Anonymity

Christian Grothoff
christian@grothoff.org
http://grothoff.org/christian/

"The problem with losing your anonymity is that you can never go back." –Marla Maples

Agenda

- **Definitions and Metrics**
- Techniques, Research Proposals and Systems

Motivation

Suppose Alice and Bob communicate using encryption.

What can Eve still learn here?

Motivation

Eve cannot read the data Alice and Bob are sending, but:

- Eve knows that Alice and Bob are communicating.
- Eve knows the amount of data they are sending and can observe patterns.

$\Rightarrow$ Patterns may even allow Eve to figure out the data

Anonymity Definitions

Merriam-Webster:

1. not named or identified: “an anonymous author”, “they wish to remain anonymous”

2. of unknown authorship or origin: “an anonymous tip”

3. lacking individuality, distinction, or recognizability: “the anonymous faces in the crowd”, “the gray anonymous streets” – William Styron

Anonymity Definitions

Andreas Pfitzmann et. al.:

> "Anonymity is the state of being not identifiable within a set of subjects, the anonymity set."

Anonymity Definitions

EFF:

"Instead of using their true names to communicate, (...) people choose to speak using pseudonyms (assumed names) or anonymously (no name at all)."

Anonymity Definition

Mine:

A user's action is anonymous if the adversary cannot link the action to the user's identity

The user's identity

includes personally identifiable information, such as:

- real name
- fingerprint
- passport number
- IP address
- MAC address
- login name
- ...

Actions

include:

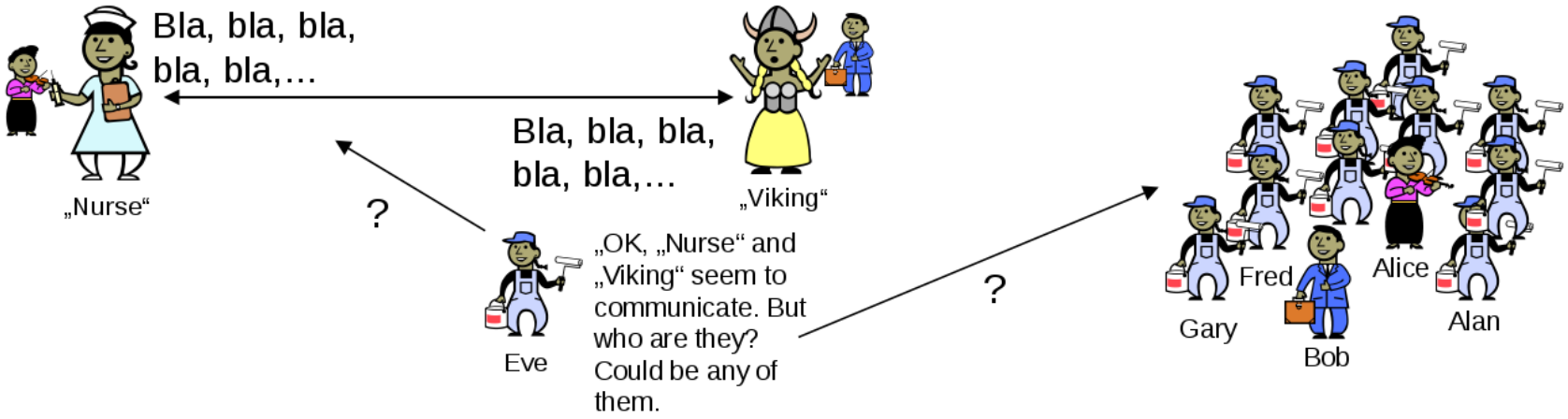
- Internet access
- speach
- participation in demonstration
- purchase in a store
- walking across the street
- ...

Anonymity: Terminology

- Sender Anonymity: The initiator of a message is anonymous. However, there may be a path back to the initiator.

- Receiver Anonymity: The receiver of a message is anonymous.

Pseudonymity



Pseudonymity

- A pseudonym is an identity for an entity in the system. It is a "false identity" and not the true identity of the holder of the pseudonym.
- Noone, but (maybe) a trusted party may be able to link a pseudonym to the true identity of the holder of the pseudonym.
- A pseudonym can be tracked. We can observe its behaviour, but we do not learn who it is.

Basic adversary characteristics

- Position
  - External: "sits" on the wire
  - Internal: participates in the anonymizing system
- Geographic
  - Global: sits on all wires
  - Local: sits on some local wires
  - Partial: controls parts of the network
- Participation
  - Passive: only observes traffic
  - Active: may send, modify, and drop messages

Typical Adversary Models

- Global Passive Adversary (GPA)
  - Observes and analyses the complete network
  - No active participation in the network
  - External attacker
- Global Active Adversary
  - Also performs active attacks
- Partial Passive Adversary (PPA)
  - Observes only parts ($<<$ 50%) of the network
  - External attacker
- PPA or GPA with some active nodes
- Local observer

Evaluationg Anonymity

How much anonymity does a given system provide?

- Number of known attacks?
- Lowest complexity of successful attacks?
- Information leaked through messages and maintenance procedures?
- Number of users?

Anonymity: Basics

- **Anonymity Set** is the set of suspects
- Attacker computes a **probability distribution** describing the likelyhood of each participant to be the responsible party.
- Anonymity is the stronger, the larger the anonymity set and the more evenly distributed the subjects within that set are.

Anonymity Metric: Anonymity Set Size

Let $\mathcal{U}$ be the attacker's probability distribution describing the probability that user $u \in \Psi$ is responsible.

$$ASS = \sum_{\substack{u \in \Psi \\ u > 0}} 1 \tag{1}$$

Large Anonymity Sets

Examples of large anonymity sets:

- Any human
- Any human speaking English
- Any human with phone access
- Any human with Internet access
- Any human speaking English with Internet access

Anonymity Metric: Maximum Likelihood

Let $\mathcal{U}$ be the attacker's probability distribution describing the probability that user $u \in \Psi$ is responsible.

$$ML = \max_{u \in \Psi} p_u \tag{2}$$

Anonymity Metric: Maximum Likelihood

- For successful criminal prosecution in the US, the law requires $ML\ close$ to 1 ("beyond reasonable doubt")
- For successful civil prosecution in the US, the law requires $ML > \frac{1}{2}$ ("more likely than not")
- For a given anonymity set, the best anonymity is achieved if

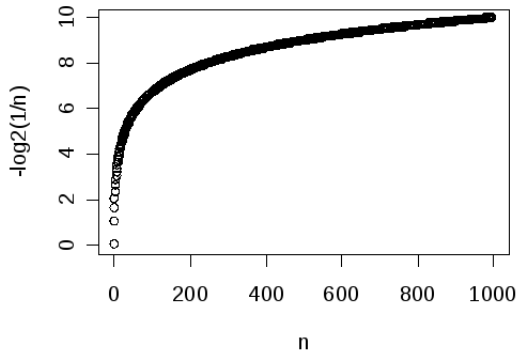
$$ML = \frac{1}{ASS} \tag{3}$$

Anonymity Metric: Entropy

Let $\mathcal{U}$ be the attacker's probability distribution describing the probability that user $u \in \Psi$ is responsible. Define the effective size S of the anonymity distribution $\mathcal{U}$ to be:

$$S = -\sum_{u \in \Psi} p_u \log_2 p_u \tag{4}$$

where $p_u = \mathcal{U}(u)$.



Interpretation of Entropy

$$S = -\sum_{u \in \Psi} p_u \log_2 p_u \tag{5}$$

This is the *expected* number of bits of additional information that the attacker needs to definitely identify the user (with absolute certainty).

Entropy Calculation Example

Suppose we have 101 suspects including Bob. Furthermore, suppose for Bob the attacker has a probability of 0.9 and for all the 100 other suspects the probability is 0.01.

What is S?

Entropy Calculation Example

- For 101 nodes $H_{max} = 6.7$
-

$$S = -\frac{100 \cdot \log_2 0.001}{1000} - \frac{9 \cdot \log_2 0.9}{10} \tag{6}$$

$$\approx 0.9965 + 0.1368 \tag{7}$$

$$= 1.133... \tag{8}$$

Agenda

- Definitions and Metrics
- **Techniques, Research Proposals and Systems**

Attacks to avoid

Hopeless situations include:

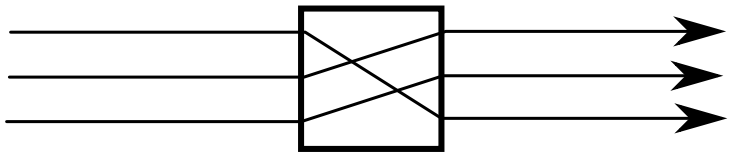
- All nodes collaborate against the victim
- All directly adjacent nodes collaborate
- All non-collaborating adjacent nodes are made unreachable from the victim
- The victim is required to prove his innocence

Anonymity: Dining Cryptographers

"Three cryptographers are sitting down to dinner. The waiter informs them that the bill will be paid anonymously. One of the cryptographers maybe paying for dinner, or it might be the NSA. The three cryptographers respect each other's right to make an anonymous payment, but they wonder if the NSA is paying." – David Chaum

Mixing

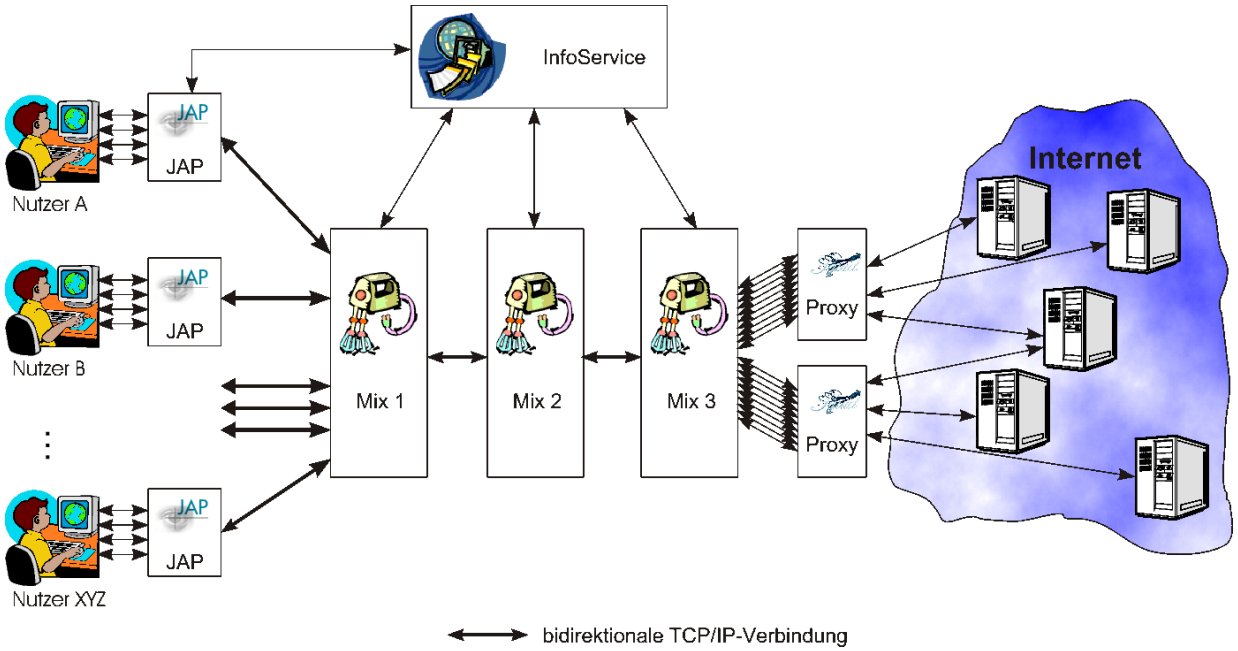
David Chaum's mix (1981) and cascades of mixes are the traditional basis for destroying linkability:



Mixing

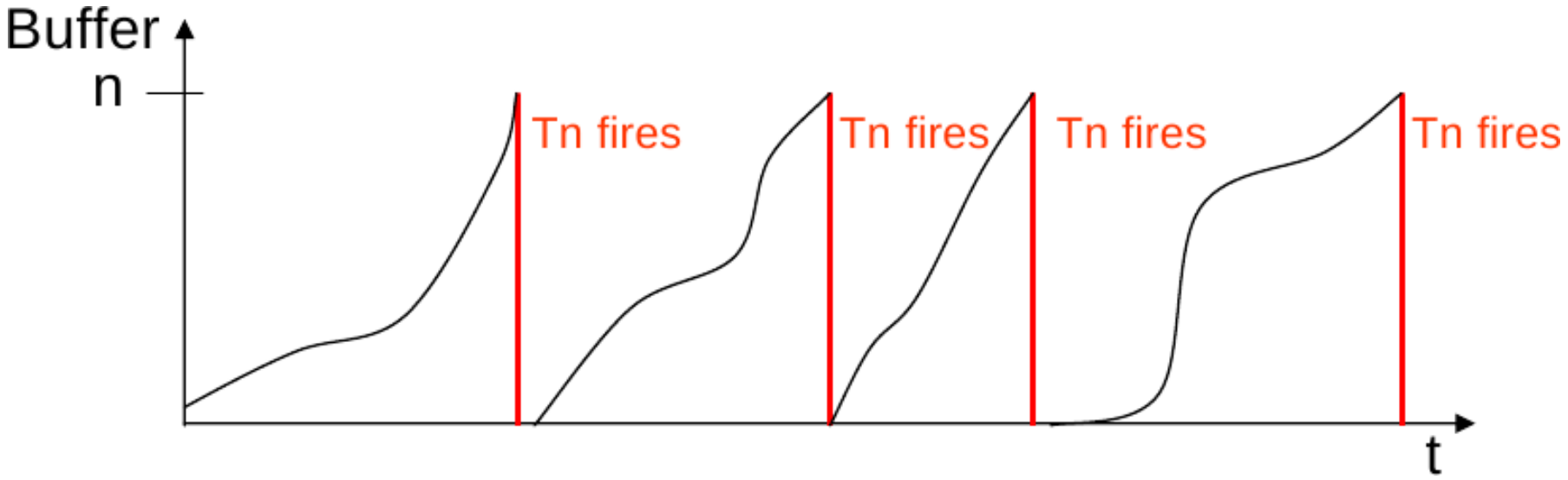
David Chaum's mix (1981) and cascades of mixes are the traditional basis for destroying linkability:

JAP[1]

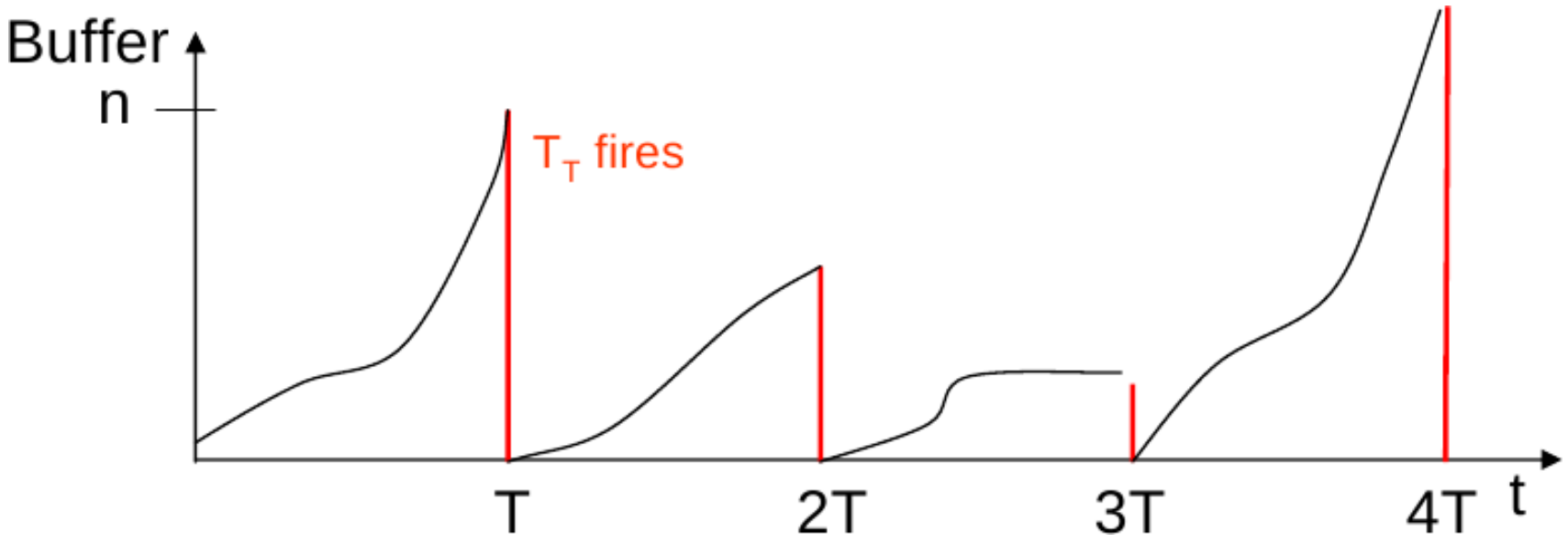


[1]From Stefan Köpsell: "AnonDienst – Design und Implementierung", 2004

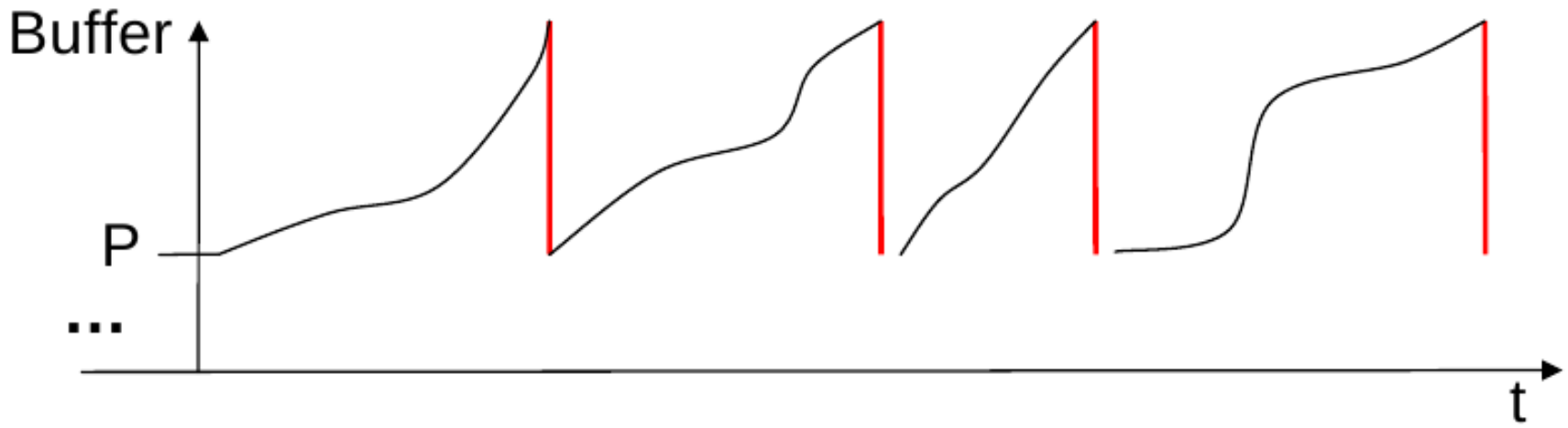
Threshold Mix



Timed Mix



Pool mix



The Number of Hops

What is more secure: more hops or fewer hops?

Path lifetime

What is more secure:
short-lived paths or long-lived paths?

Copyright