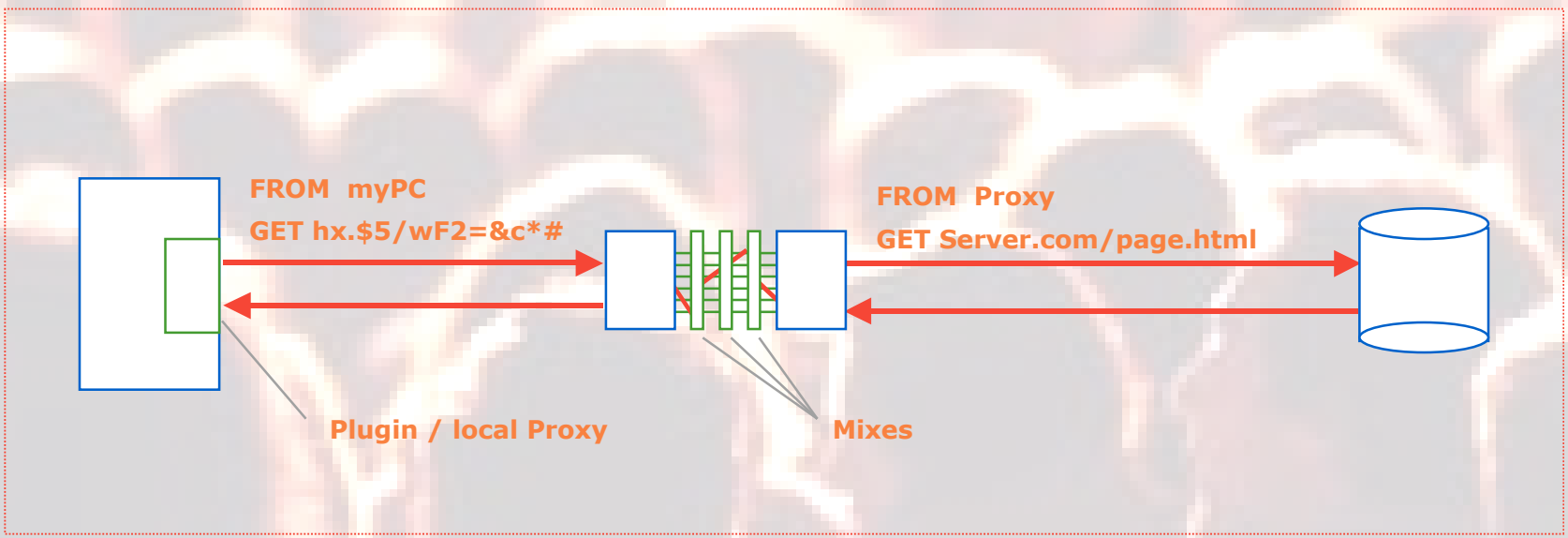
GET http://anon.nowhere.com/
>please type in your name
>set cookie

> Privacy enhancing technologies in the Internet

Hannes Federrath



> Logging and Observation of user actions

Logging of e-mail communication

```
>tail syslog
Oct 15 16:32:06 from=<feder@tcs.inf.tu-dresden.de>, size=1150
Oct 15 16:32:06 to=<hf2@irz.inf.tu-dresden.de>
```

Logging of web access

```
wwwtcs.inf.tu-dresden.de>tail access_log
amadeus.inf.tu-dresden.de - - [15/Oct/1997:11:50:01] "GET
/lvbeschr/winter/TechnDS.html HTTP/1.0" - "http://wwwtcs.inf.tu-
dresden.de/IKT/" "Mozilla/3.01 (X11; I; SunOS 5.5.1 sun4u)"
```

Linkage of user actions

```
ithif19 logs 17 >finger @amadeus.inf.tu-dresden.de
[amadeus.inf.tu-dresden.de]
Login       Name                TTY         Idle    When
feder    Hannes Federrath       console             Wed 11:56
```

> Logging and Observation of user actions

Log

```
>tai
Oct
Oct
```

Log

```
www
amad
/lv
dres
```

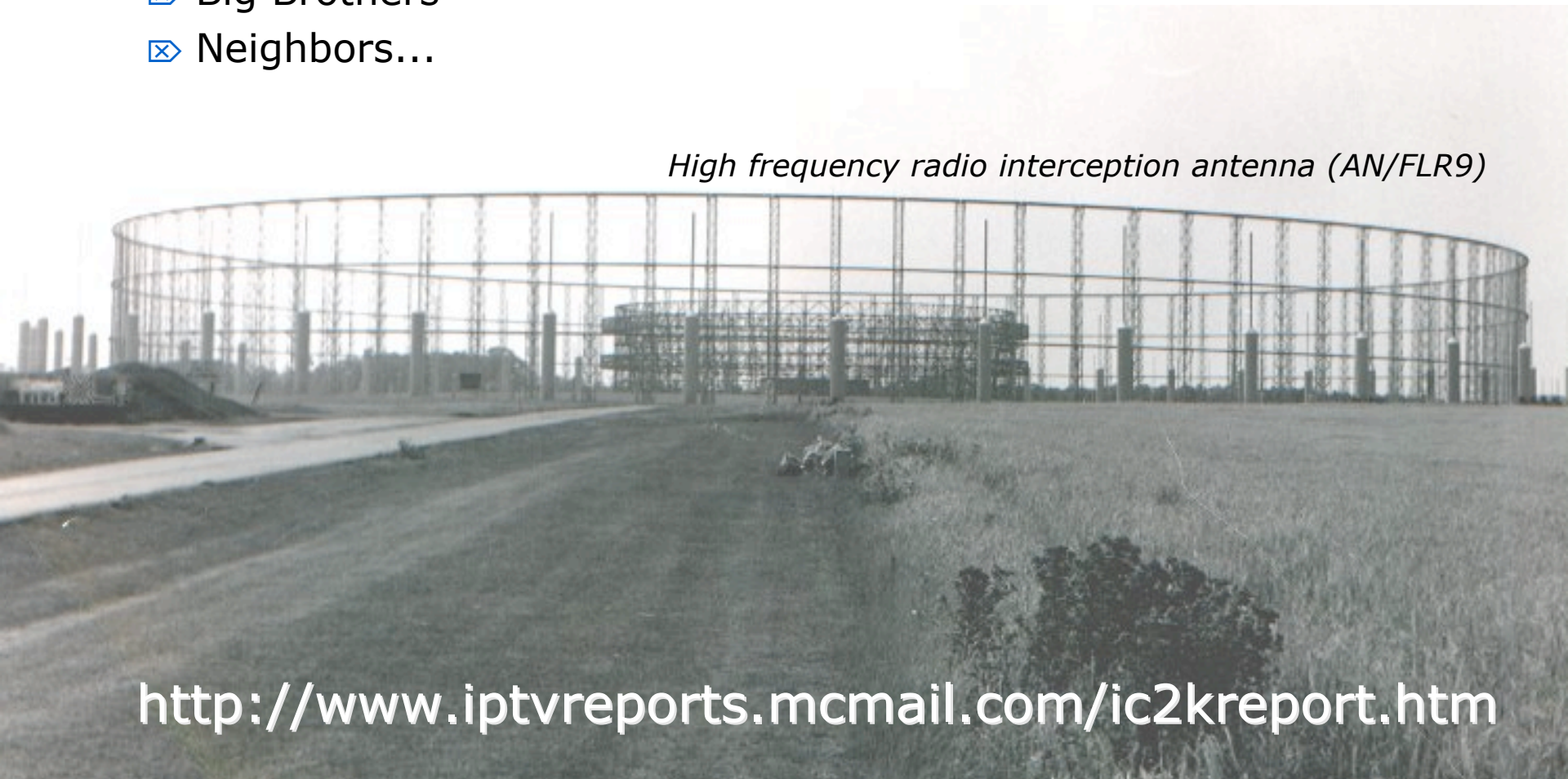
Lin

```
ithi
[ama
Logi
fede
```

(213.68.175.4)	...	Heute 16:17 Uhr
p3e9baca6.dip.t-dialin.net (62.155.172.166)	...	Heute 13:41 Uhr
voss.mat.tu-harburg.de (134.28.61.22)	...	Gestern 14:13 Uhr
pec-120-252.tnt10.me2.uunet.de (149.225.120.252)	...	Gestern 13:50 Uhr
(212.100.36.50)	...	Gestern 13:35 Uhr
gw3.telekom.de (194.94.109.2)	...	Gestern 9:32 Uhr
wzl214.wzl.rwth-aachen.de (137.226.193.214)	...	Gestern 9:09 Uhr
n2-146-189.dhcp.mcphu.edu (144.118.146.189)	...	Gestern 4:04 Uhr
acb08d7f.ipt.aol.com (172.176.141.127)	...	04.06.2001 16:46 Uhr
pd9009416.dip.t-dialin.net (217.0.148.22)	...	04.06.2001 13:44 Uhr
dialppp-7-56.rz.ruhr-uni-bochum.de (134.147.7.56)	...	04.06.2001 8:24 Uhr
(194.64.244.18)	...	03.06.2001 23:19 Uhr
(62.2.58.8)	...	03.06.2001 20:25 Uhr
f-226-182.bielefeld.ipdial.viaginterkom.de (62.180.182.226)	...	03.06.2001 10:30 Uhr

> Anonymity in the Internet is an illusion

- Know your enemy!
  - Competitors
  - Security Agencies of foreign countries
  - Big Brothers
  - Neighbors...



High frequency radio interception antenna (AN/FLR9)

http://www.iptvreports.mcmail.com/ic2kreport.htm

> Anonymity in the Internet is an illusion

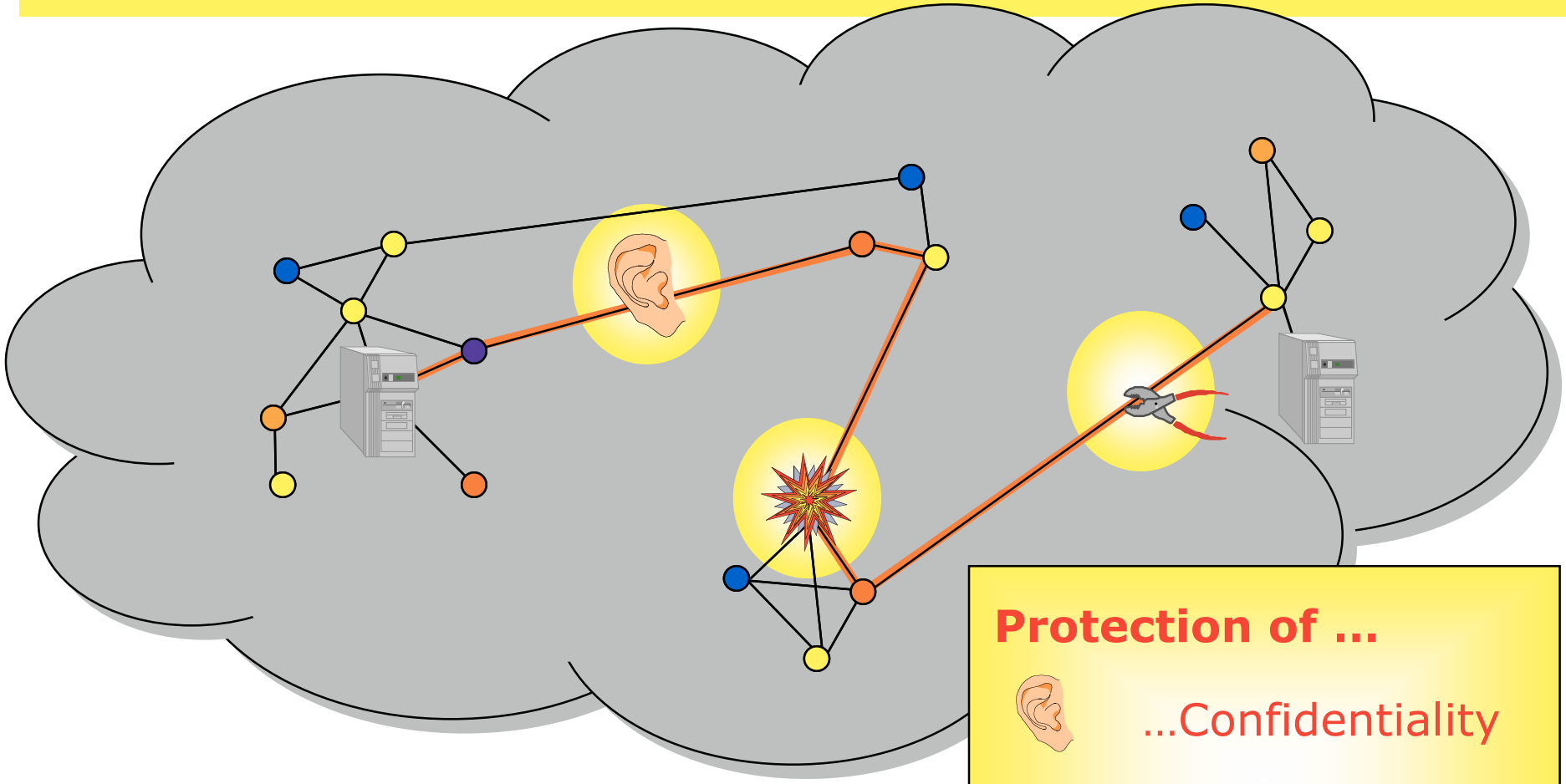
- Know your enemy!
  - Competitors
  - Security Agencies of foreign countries
  - Big Brothers
  - Neighbors...

Bad Aibling Interception facility of the ECHELON system

Source: http://ig.cs.tu-berlin.de/w2000/ir1/referate2/b-1a/

Subject of communication WHAT?	Circumstances of comm. WHEN?, WHERE?, WHO?
Confidentiality **Contents**	**Anonymity** **Unobservablity** **Sender** **Location** **Recipient**
Integrity **Contents**	**Accountability** **Legal Enforcement** **Sender** **Billing** **Recipient**

The Internet



- Telecommunication networks:
  - many operators
  - many users

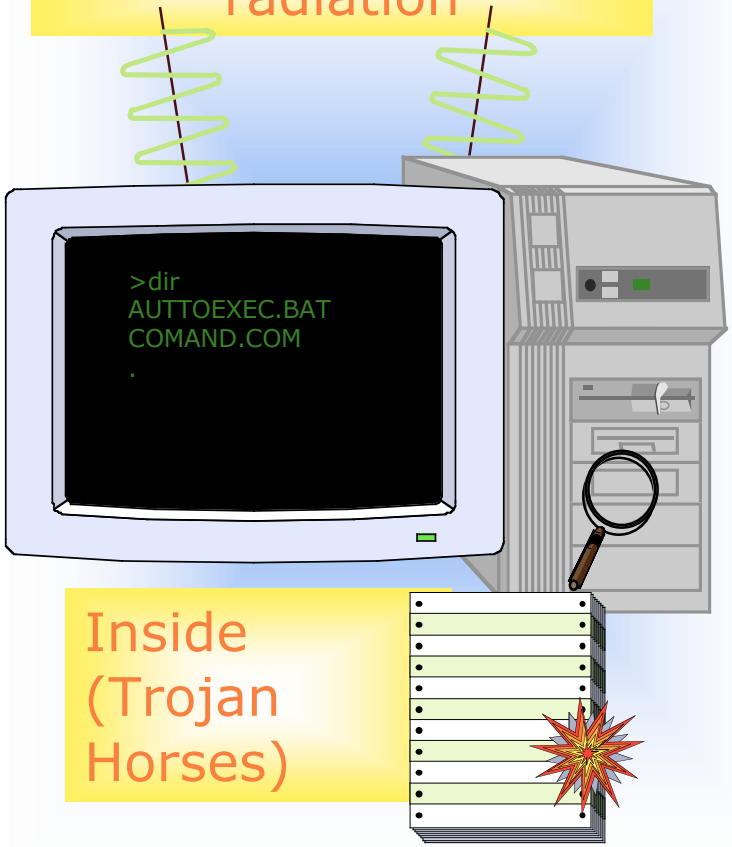
Protection of ...

- ...Confidentiality
- ...Integrity
- ...Avaliablity

> “Access points”

Computer

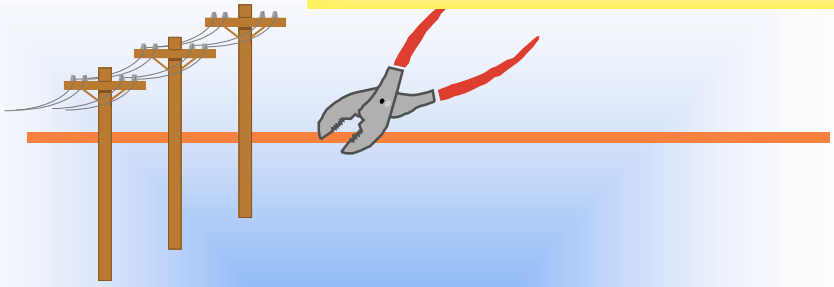
Electromagentic radiation



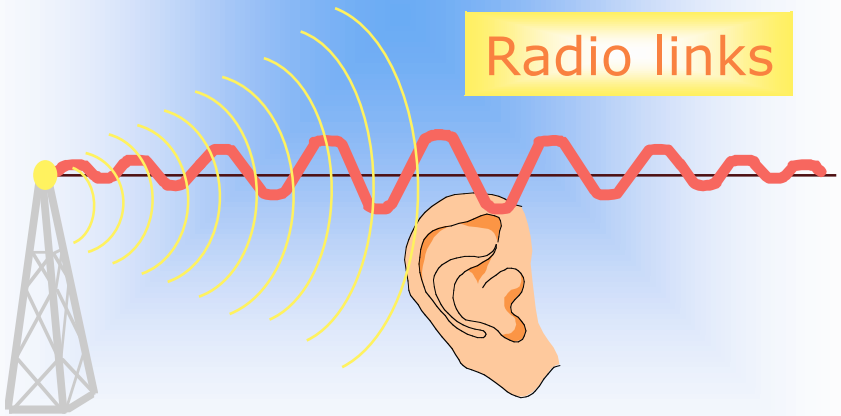
Inside (Trojan Horses)

Transmission

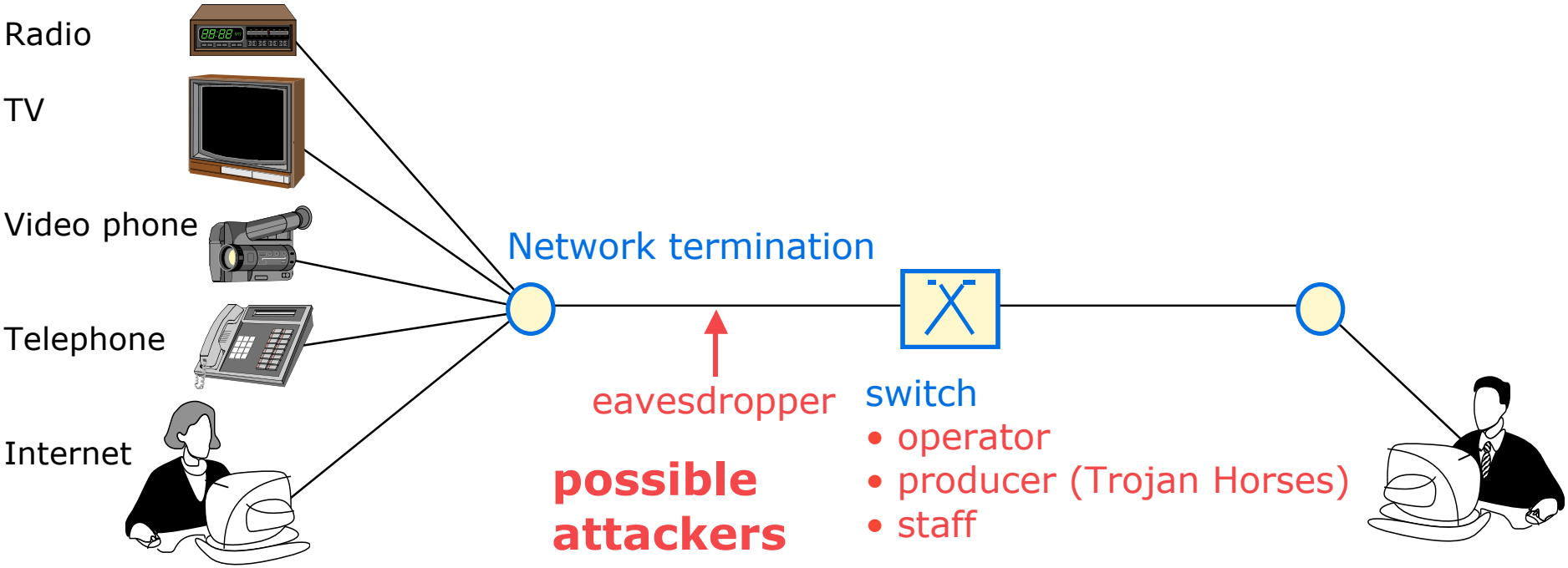
Copper or fiber lines



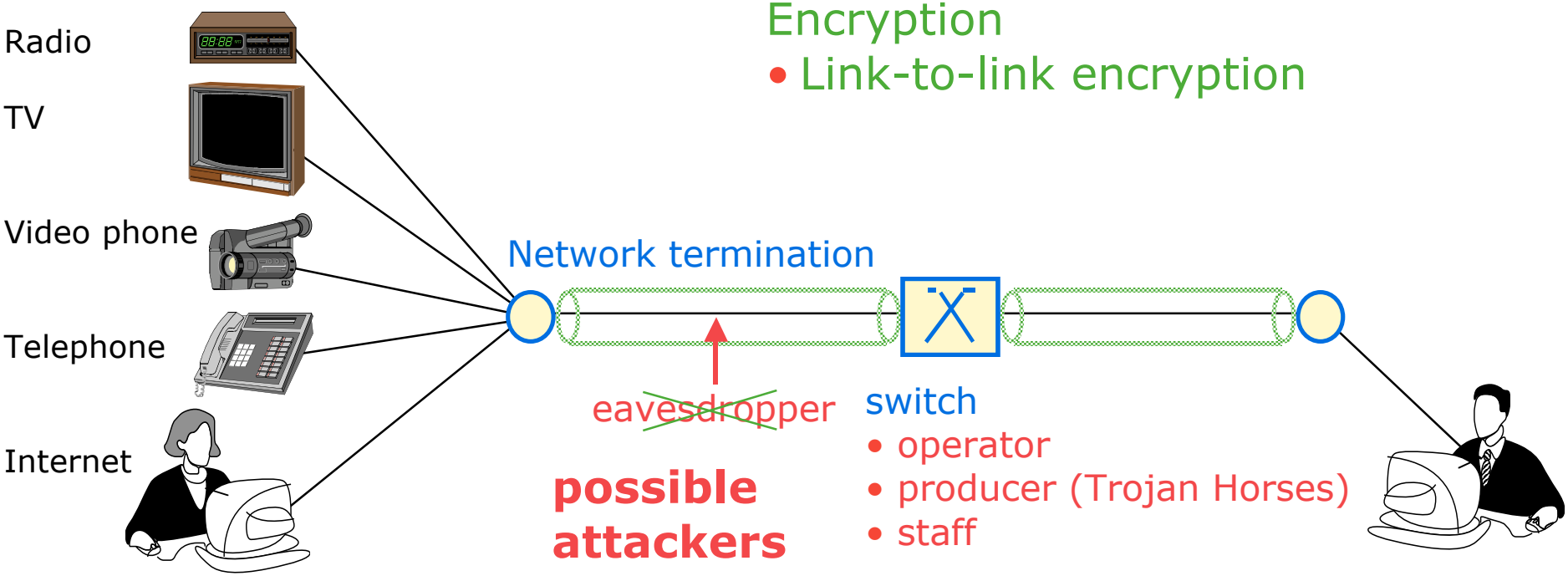
Radio links



> Observation of users in switched networks

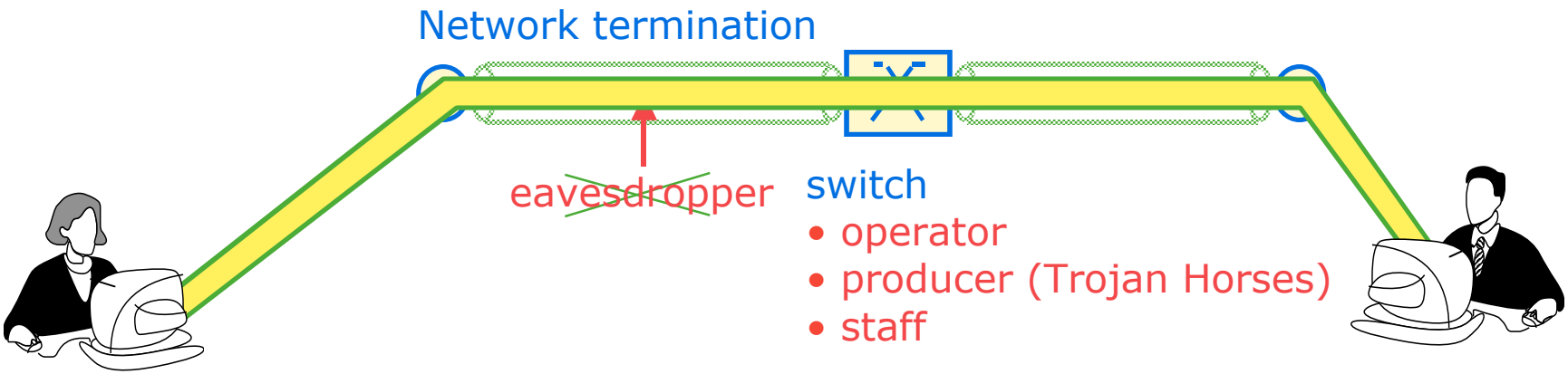


>>Observation of users in switched networks



>>> Observation of users in switched networks

Encryption
- Link-to-link encryption
- End-to-end encryption of contents



Problem – Traffic data:

Who communicates with whom, how long, where?

Who ist interested in which contents?

We need concepts that hide traffic data (or avoid it).

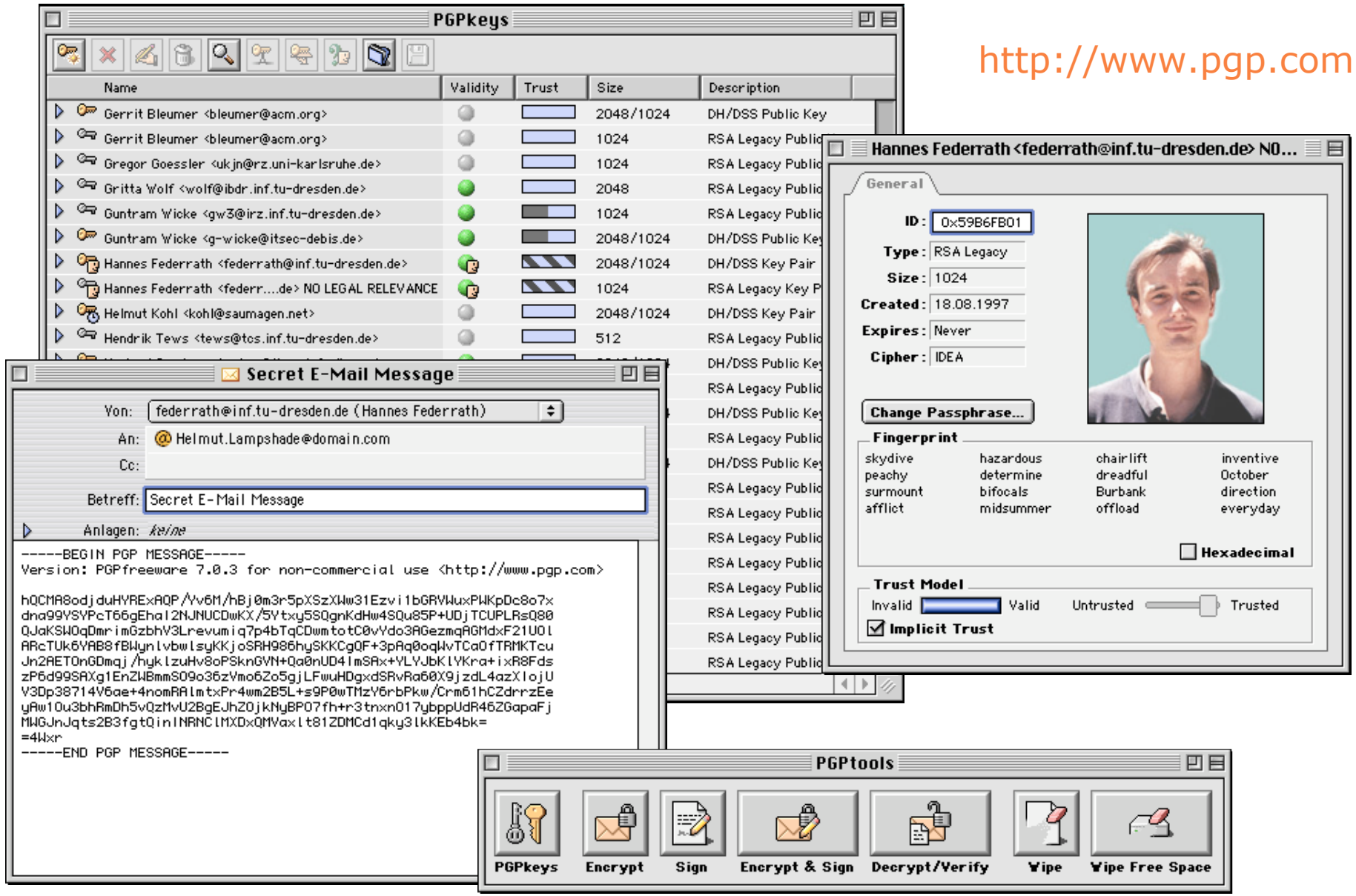
Confidentiality of content by means of Encryption

- Symmetric Encryption, e.g. DES, IDEA, AES
  - Both communication partners share a *secret* key for encryption and decryption
  - Security is based on a „chaos machine“
  - Key length approx 128 bits
- Asymmetric Encryption (Public Key Encryption), e.g. RSA
  - Each user generates a key pair:
    - *public* encryption key
    - *private* (and secret) decryption key
  - Security is based on hard problems in number theory
  - Key length > 1024 bits
    new: elliptic curve cryptography approx. 160 bits

- Well-known encryption software:
  - Pretty Good Privacy
  - http://www.pgp.com

> Pretty Good Privacy (PGP)

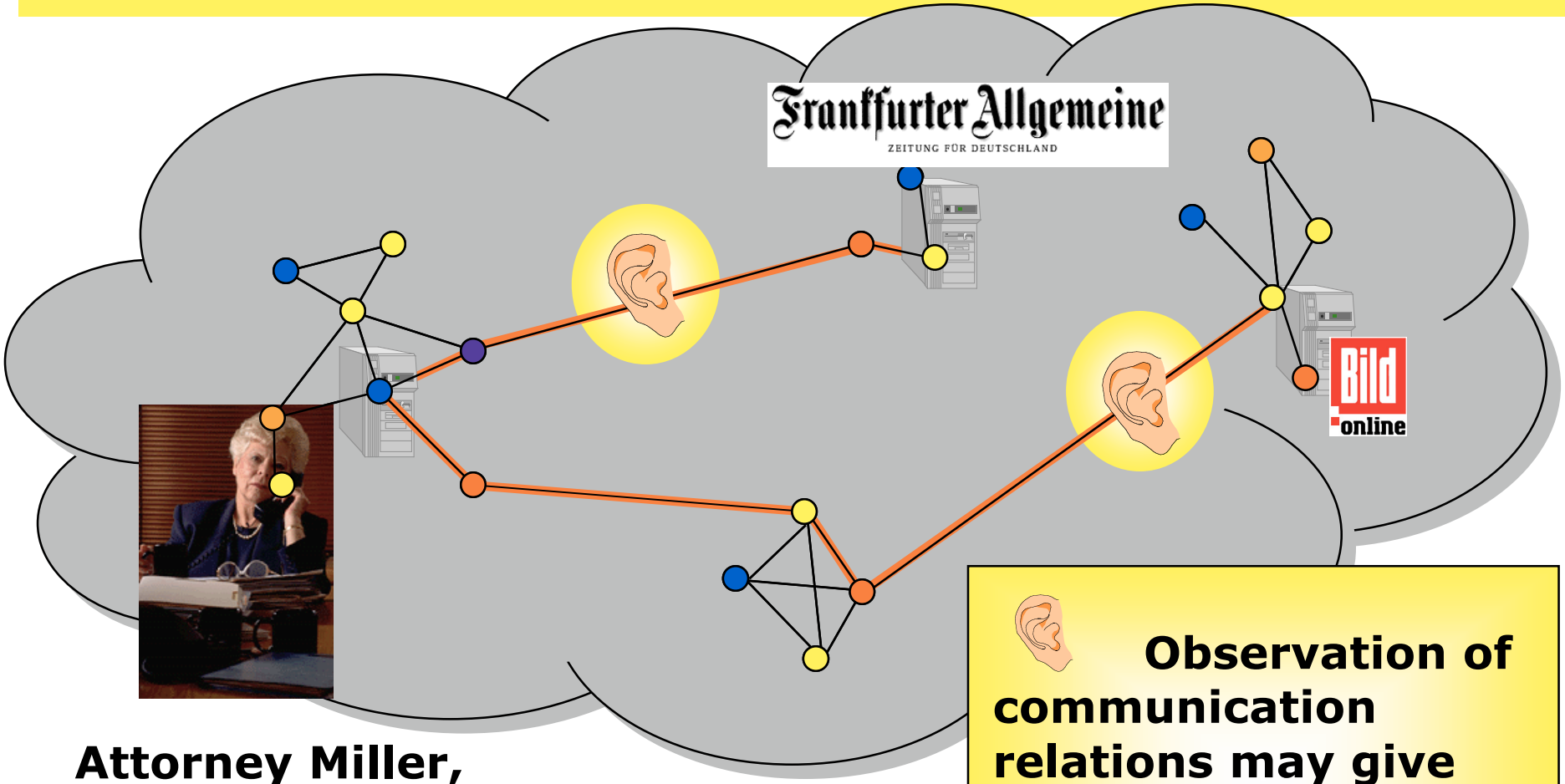
http://www.pgp.com



> Protection against observation?

- New challenges:
  - Privacy in the Internet:
  - Protection against "Profiling" and commercial use of private data without consent.
- Part of Privacy; here: confidentiality of traffic data
- Encryption does not help against observation
  - Who is communicating with whom?
- **Anonymity:**
  - The sender and/or recipient stay anonymous to each other.
- **Unobservability:**
  - All parties (incl. network operators) cannot trace communication relations.
  - Sending and/or receiving of messages is unobservable
- Remarks:
  - A single event caused by a single user cannot be anonymous or unobservable.
  - We need a group of users where all users behave similarly.

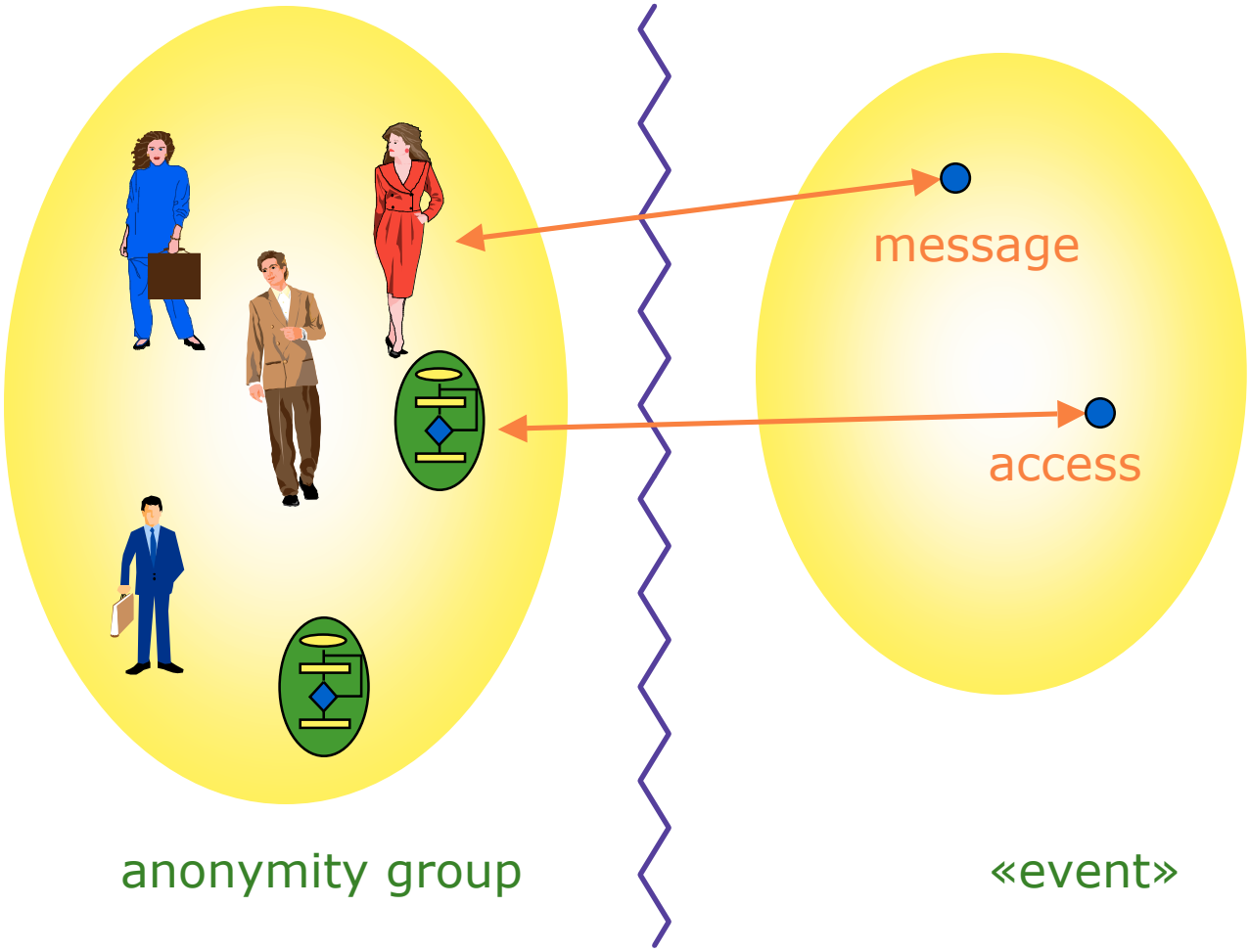
> Why encryption is not enough



Attorney Miller, specialized in mergers

Observation of communication relations may give information about contents

> Anonymity and unobservability

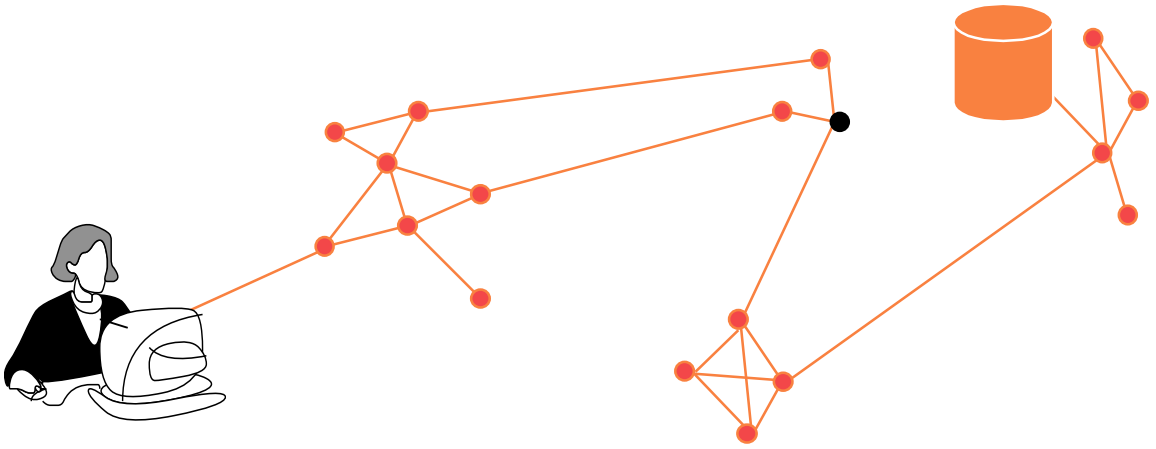


Everybody can be the originator of an «event» with an equal likelyhood

> Our attacker model

- Attacker may:
  - observe all communication links,
  - send own messages,
  - operate anonymity services (all but one ...)
  - operate a server (web server)
- Attacker cannot:
  - break into cryptographic systems,
  - attack the users personal machine,
  - has limited time and computing power

Assuming a very strong attacker is the best way to achieve real security.



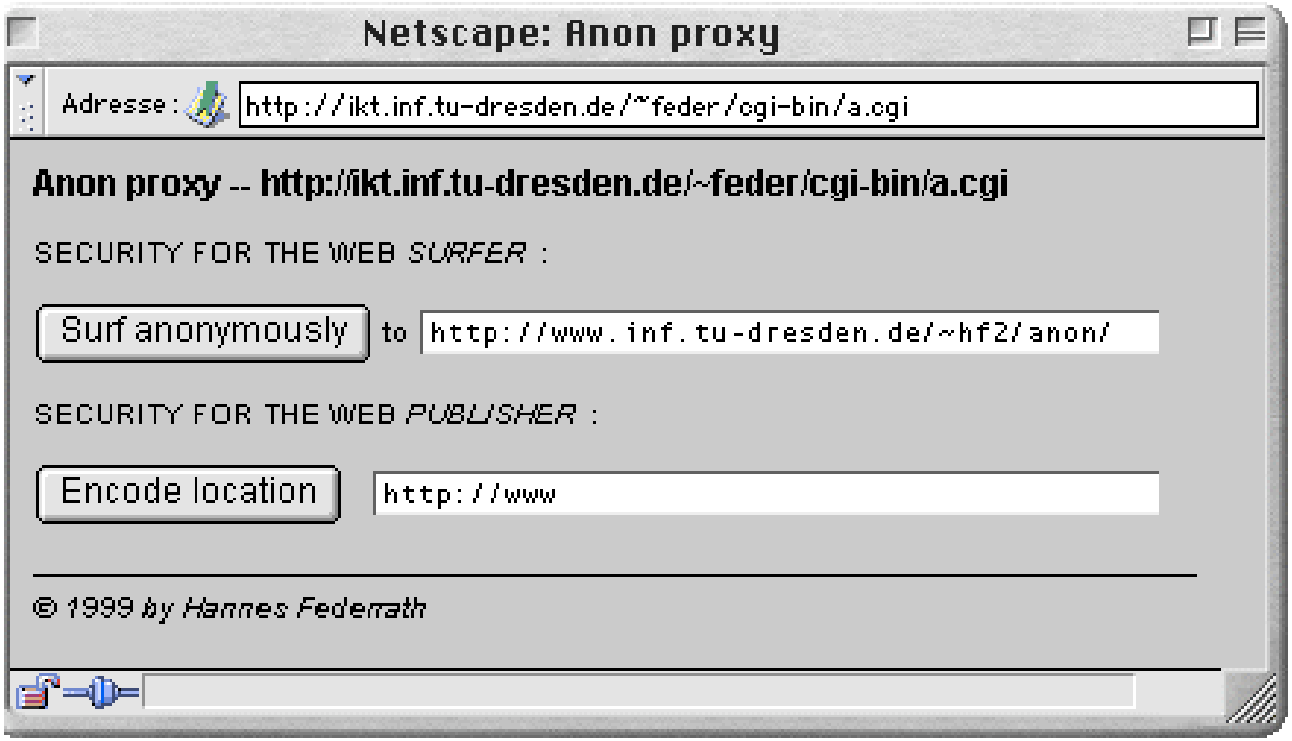
Existing systems for HTTP (real-time communication)

- Simple Proxies (partly with filtering functions: Cookies, JavaScript, active content)
  - Anonymizer.com (Lance Cottrel)
  - Aixs.net
  - ProxyMate.com (Lucent Personal Web Assistant, Bell Labs)
  - Rewebber.com (Andreas Rieke, Thomas Demuth, FernUni Hagen)
  - Anon proxy (Hannes Federrath)
  - Each appropriate configured web server with proxy functions
- Systems considering traffic analysis
  - Crowds (Mike Reiter, AT&T)
  - Onion-Routing (Naval Research Center)
  - Freedom (Ian Goldberg, Zero-Knowledge Inc.)
  - WebIncognito (Privada)
  - WebMixes (TU Dresden)

> Simple Proxies

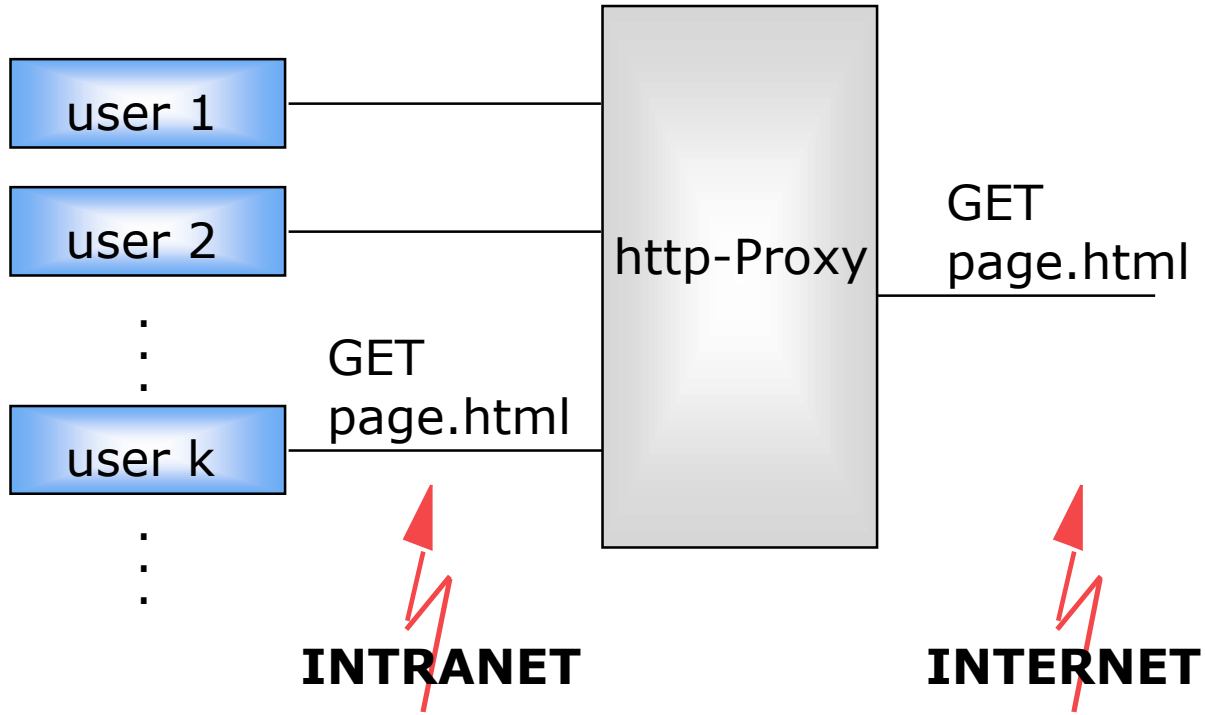
- Server has no information about the real originator of request
- **No protection against the operator**
- **No protection against traffic analysis**

- **Principles for Web access:**
  1. **Form-based**
     - Type in URL
     - Proxy gets the URL on behalf of user
  2. **Change browser config**
     - „use proxy"



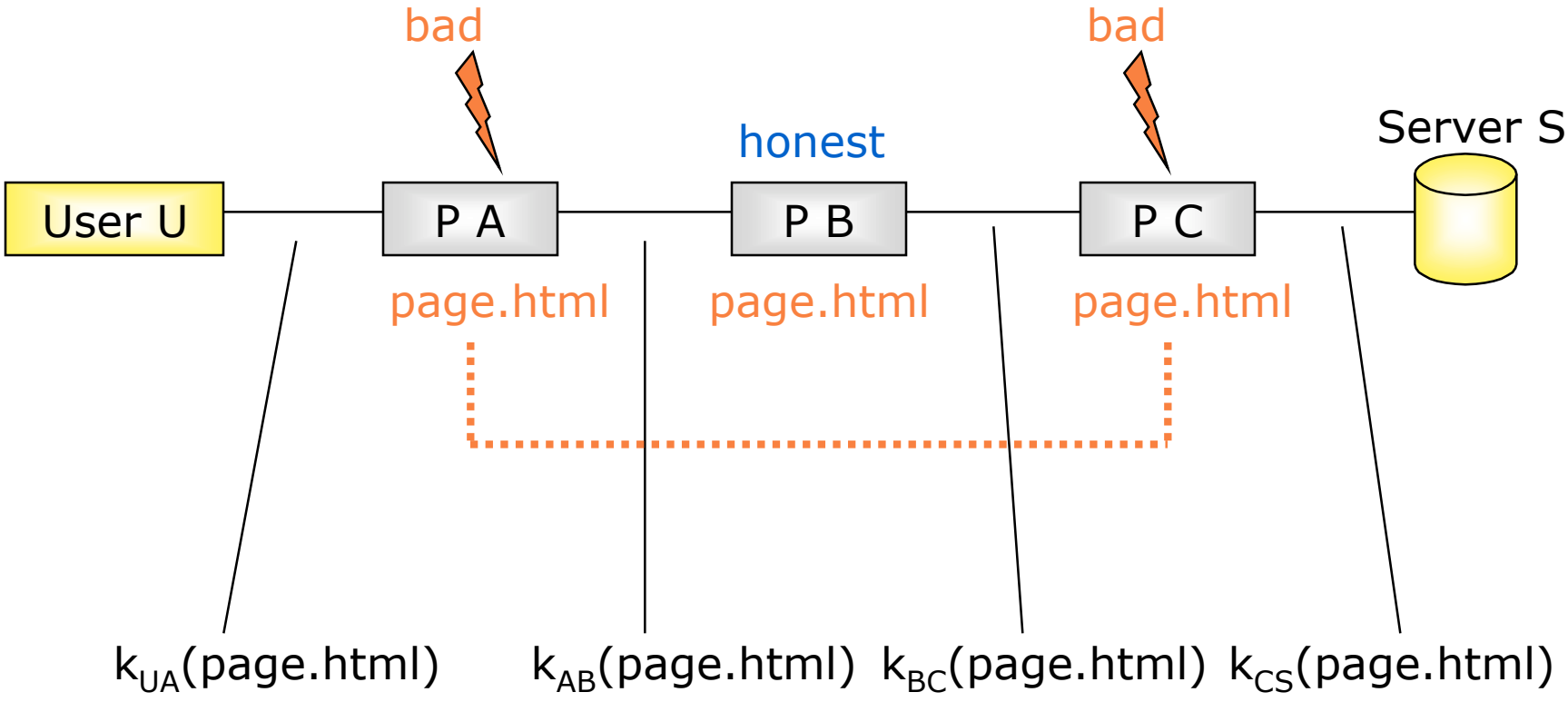
>> Simple Proxies

- Proxy gets to know all contents!!!
- Observation is possible
  - Timing correlation of incoming and outgoing requests
  - Correlation by message length and coding
  - Simple encryption between user and proxy is not sufficient because of the correlation of timing and length and it does not help against the operator

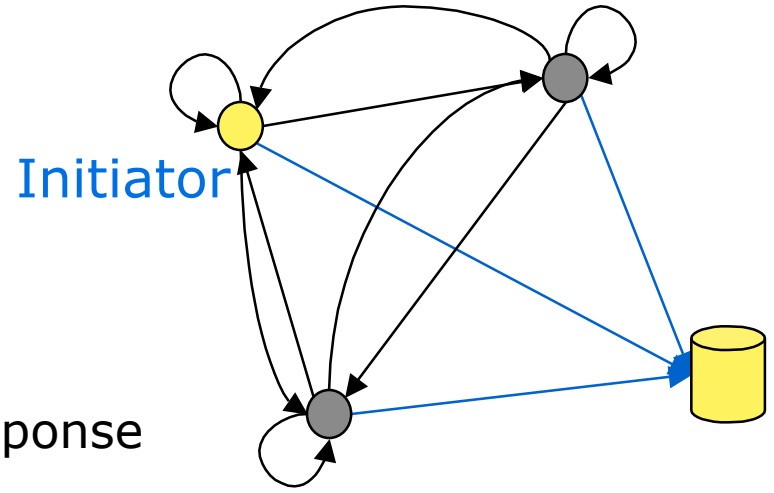


> Cascading Simple Proxies

- Link-to-link encryption between proxies
- Does not help to avoid observation by operators



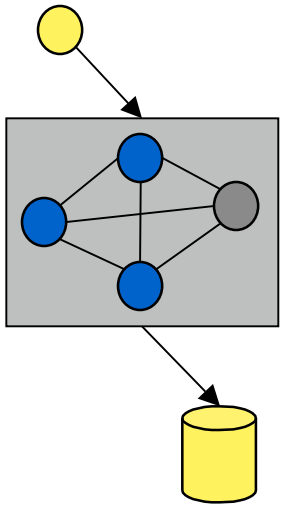
- Each communication request is sent directly to the server with a probability of P
- Else the request is sent to another user (Jondo) of the crowd (with 1-P)
- Symmetric link-encryption between the users
  - Avoid linkability
  - However: timing coincidence
- Enbedded objects (images etc.) are requested by the last Jondo
  - Suppress bursts of requests
- Security goal:
  - Every user can deny that he or she is the originator of a certain request
- Problem:
  - Jondos get to know about content of a request and response

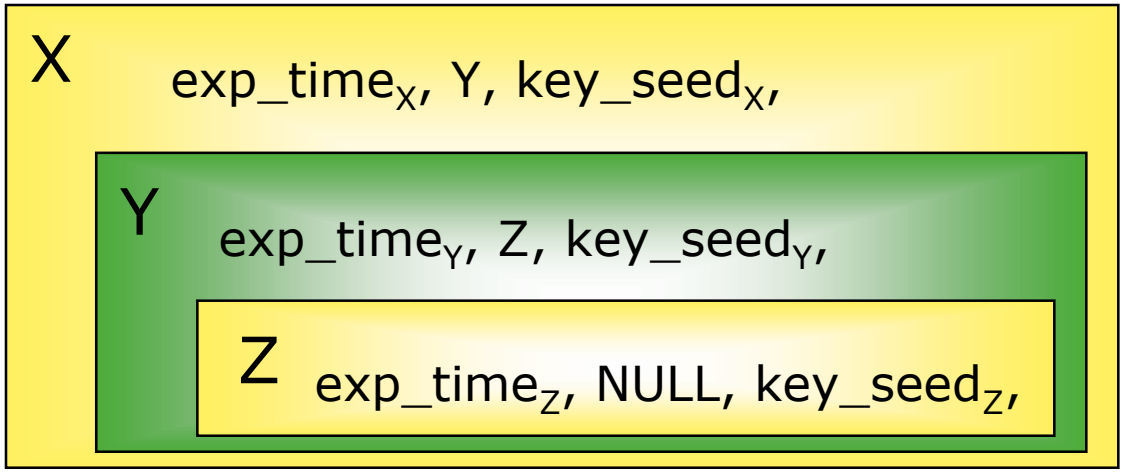


> Onion Routing

US Naval Research Center

- Hiding of routing information in connection oriented communication relations
- Nested public key encryption
- Uses an expiration_time field to reduce cost of replay detection
- Dummy traffic between MIXes (Onion Routers)
- First/Last-Hop-Attacks:
  - Timing correlations
  - Message length



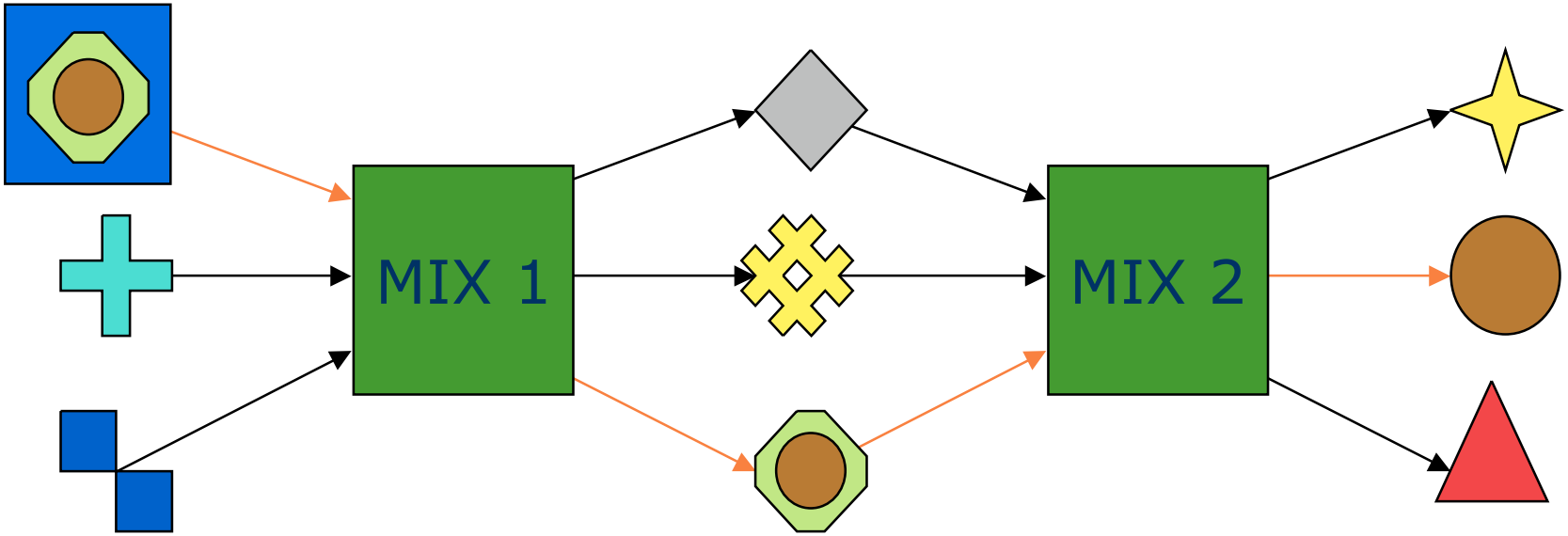


> Attacks

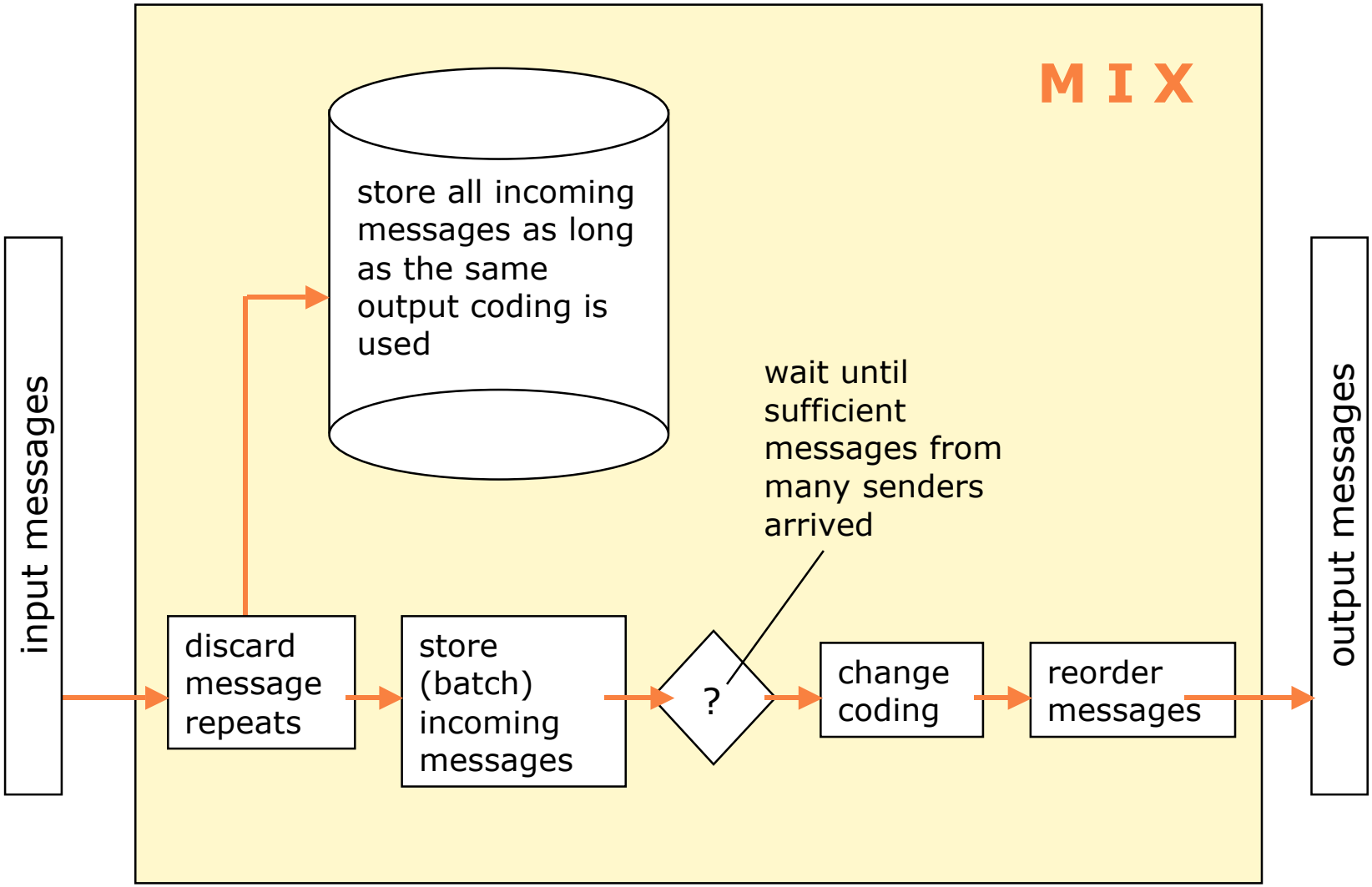
- Systems considering traffic analysis have to avoid all of the following possible attacks
  - MIX — Timing attacks: Observe the duration of a communication by linking the possible endpoints of a communication and wait for a correlation between the creation and/or release event at all possible endpoints.
  - MIX — Message volume attacks: Observe the amount of transmitted data (i.e. the message length) and correlate input and output.
  - ? — Flooding attacks: Each message can only be anonymous in a group of messages (batch). Under normal circumstances, each sender sends one message per batch. A good system has to avoid that the batch can be flooded by an attacker in order to separate a certain message.
  - ? — Linking attacks: Because of online/offline-periods of the users an attacker may create intersections of anonymity groups by observation over a long period.
- At this time, no existing system withstands all attacks

Mixes (David Chaum, 1981)

- Basic idea:
  - Sample messages in a batch, change their coding and forward them all at the same point oftime but in a different order. All messages have the same length.
  - Use more than one Mix, operated by different operators.
  - At least one Mix should not be corrupt.
- Then:
  - Perfect unlinkability of sender and recipient.

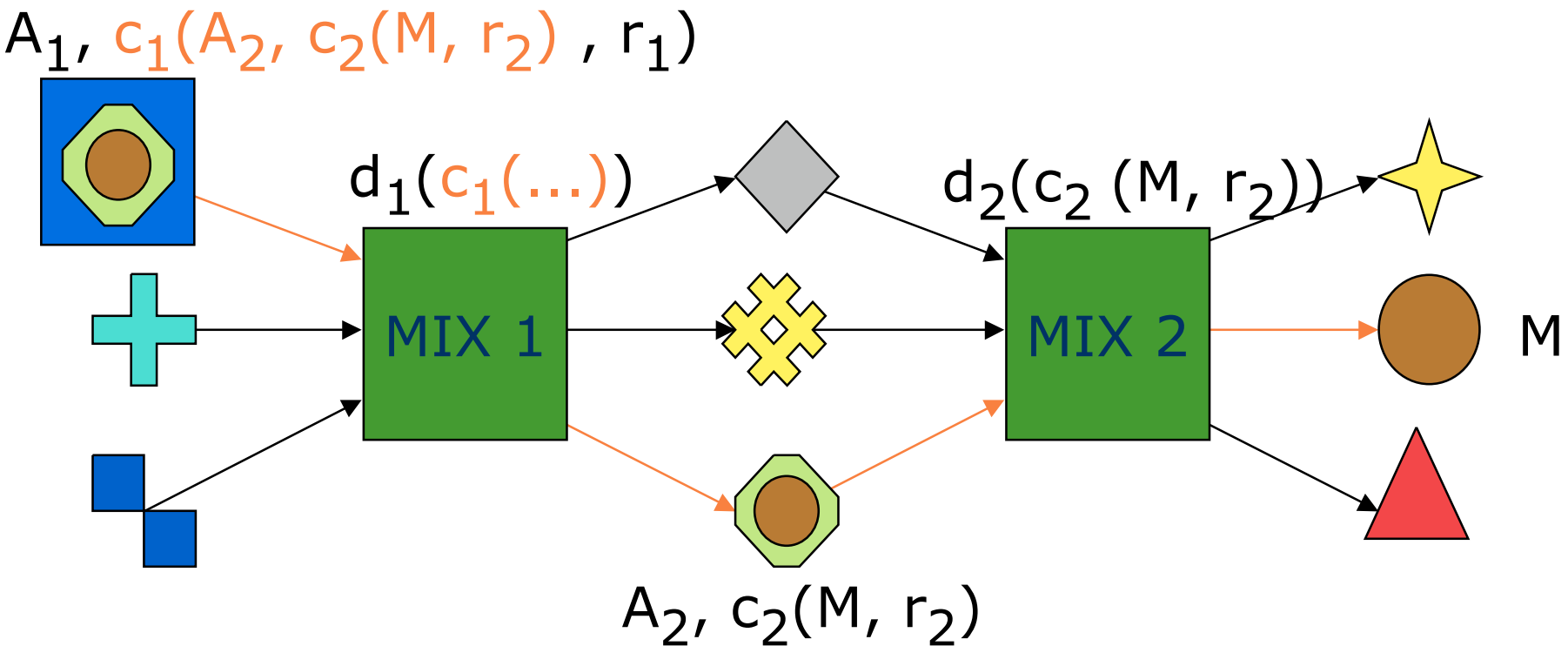


> How a MIX works



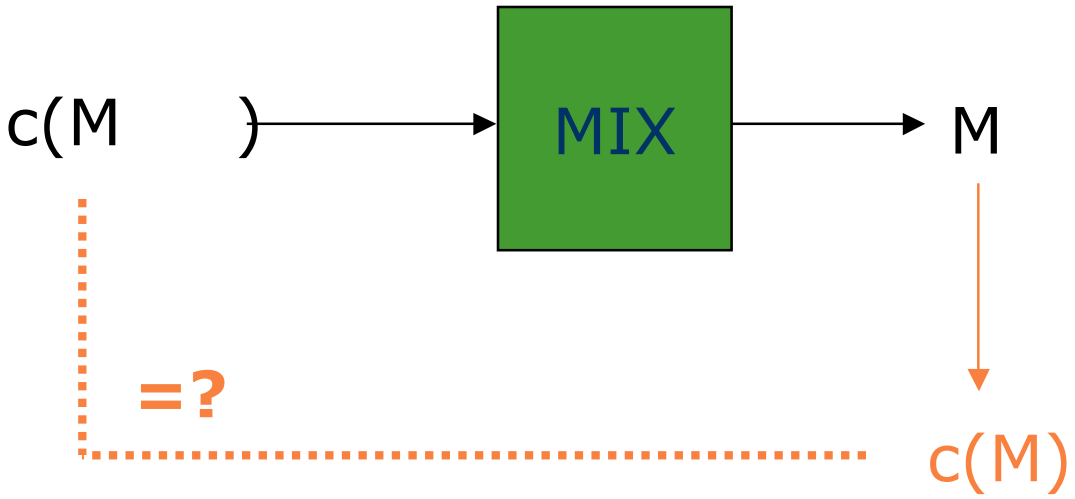
Mixes: some cryptography

- Use a public key cryptosystem:
- $c_i(...)$ is an encrypted message for Mix i (everybody can encrypt messages for Mixes using this function)
- $d_i(...)$ is the private function of Mix i to decrypt messages (only Mix i can decrypt his messages, nobody else)
- A_i is the address of Mix i; r_i are random numbers (dropped by the Mix)
- M is the message for the recipient (including his address)



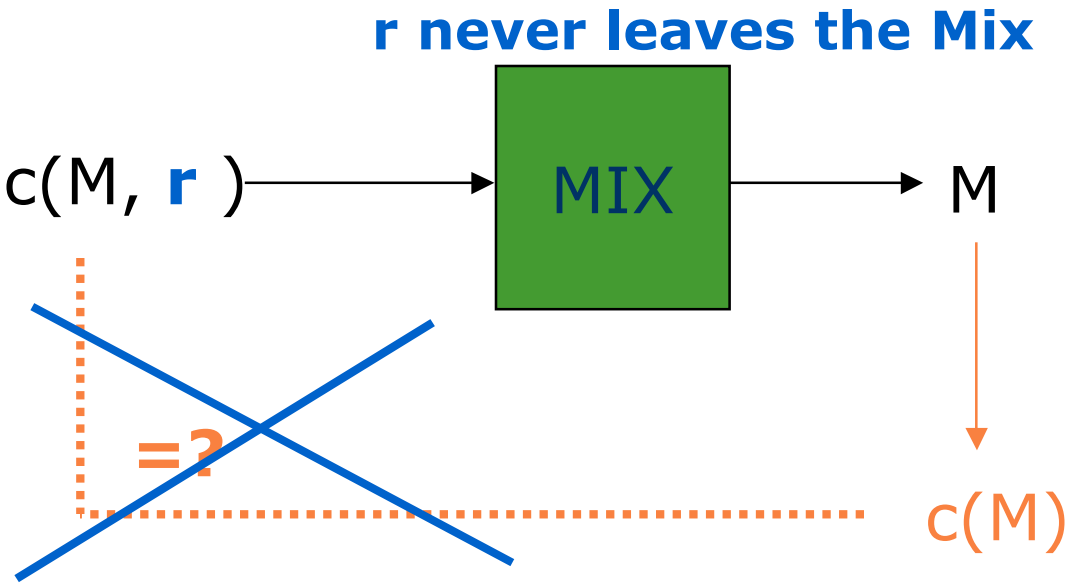
> Mixes: Why do we need random numbers?

- If no random numbers r used:
  - Everyone can encrypt the output messages of a Mix because c(...) is public
  - Compare results with all incoming messages
  - Need a indeterministic encryption scheme (or use random numbers)



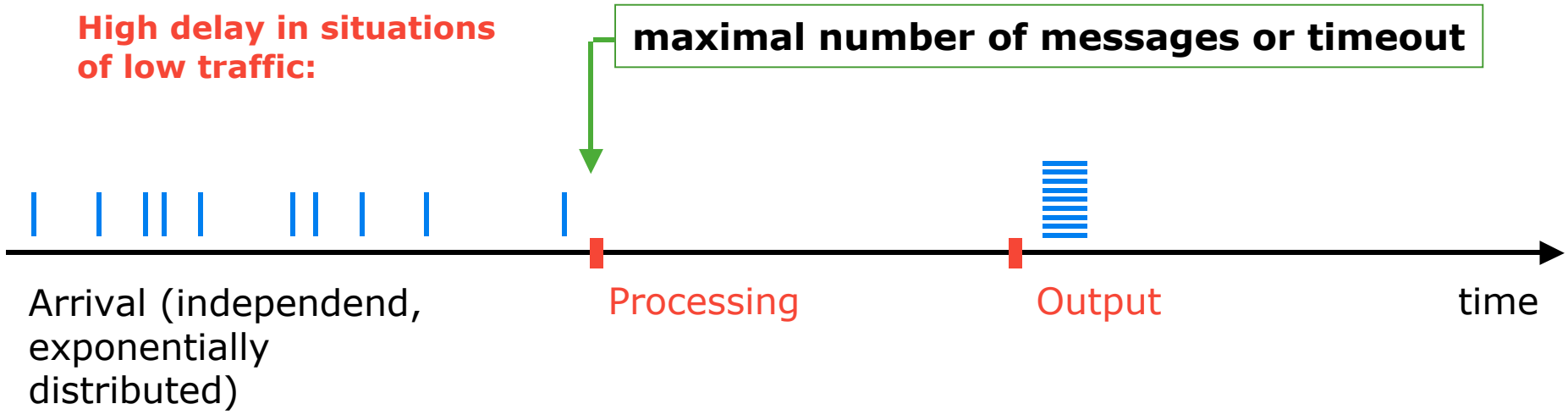
>> Mixes: Why do we need random numbers?

- If no random numbers r used:
  - Everyone can encrypt the output messages of a Mix because c(...) is public
  - Compare results with all incoming messages
  - Need a indeterministic encryption scheme (or use random numbers)



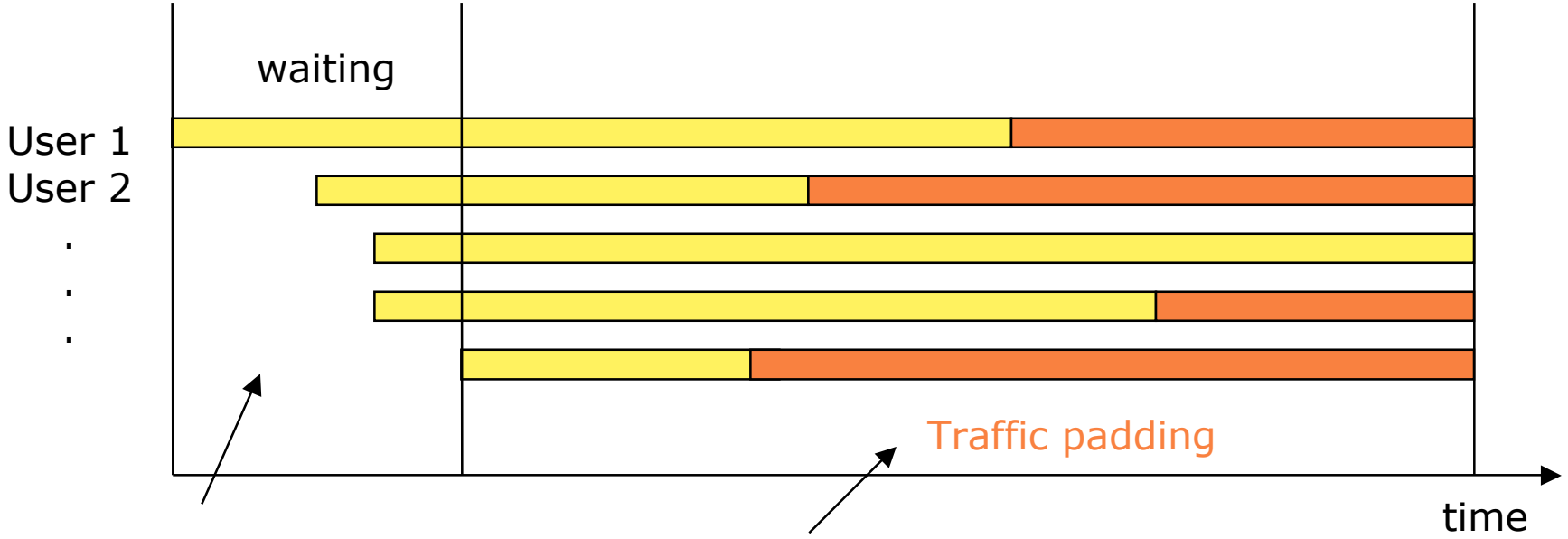
The problem of anonymous real-time communication

- Plain Mixes are good for non-real-time communication: E-Mail
- But not sufficient for real-time communication: Web, Ftp, Internet Phone
  - Sampling of messages means high delay, because a Mix is waits for (another) messages the most of time.
  - Message lengths vary in a very large interval or no support of connection oriented services
- We need a few improvements



> Traffic padding

- Hide from the attacker, when a certain communication ends
- But: nobody knows, when the last user wants to end his communication



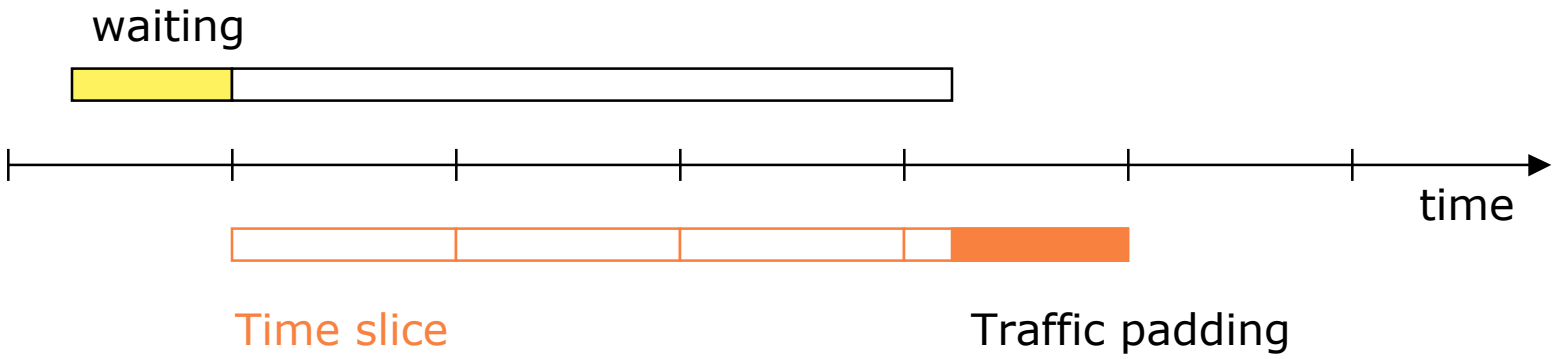
1. Users have to wait until enough users want to communicate (creation of the anonymity group)
   Example: 5 users

2. End of communication but users have to send random data until the last user has finished his connection

3. However: Nobody knows when the last user wants to end his communication – because nobody can distinguish real traffic from traffic padding

> Time slices and traffic padding

- Chopping of long communications into small pieces (connections or packet size)
  - Unobservability in the group of all processed messages at one time slice
  - Long communications consist of more than one time slice
  - No linkability of time slices



> Dummy traffic

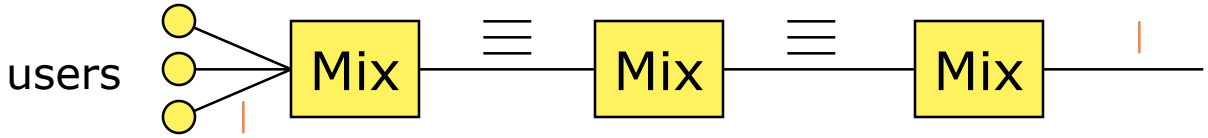
- Increase the amount of traffic in situations of low traffic

- Sometimes the number of users is not sufficient to fill the batch.
- This can happen in times of low traffic.
- In that case,
  - either the use has to wait until enough messages arrive (leads to likely high delay)
  - or accepts, that he cannot remain anonymous,
  - or other users send dummy traffic.
- **Def.: Dummy traffic**. A user sends messages at all times. When he doesn't want to send messages, he sends random numbers. Nobody can make a distinction between real encrypted messages and the random numbers.

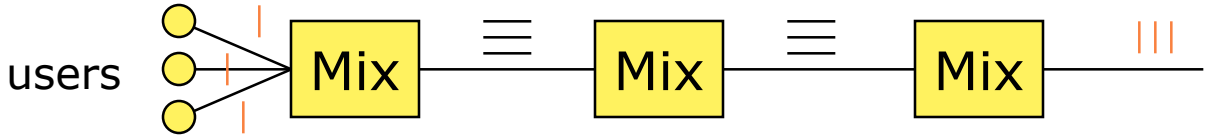
>> Dummy traffic

- Increase the amount of traffic in situations of low traffic

- Dummy traffic only between Mixes is not sufficient



- Dummy traffic has to be generated by the users



> Remaining attacks

- Systems considering traffic analysis have to avoid all of the following possible attacks:
  - Timing attacks

MIX

  - Message volume attacks

MIX

  - Flooding attacks: Each message can only be anonymous in a group of messages (batch). Under normal circumstances, each sender sends one message per batch. Avoid that the batch can be flooded by an attacker in order to separate a certain message.

?

  - Linking attacks: Because of the online/offline-periods of the users an attacker may create intersections of anonymity groups by observation over a long period.

?

> The Problem of flooding Mixes

- Batch size n
- Flooding: Attacker tries to flood the Mix with his own (n-1) messages, except one message that he wants to observe
- Attacker knows (n-1) outgoing messages. The only unknown message is the observed message.
- In that case, the sender and recipient are uncovered.

- Solution (first hack):
  - All incoming messages need a ticket to be processed by a Mix.
  - Now, the attacker needs help of the (n-1) other users. However, we assume the users will never harm themselves.
  - Very similar to an anonymous payment system.
  - Digital coin not traceable neither by the Mix nor the Bank.
  - Additionally, solves the problem of payment for anonymity systems

> The Problem of long-term observation of users

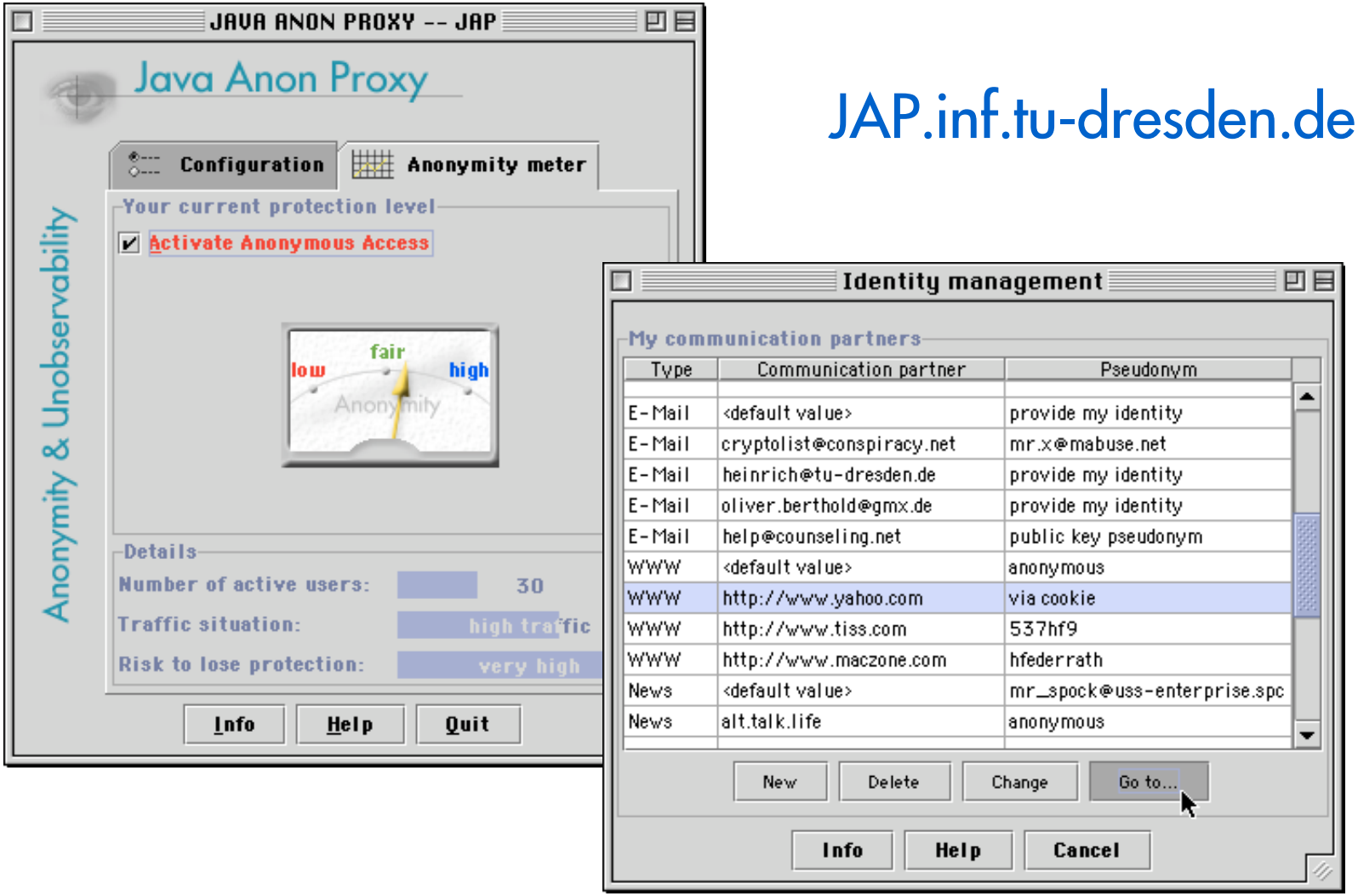
- Supposed:
  - A user shows a nearly constant online-offline behavior (from 8 - 10 PM online everyday)
  - Requests certain contents (web pages, his e-mail account) during this time
  - A lot of other people are also online and use the anonymity service
- Attacker observes all communication links and servers, except the anonymity service over a long time period.
- Long-term observation leads to intersections of anonymity groups and uncovers the users behavior.
- How long it takes that an attacker to link the user actions with a high probability depends on the size of the anonymity group and its behavior.
- Simulation of that attack
- No good solution at this time to defend this attack.

> Web Mixes: Anonymous real-time communication

University of Technology Dresden

- Anonymous and unobservable transport system
  - Mix-based proxies with additional functions to provide real-time communication
  - Should withstand strong (big brother) attacks
- Information service (impossible to operate a perfect Anon system)
  - Current level of protection (Anonymity level)
  - Trade-off between performance and protection should be decided by the user
- Open source, as soon as core functions have been completely implemented
  - Client software: Java (platform independent)
  - Server software: C/C++ (Win/NT, Linux/Unix)
- Technical and jurisdictional knowledge to serve legal issues
- Test application:
  - anonymous drug counseling site, supervised by an counselor, but without revealing identities

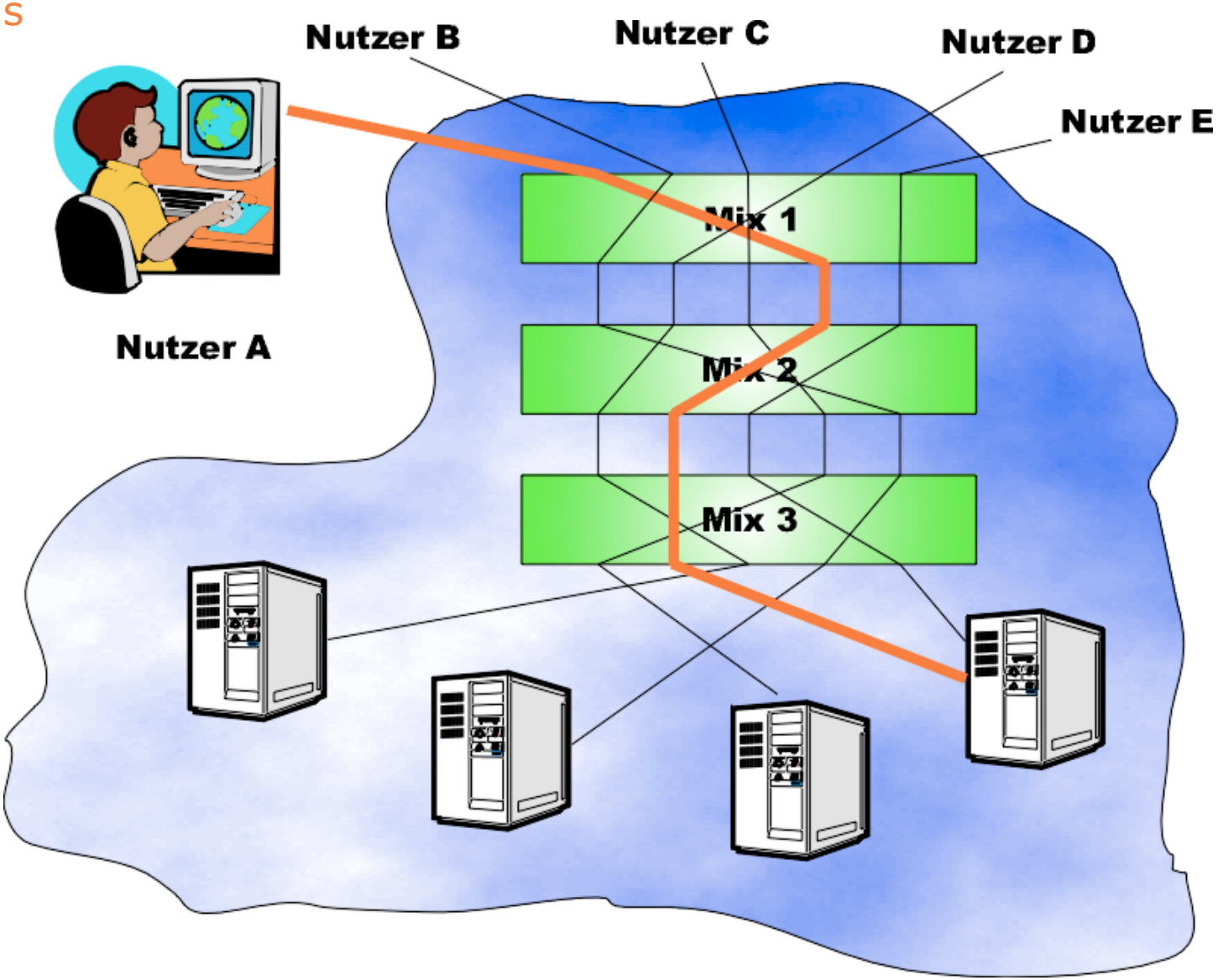
JAP.inf.tu-dresden.de



> How does it work?

University of Technology Dresden

- JAP acts as a local proxy on the local machine

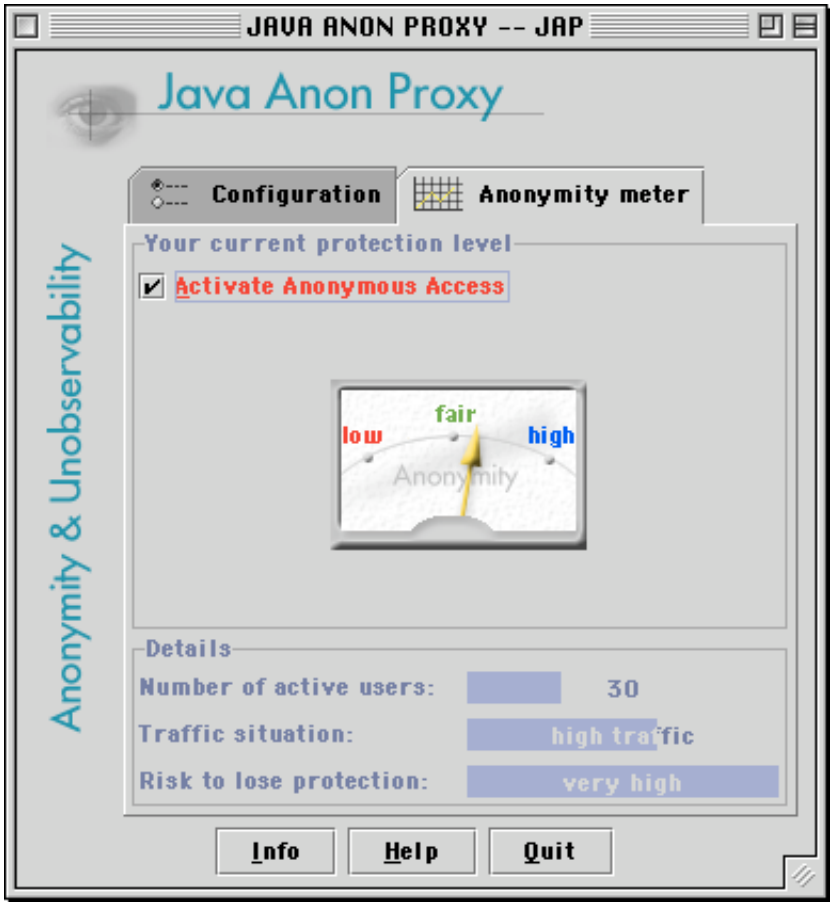


Some practical experiences

University of Technology Dresden

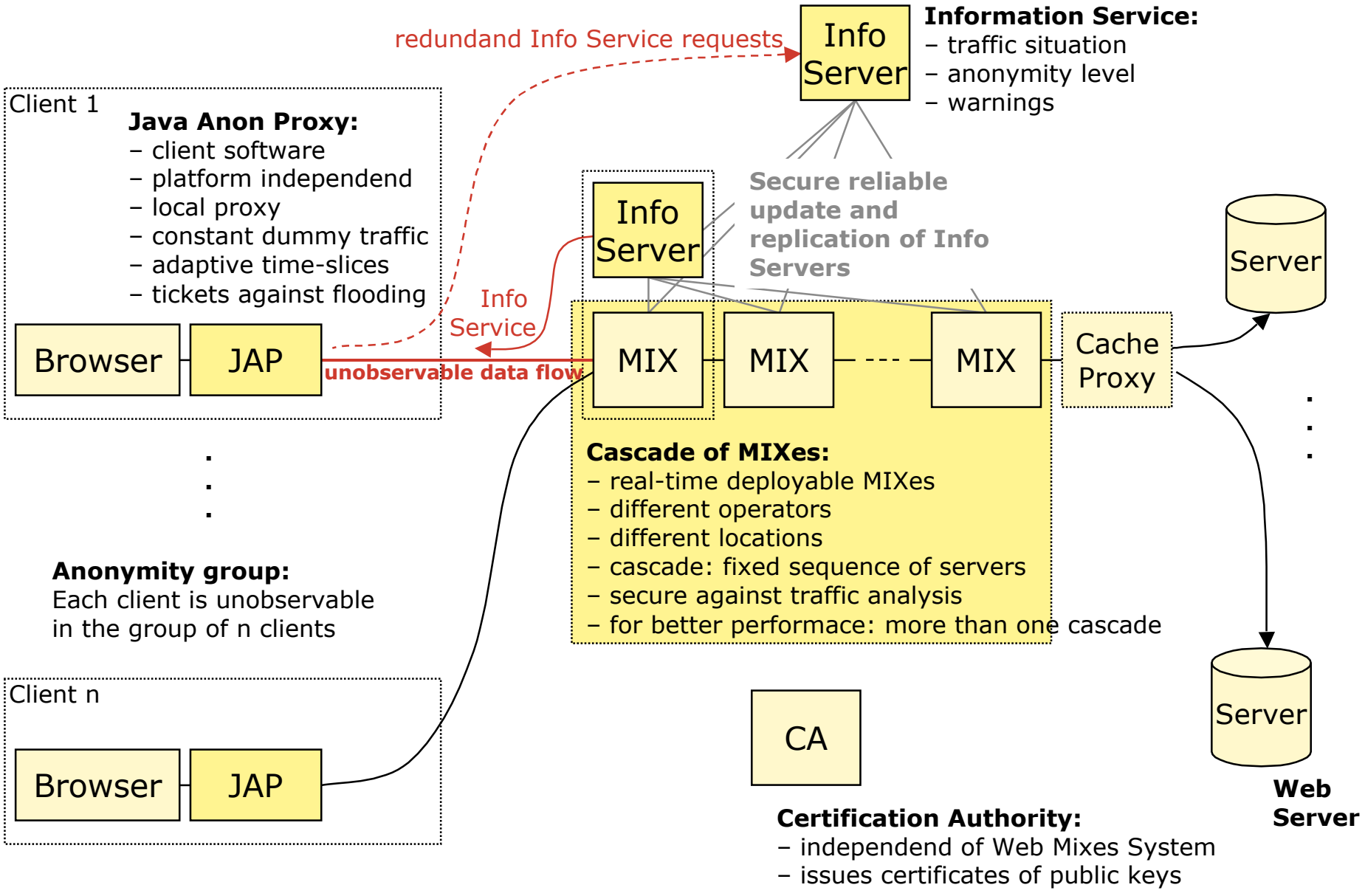
JAP.inf.tu-dresden.de

- First test version has been launched in October 2000
- Full service has been running since February 2001
- Hybrid encryption system of 128 bit encryption by AES (Rijndael) and RSA/1024 bit public key encryption
- 3 mix casades are running
- Busy hour: 500 users at the same time are online
- about 5000 – 8000 users
- about 120 gigabyte troughput per week



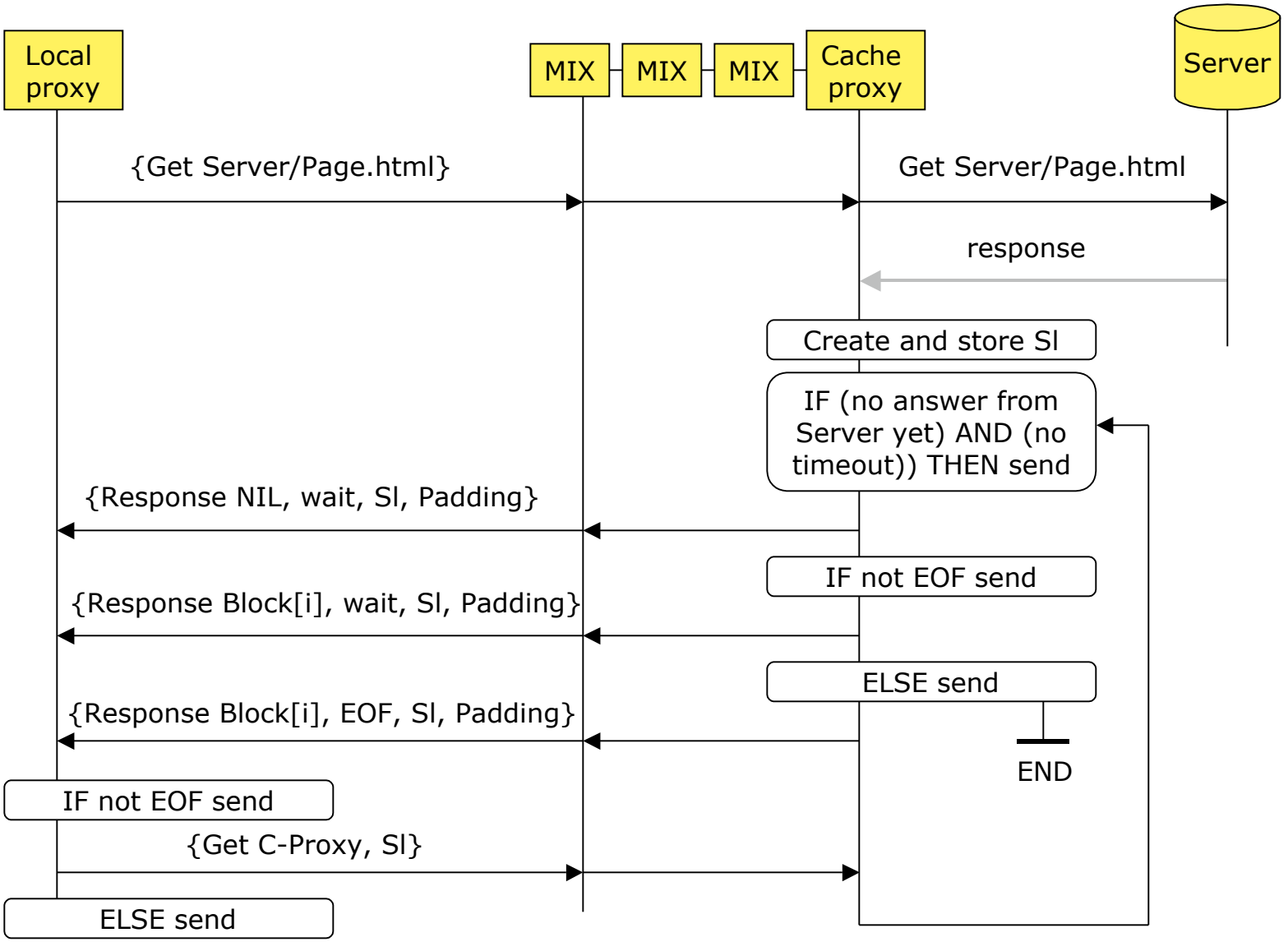
> Architecture of Web Mixes

University of Technology Dresden



> Time Slice protocol

University of Technology Dresden



> Some remarks about active content

- Deactivate Cookies in your browser
  - Web server can track all activities of a user
  - Additional filter software is very useful
    - http://www.webwasher.com/
    - http://www.junkbusters.com/ijb.html
  - Filter additional “bugs” that reveal your behavior
  - Example:very small (1x1) transparent pictures on a website

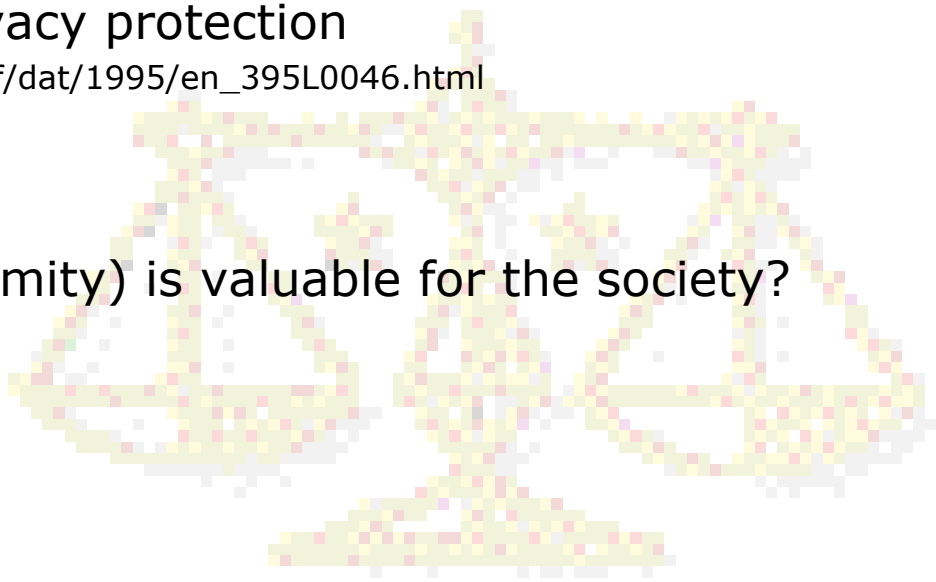
- Deactivate all sorts of active content in your browser
  - Java, JavaScript, ActiveX
  - IP-Address can be observed by an attacker
  - Unauthorized access to hard drive by ActiveX components

> Concluding remarks

- Anonymity and unobservability in the Internet is hard to realize.
- All commercial systems like Anonymizer, Freedom etc. suppose a weaker attacker model. They base their model on the assumption, that the strong attacks are not realistic in the Internet.
- In 95 or more percent of observation this assumption may be right, but not in the remainig 5 or less percent. Let's give an example of what we mean:
  - Assuming that an encryption tool sufficiently encrypts 99 of 100 messages, but in one case the message is sent in clear text. – Nobody will rely on that tool…
- That is exactly the situation using one of the existing systems.
- However, in some cases (or to defend some attacks) we do presently not know how a secure system has to be built.

> Political and social context

- Legal enforcement of communications
  - German Telekommunikationsüberwachungsverordnung (TKÜV)
    - http://www.bmwi.de/Homepage/download/telekommunikation_post/TKUEV-Entwurf.pdf
  - European Cybercrime Convention
    - http://conventions.coe.int/treaty/en/projets/cybercrime.htm
- Privacy laws
  - German (new) Bundesdatenschutzgesetz (BDSG)
    - http://www.bfd.bund.de/information/bdsg_hinweis.html
  - European directive on privacy protection
    - http://europa.eu.int/eur-lex/en/lif/dat/1995/en_395L0046.html
- Open question
  - How much privacy (anonymity) is valuable for the society?



>>> Privacy and Anonymity

Anonymous communication secure against traffic analysis

Information online ?

http://www.inf.tu-dresden.de/~hf2/anon/

- Demonstrations
- Downloads
- Links