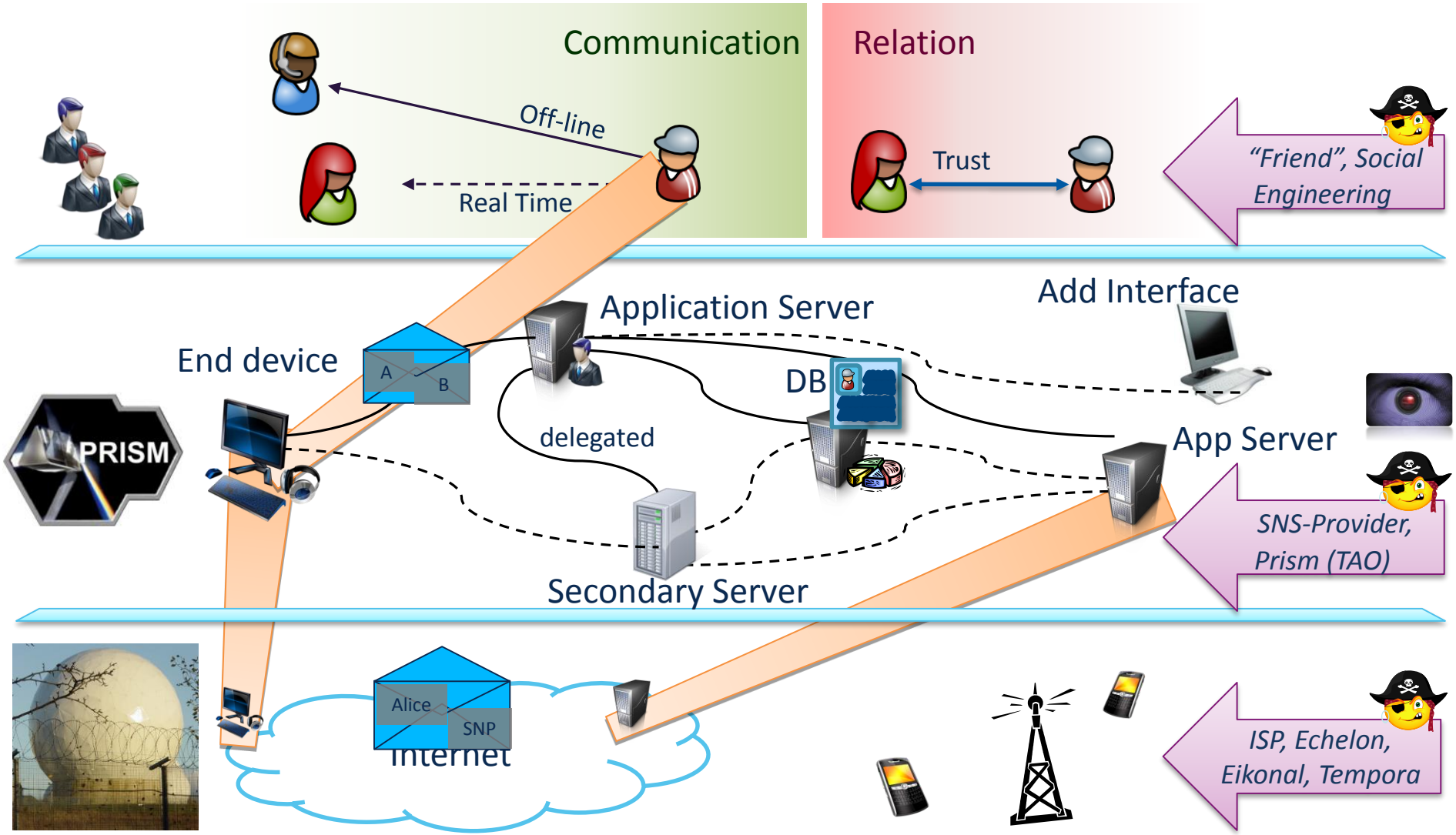
Public



Network Provider



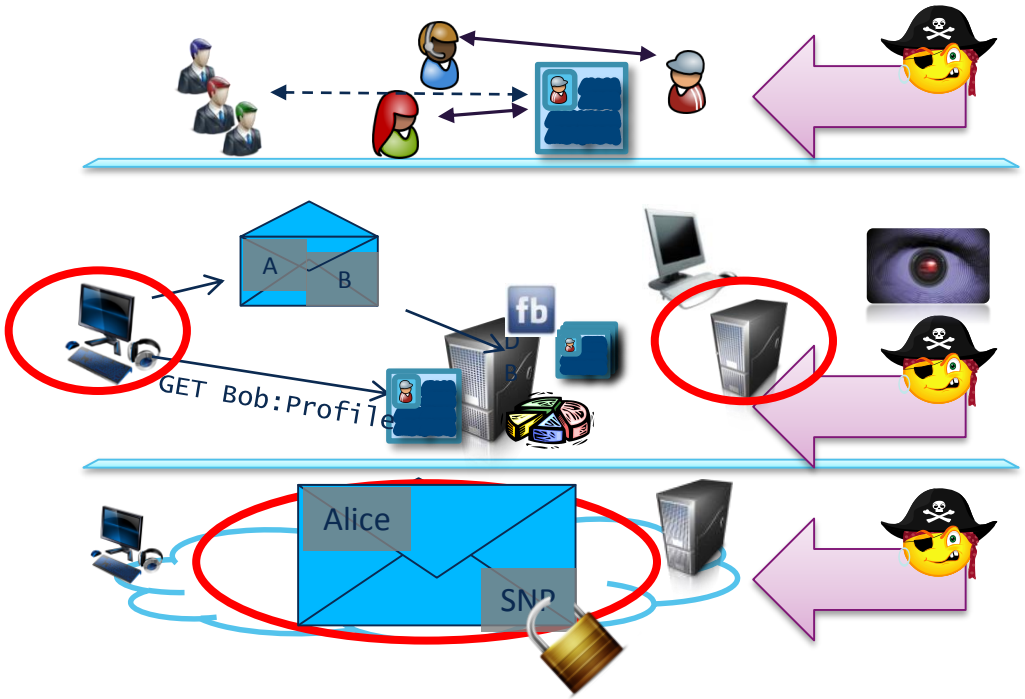
Model and Adversaries



Solution Classes / Research Clusters

- ***Network Security***
  - Protecting the transmission
  - Protecting the network

- ***Surveillance Prevention***
  - Network anonymization
  - Anonymized services



Dezentralization against Censorship

Entire Distribution of Data and Control

- *Decentralize completely*
- *Use explicitly trusted services only*

Common system classes

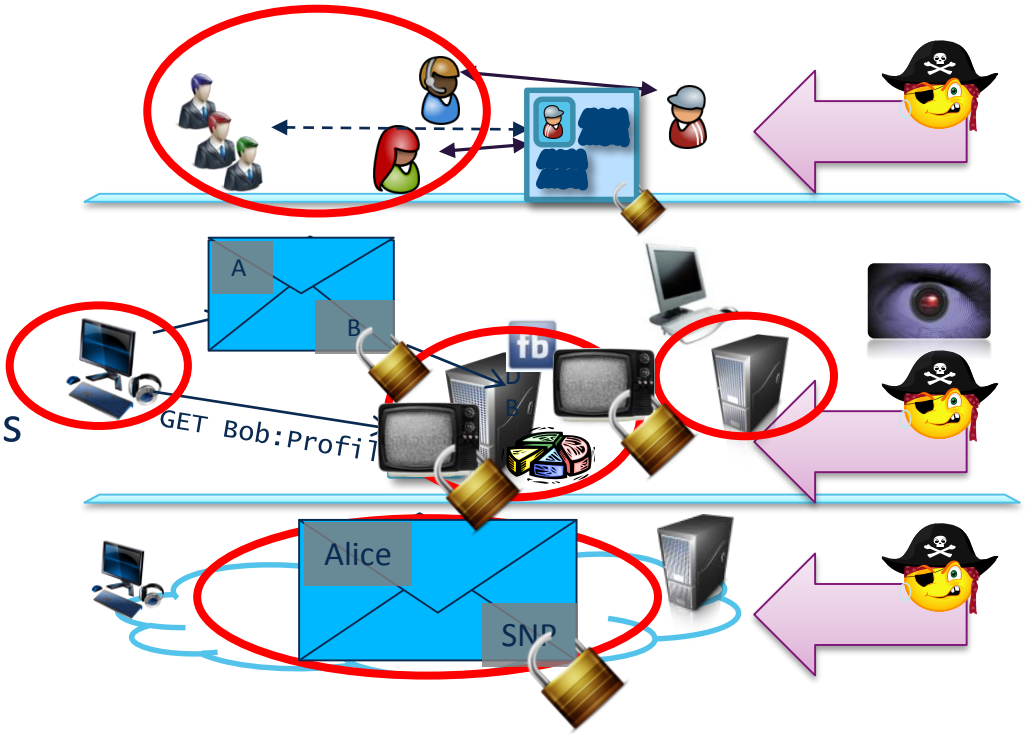
- Federated SNS
- P2P / D-OSN
- Social Overlays and Darknets

diaspora*

PeerSoN

Solution Classes / Research Clusters

- ***Network Security***
  - Protecting the transmission
  - Protecting the network
- ***User/System Understanding***
  - Assessing privacy (inference)
  - Intention recognition
  - Support and useable security
- ***Surveillance Prevention***
  - Network anonymization
  - Anonymized services
- ***Secure Computations***
  - Trusted Execution Environments (Intel SGX)
  - Homomorphic crypto



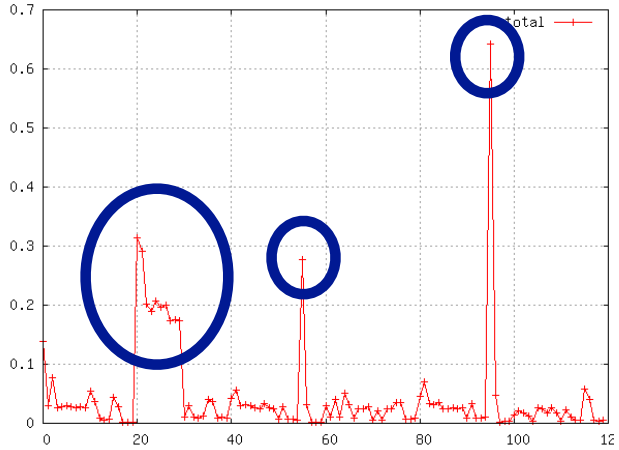
Cluster Data Collection and Analysis: Goals

- Understanding the user
  - Intention recognition
  - Protection of privacy
  - Adaptation of privacy settings
- Social media analysis
  - Social-bot detection
  - Echo chambers and filter bubbles
- User support
  - Phishing prevention/trainings
  - Usable interfaces
- Anomaly detection in dynamic systems

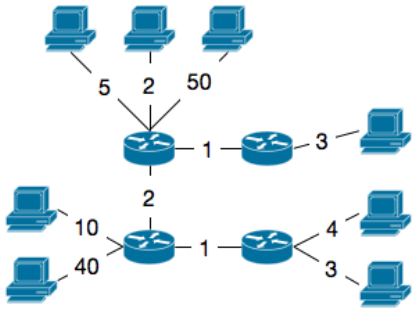
Motivation

SYSTEMS

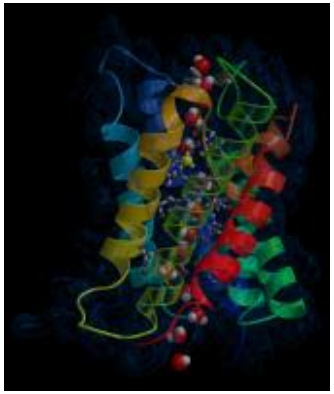
Unterstanding Users



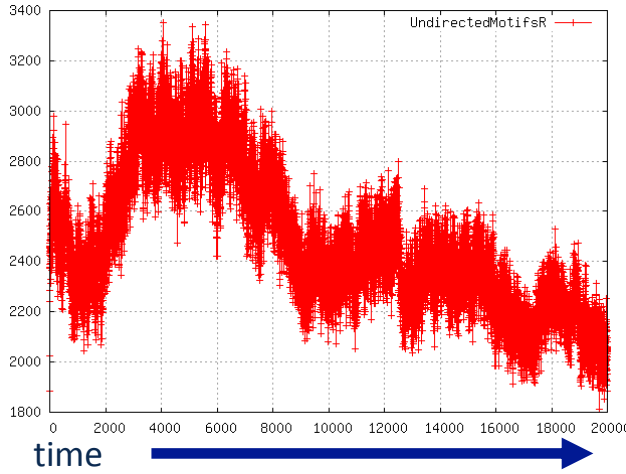
Anomaly Detection



Intention Recognition



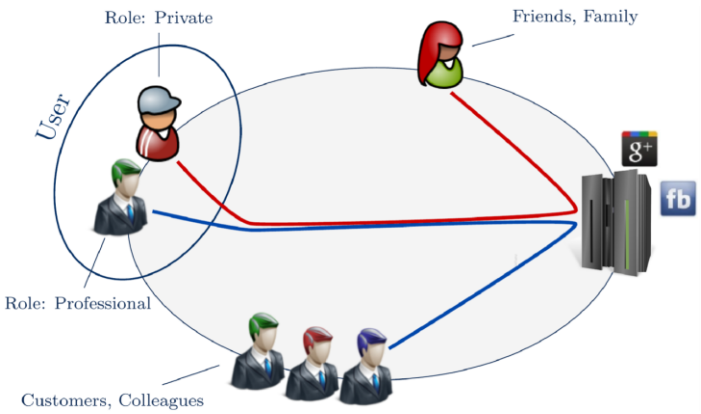
Privacy Protection



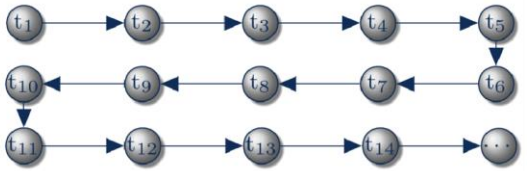
System Analysis

Analysis - Machine Learning

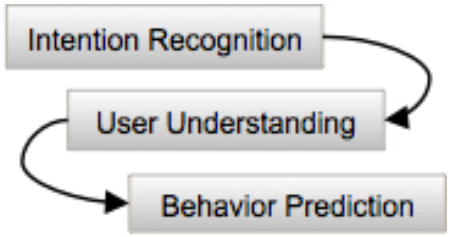
Idea

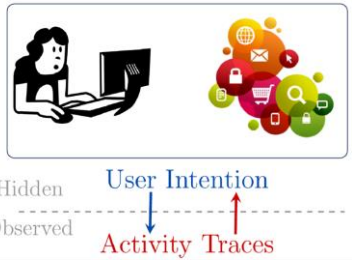


Given

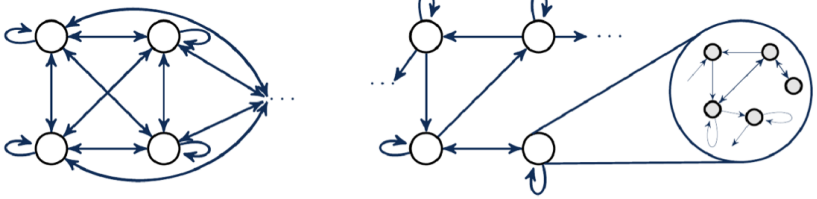


Wanted

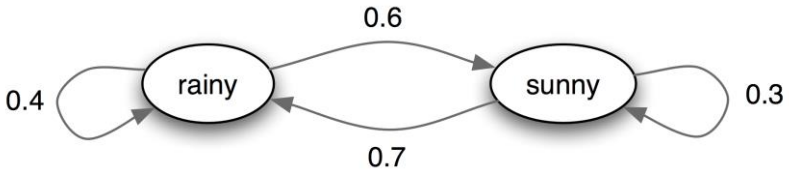




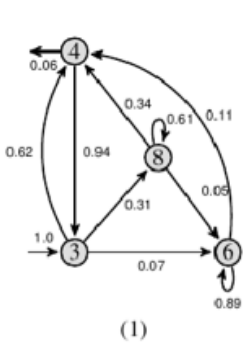
IMMC



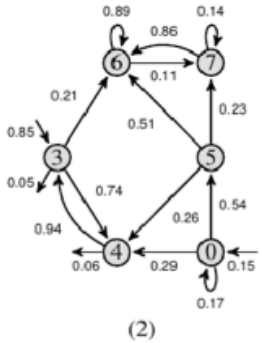
Markov Chains



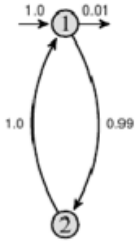
Model | Results



(1)



(2)



(3)

Legend

0:	gotoFacebook
1:	deviceWakeUp
2:	deviceSleep
3:	activateTab
4:	deactivateTab
5:	visitNewsFeed
6:	viewNFEntry
7:	loadMoreNews
8:	sendMessage

Network Security: Motivation

Networks are increasingly targets of attacks

- Internet of things
- Content distribution
- High Performance Computing

Security is essential

- Basis: protection against errors at physical layer (channel coding) – otherwise, security measures are useless …
- Without ensuring confidentiality (C), integrity (I), and availability (A), information is useless
- Threats: eavesdropping (C), Denial of Service attacks (A), pollution attacks (I, A), ...

Security implies costs

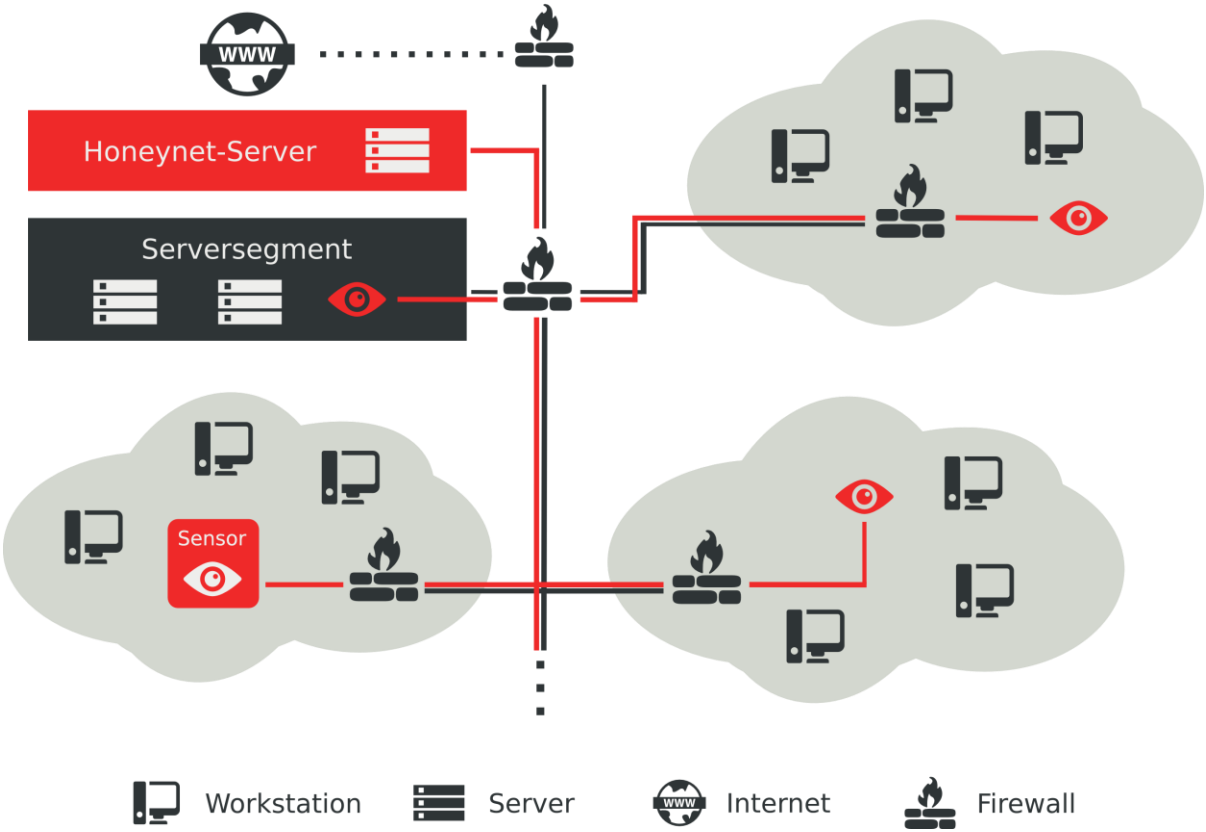
- Computational overhead
- Communication overhead
- Increased delays

→ We need secure solutions that are also efficient!

Network Security: Selected Results

Detection and Mitigation of network-based attacks

- Analysis of network attack vectors and development of defensive architectures, e.g. using honeypots



Cluster Surveillance Prevention

Goals

- Freedom of speech
- Censorship resistance
- Privacy despite 3-letter agencies

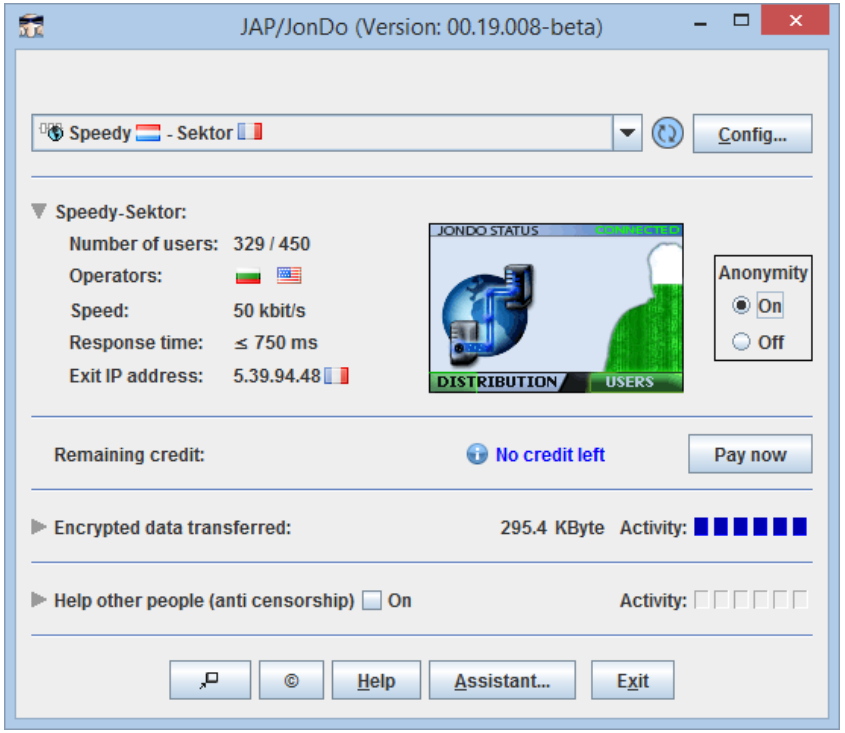
Approaches

- Decentralized service provision
  - Distributed Social Networking
  - Darknets
- Network layer anonymization

The AN.ON Project – ANonymity.ONline since 2000

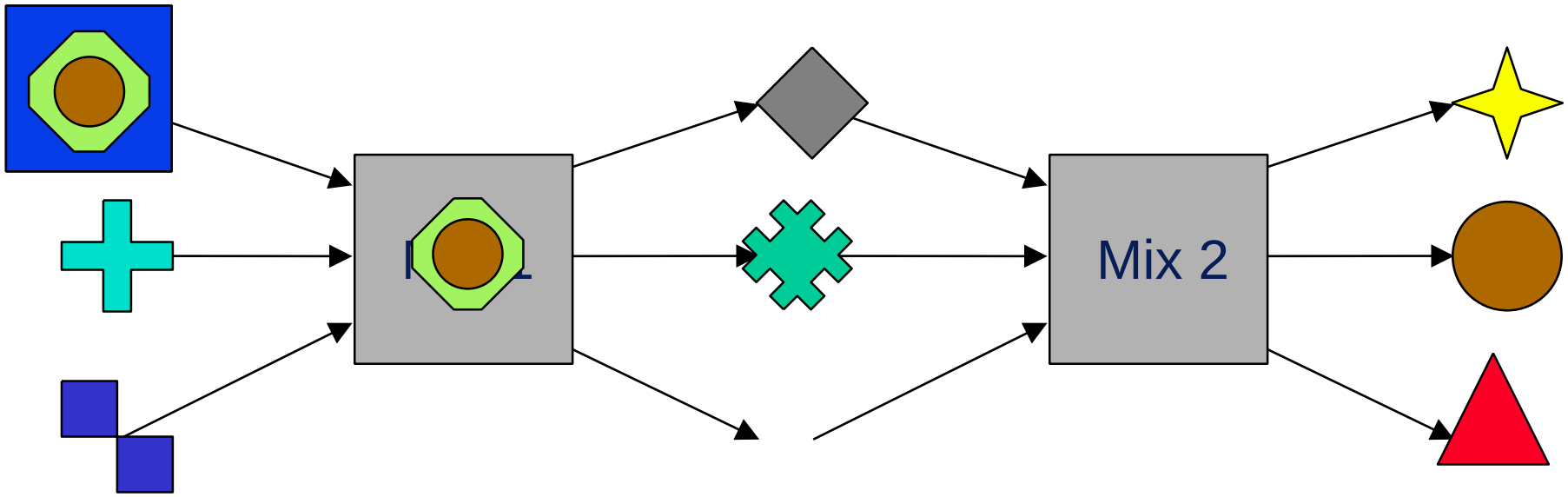
Network Layer Anonymization:

- long track record in the area of “anonymous and unobservable communication”
- holistic view:
  consideration of complex requirements (law enforcement, censorship resistance, etc.)
- since 2001 practical realisation within the project “AN.ON”
- implementation and operation of a anonymisation service based on Mixes



The Mix-technique

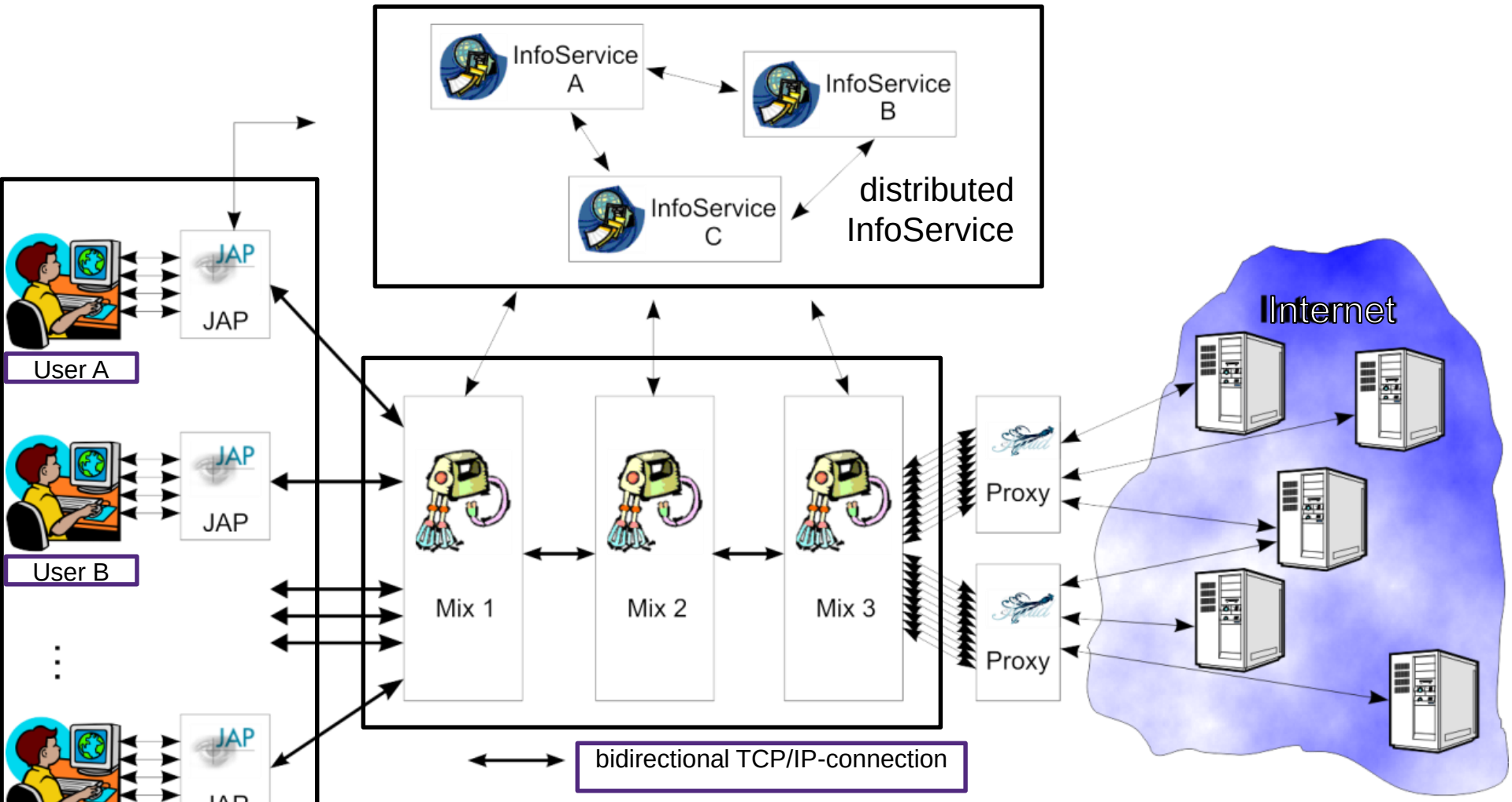
Main Idea: Provide Unlinkability between incoming and outgoing messages !



A Mix samples messages in a batch, changes their coding and forwards them in a different order.

Only if all Mixes work together they can deanonymise a communication relation

Overview of the AN.ON system



Dezentralization against Censorship

Entire Distribution of Data and Control

- *Decentralize completely*
- *Use explicitly trusted services only*

Common system classes

- Federated SNS
- P2P / D-OSN
- Social Overlays and Darknets

diaspora*

PeerSoN

Resilient Social Overlays

Prevent identification, censorship and retribution.

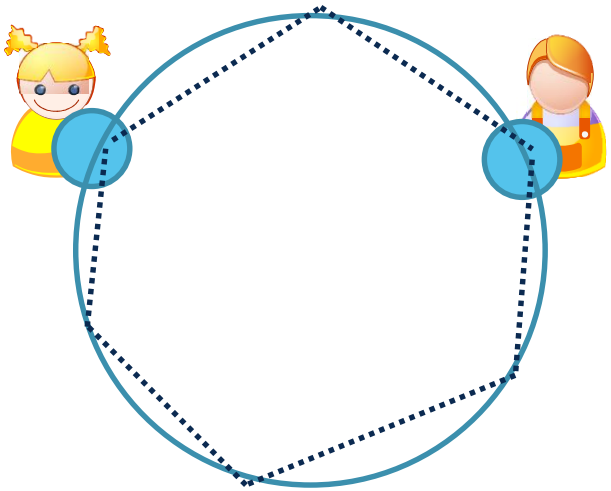
From DOSN to darknets: Tightening requirements

- Concealed participation
- Unobserveability
- Metadata privacy (sender-, receiver-, relationship anonymity)

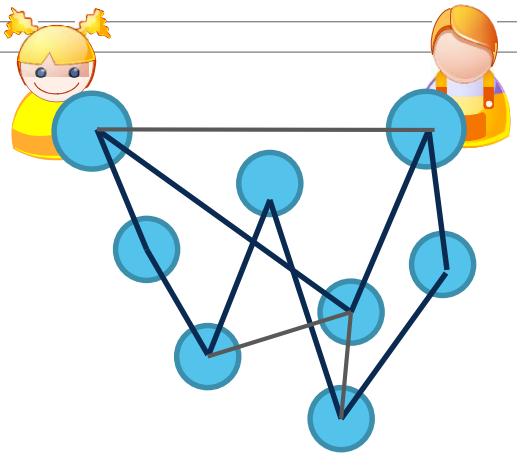
So where's the problem?

Classic overlays:

- Disclosure of IP address
- Eclipse, X-hole attacks



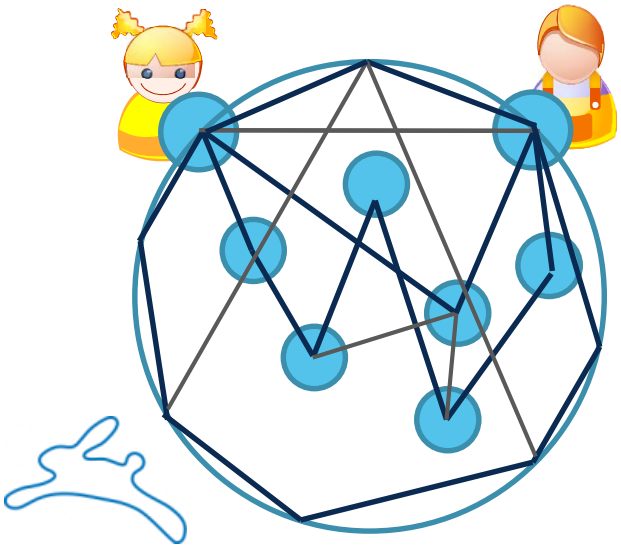
Social Overlays: Embedding/Virtual Overlays



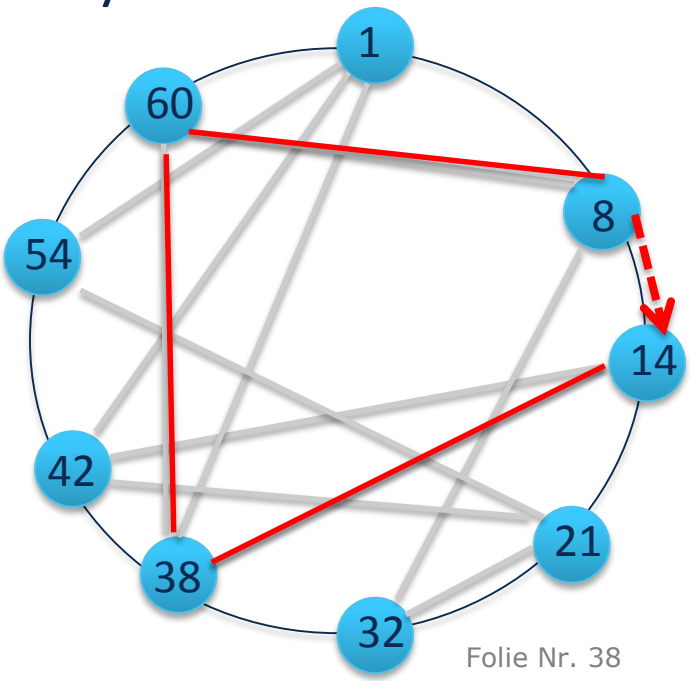
Concepts of social overlays:

- Constrain connectivity to social links
- Contain information
- Attempt to route messages

Embeddings



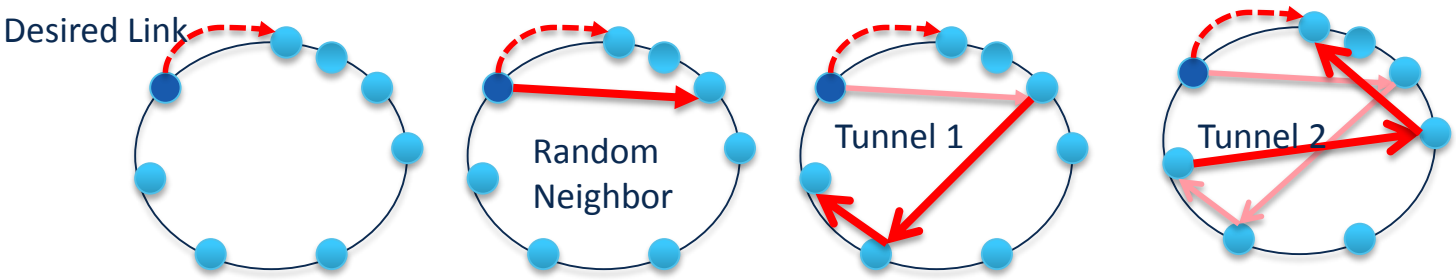
Virtual overlays



Virtual Overlays – Tunnel Maintenance

Establishment & Maintenance of tunnels („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?*

[1] **Roos and Strufe: INFOCOM 2015**

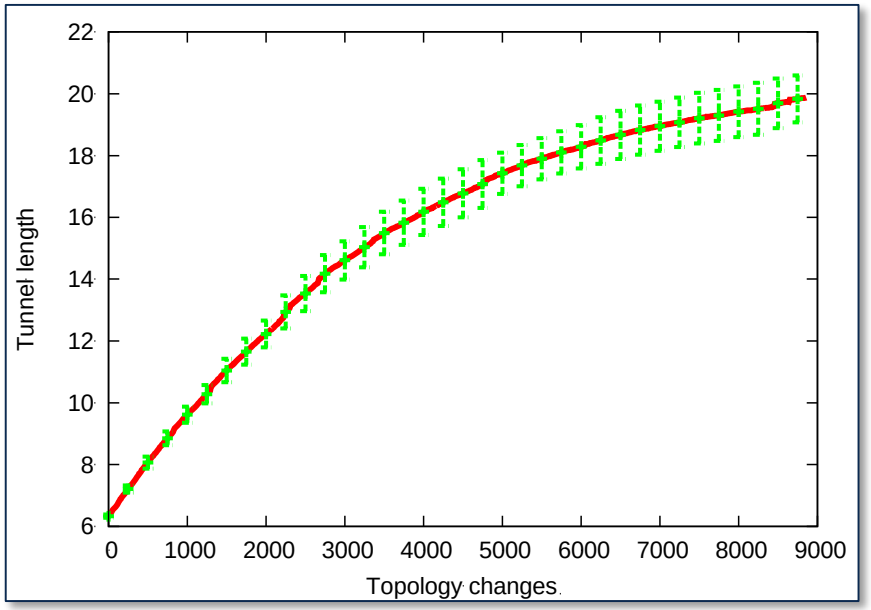
Can Virtual Overlays be Cost-Efficient (Polylog)?

Flooding: no-brainer

Concatenation of trails: Proof by contradiction

1. Model dynamic virtual overlays as a stochastic process
2. Assume polylog stabilization
3. Show tunnel length increases beyond polylog

→ *New trail is longer than removed trail with high probability*



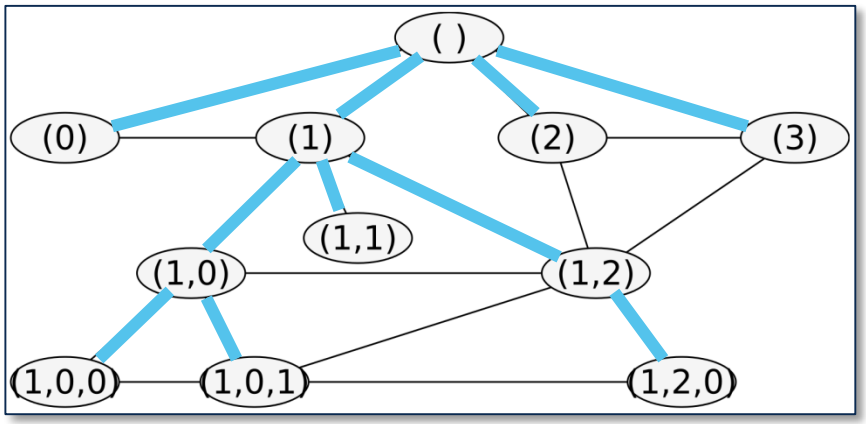
[1] Roos and Strufe: INFOCOM 2015
[2] Roos et al.: PETS 2014

Enhancing Freenet's Embedding

Distortion extends paths

Aim: *greedy embedding*

Trees can be embedded



PIE tree embedding

1. Find spanning tree
2. Enumerate children

Distance metric:

$$d(s,t) := |s| + |t| - 2cpl(s,t)$$

Challenges:

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

Performance Evaluation

TE is a greedy embedding

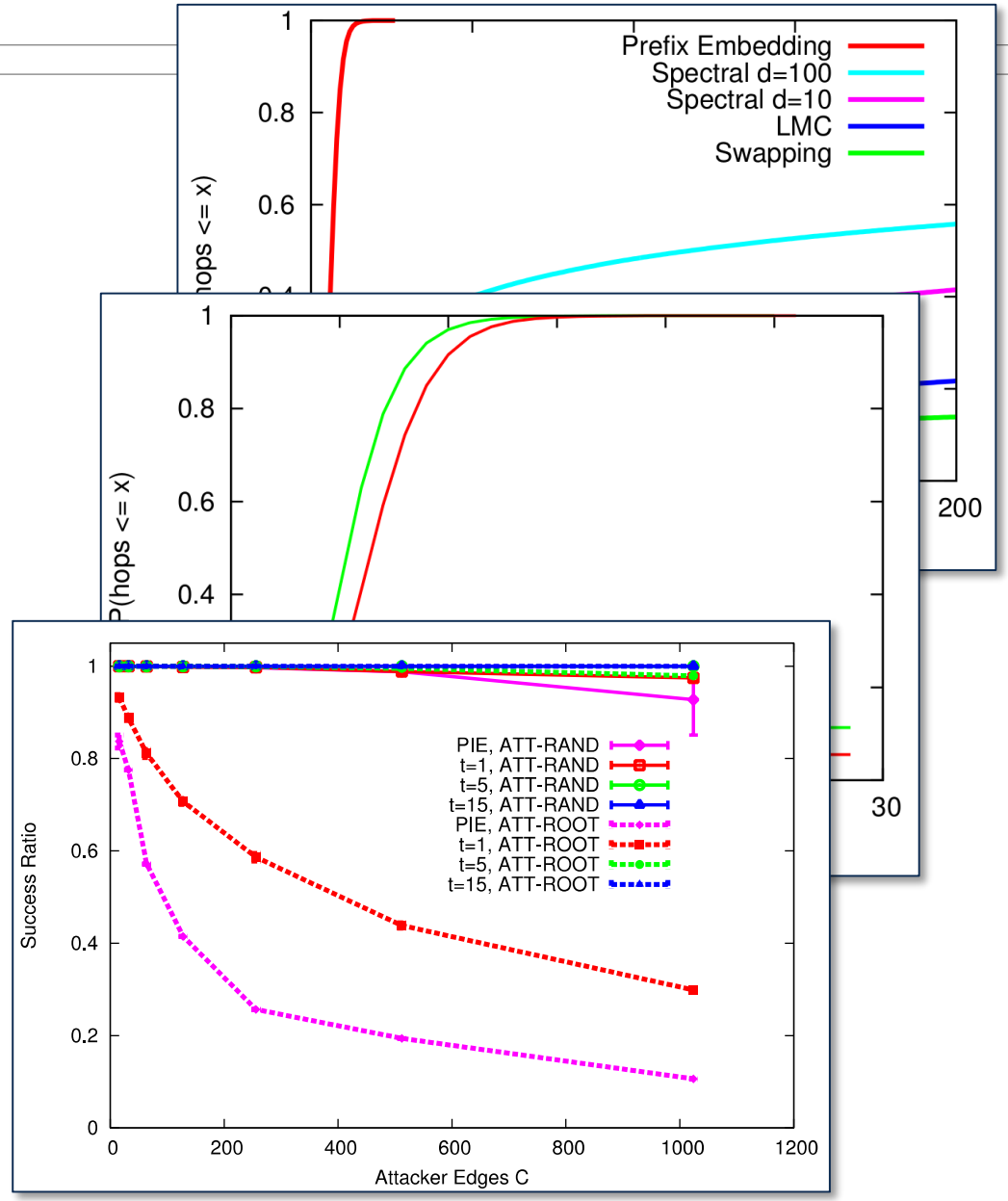
Simulation Experiment

- Topology: PGP Web of Trust
- Embeddings: Freenet/RW
- Routing: DDFS/Greedy

Is it robust?

Summary:

- *It's robust and fast!*



DuD in Your Syllabus

FS	Wintersemester	FS	Sommersemester
1		2	Informations- und Kodierungstheorie
3	Betriebssysteme & Sicherheit	4	*Forschungslinie*
5	**BAS-4** *SaC-1* / Kanalkodierung	6	**BAS-4** *SaC-2/Crypto*
7		8	**Vert-4,** ANW/AFT, Beleg *SaC-2/Crypto/Resilient Networking*
9	**Vert-4**, ANW/AFT *FB-Mining*/Kanalkodierung	10	Diplom/Masterarbeit

BAS-4:

- **S**ecurity **& C**rypto 1
- **S&C** 2 (PETs)
- Crypto
- Kanalkodierung

Vert-4:

- S&C 1&2
- Crypto
- Resilient Networking
- Mining Facebook
- Kanalkodierung

B-510/B-520:

- **S**ecurity **& C**rypto 1
- **S&C 2** (PETs)
- Kanalkodierung
- Seminare/Praktika

FS	Wintersemester	FS	Sommersemester
B1		B2	Informations- und Kodierungstheorie
B3		B4	
B5	**B-510** Betriebssysteme & Sicherheit	B6	**B-520** Bachelor-Thesis
M1	**BAS-4**	M2	**BAS-4, VERT-4, ANW**
M3	**Vert-4, FPA**	M4	Master-Thesis

Thanks for your attention

We‘re looking forward to meeting you!