FILTER DESIGN FOR DELAY-BASED ANONYMOUS COMMUNICATIONS

Universidad de Vigo
GPSC Signal Processing in Communicatons Group

Simon Oya* Fernando Pérez-González* Carmela Troncoso†

*Signal Theory and Communications Dept., University of Vigo †IMDEA Software Institute

institute IMDEA software

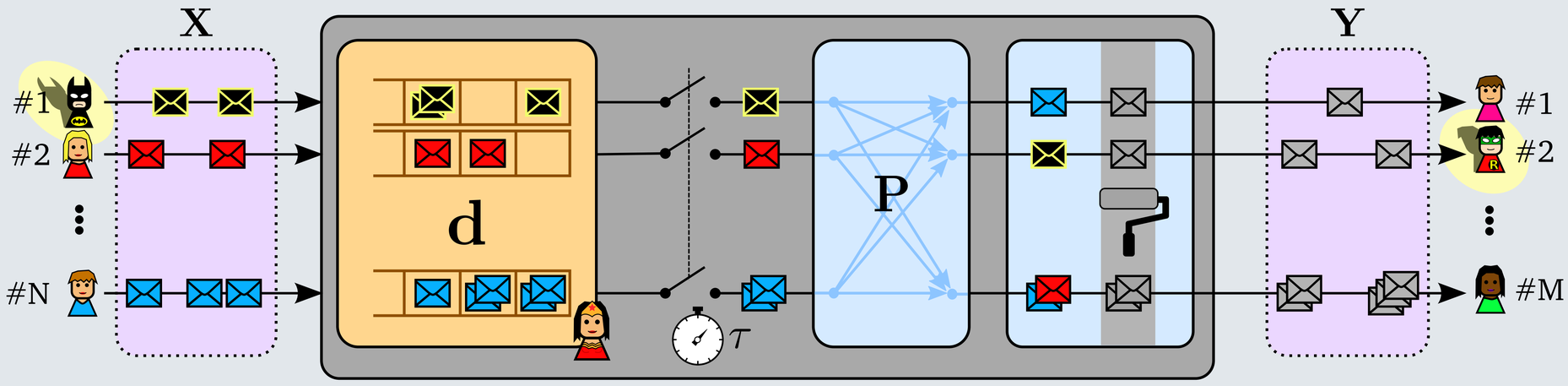
Delay-Based Anonymous Communication System

Some users send messages through the anonymous communication system

The anonymous communication system (Timed Mix) works in four steps

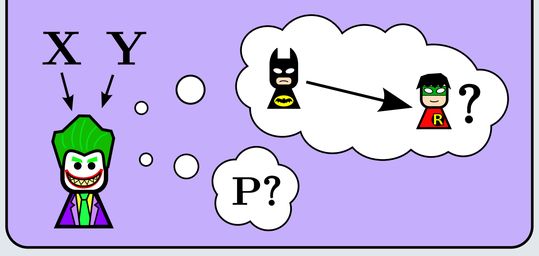
Delay messages randomly — Forward messages periodically — Route messages to their recipients — Change message appearance

The recipients receive the messages after some time



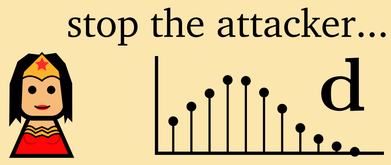
Attacker

Observes inputs and outputs



Defender

Designs the probability mass function of the delay to stop the attacker...



Another way of looking at the problem: linear filtering

From the model:

$$\mathbf{Y} = \mathbf{D} \cdot \mathbf{X} \cdot \mathbf{P} + \mathbf{N}$$

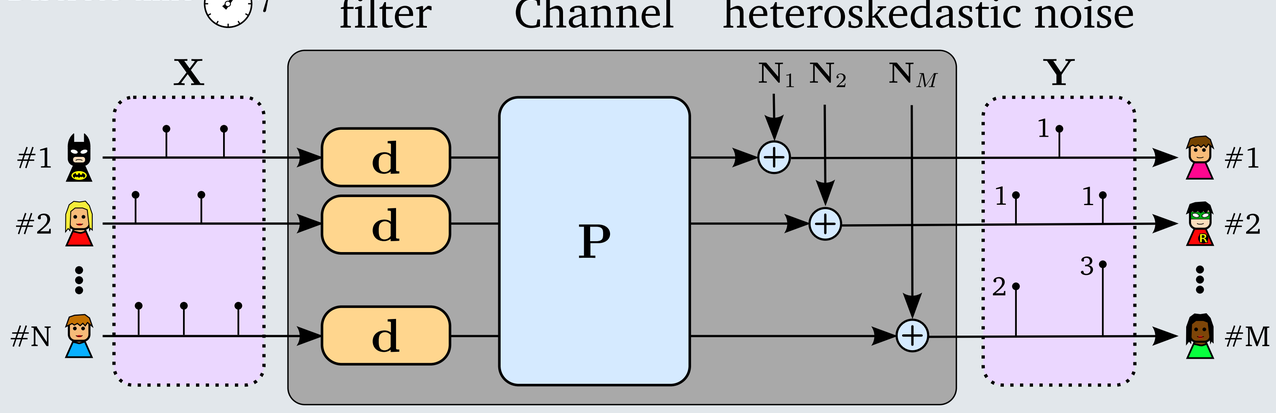
Convolution matrix that contains $\mathbf{d}$

Zero-mean noise with covariance

$$\mathbf{\Sigma_{N|X}} = \mathrm{diag}\{\mathbf{DX1}_N\} - \mathbf{D}\,\mathrm{diag}\{\mathbf{Xv}\}\mathbf{D}^T$$

This is like a MIMO system:

Discrete signals, Discrete time τ — Linear filter — MIMO Channel — Zero-mean heteroskedastic noise



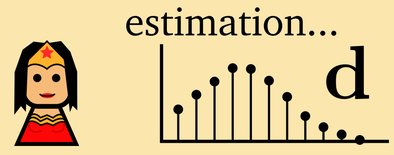
Attacker

The optimal adversary performs a least-squares estimation

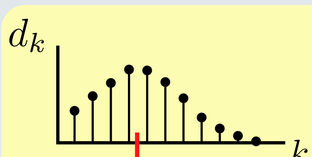
$$\hat{\mathbf{P}} = (\mathbf{X}^T\mathbf{D}^T\mathbf{D}\mathbf{X})^{-1}\,\mathbf{U}^T\mathbf{D}^T\mathbf{Y}$$

Defender

Designs the linear filter that worsens the least-squares estimation...



This is a filter design problem... with some constraints



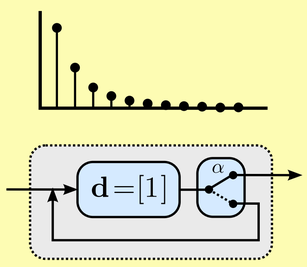
Positive taps $d_k \geq 0$ ⇒ It is easier to achieve attenuation at high frequencies; and the first DFT coefficient is larger than the others.

Adds up to one $\sum_k d_k = 1$ ⇒ The first DFT coefficient is one.

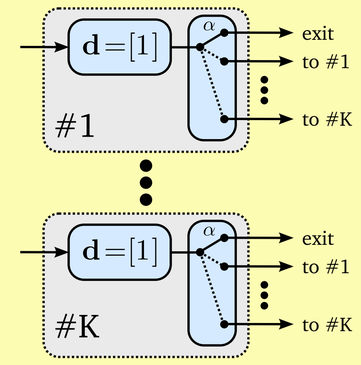
Average delay $\sum_k d_k \cdot k \leq \bar{\delta}$ ⇒ We prefer the minimum-phase solution.

Applications of filter design

An exponential delay is an IIR filter:

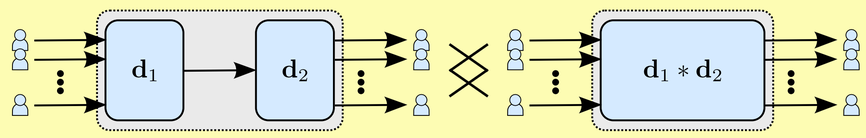


Distributed implementation of the exponential delay

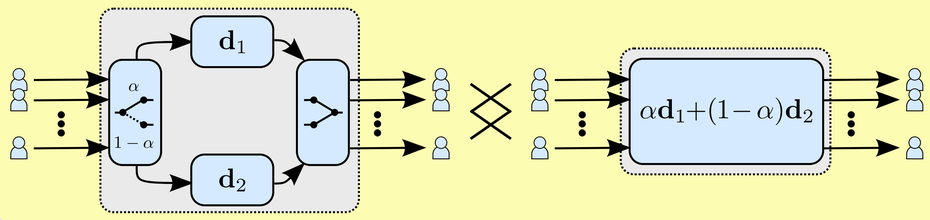


In real systems, we have networks of mixes

Cascade of mixes = convolution of filters



Mixes in parallel = sum of filters



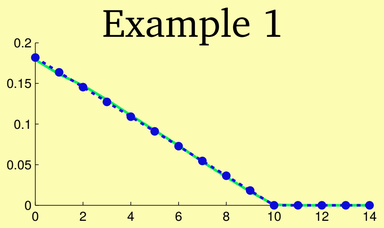
Evaluation

Depending on the system parameters, the shape of the "optimal filter" changes.

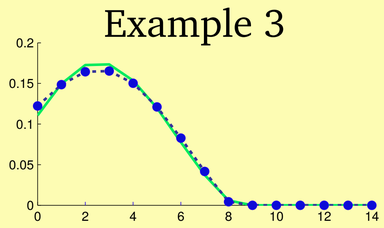
Numerical, from real data. (Enron, Gowalla, Indymedia)

Analytical, from the filter design problem.

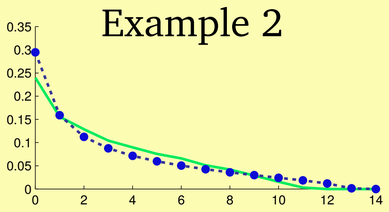
Example 1



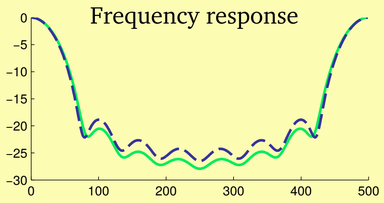
Example 3



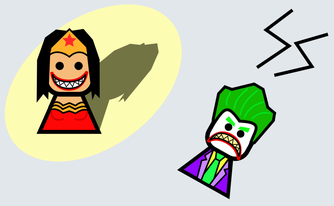
Example 2



Frequency response



The **analytical** designs are very close to the **numerical** solutions with real-data experiments!



S. Oya is funded by the Spanish Ministry of Education, Culture and Sport under the FPU grant. F. Pérez-González is funded by the Spanish Ministry of Economy and Competitiveness and the ERDF under projects TACTICA, COMPASS (TEC2013-47020-C2-1-R) and COMONSENS (TEC2015-69648-REDC), and by the Galician Regional Government and ERDF under projects GRC2013/009 and AtlantTIC. C. Troncoso is supported by EU H2020-ICT-10-2015 NEXTLEAP (GA n 688722).