Do dummies pay off?
Limits of dummy traffic protection in anonymous communication systems

Simon Oya, Carmela Troncoso and Fernando Pérez-González

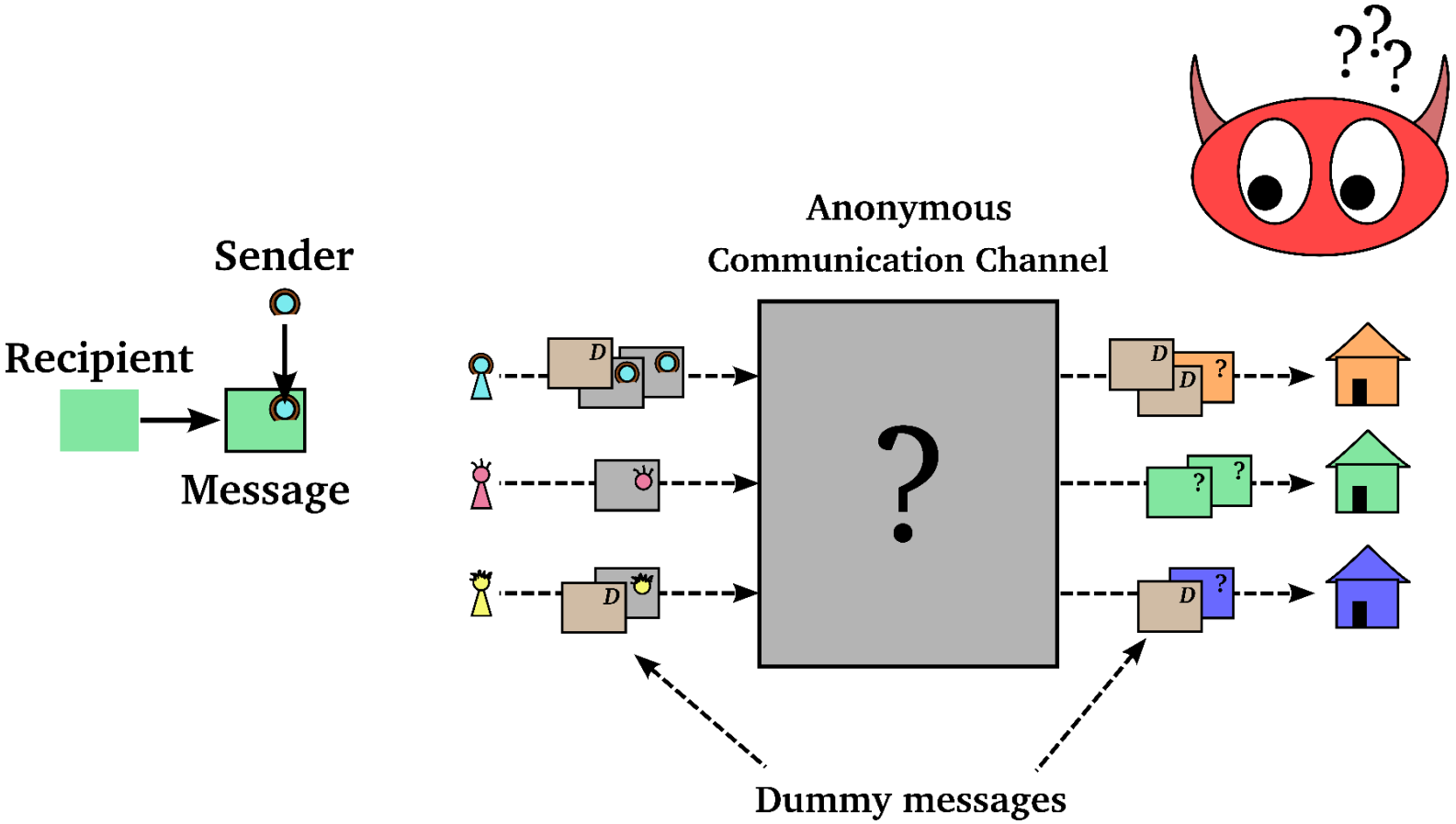
Universidad de Vigo
GPSC
Signal Processing in Communications Group

Gradiant

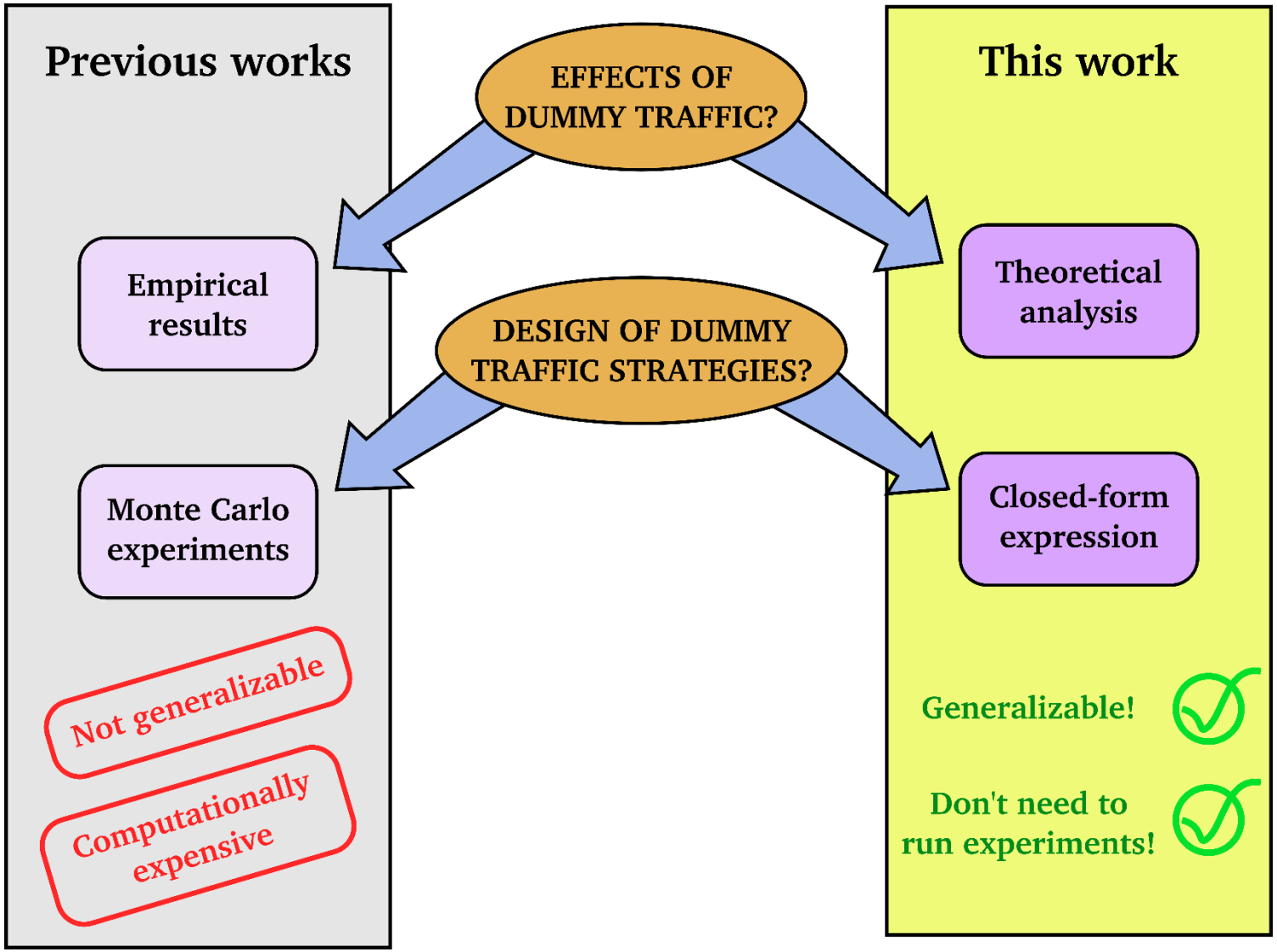
LAFTGATE

Introduction

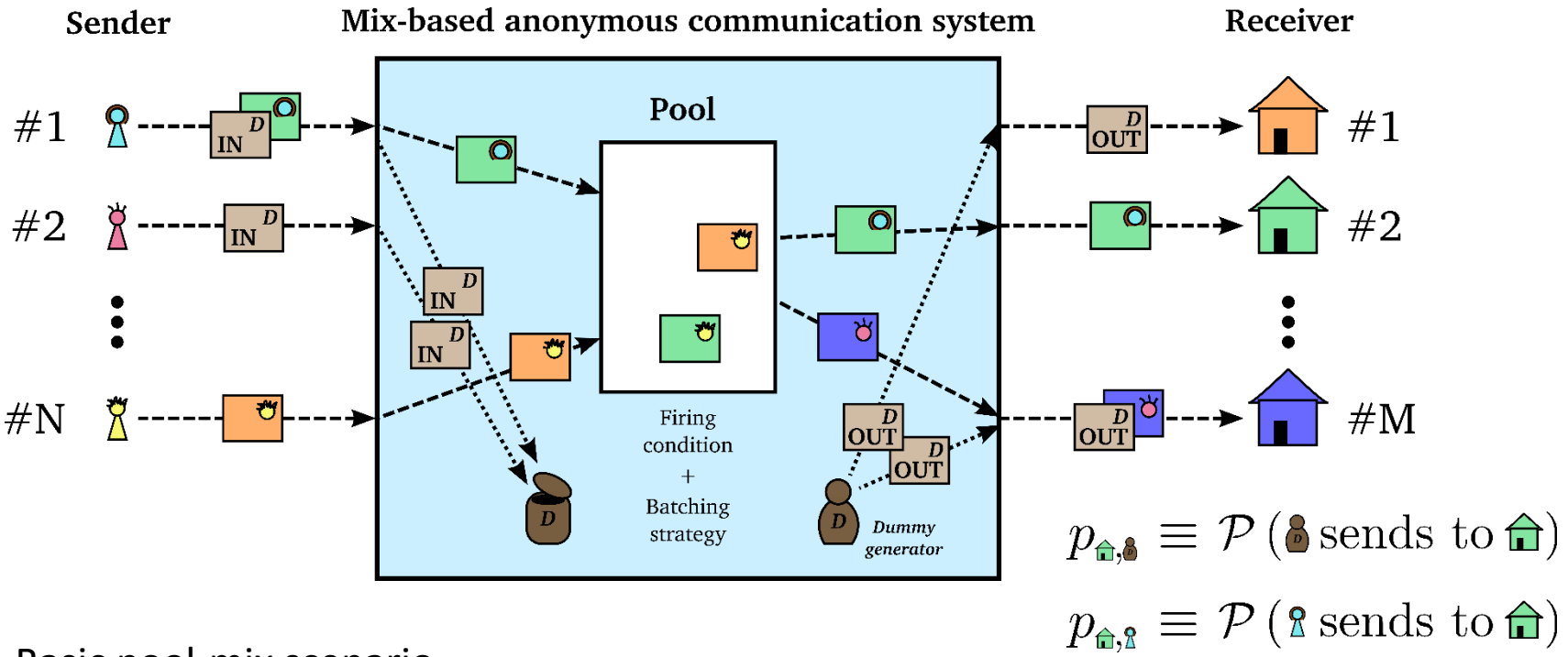
- Anonymous channels hide correspondences (input/output).
- **Dummies** are a common protection mechanism.



Motivation


Previous works
EFFECTS OF DUMMY TRAFFIC?
This work
Empirical results
Theoretical analysis
DESIGN OF DUMMY TRAFFIC STRATEGIES?
Monte Carlo experiments
Closed-form expression
Not generalizable
Computationally expensive
Generalizable!
Don't need to run experiments!

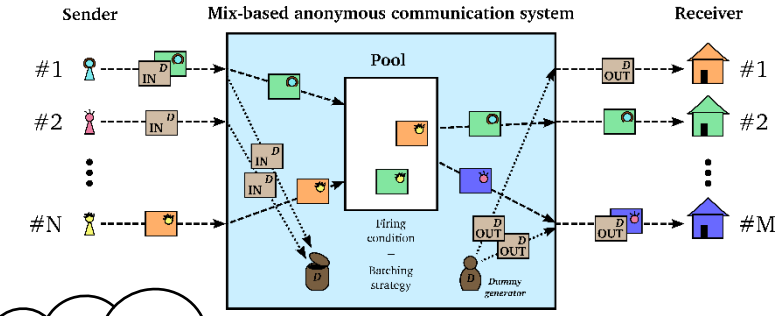
System Model



- Basic pool-mix scenario
  - Firing condition: triggers the flushing of messages.
  - Batching strategy: how messages are chosen from the pool.
- Sender-based dummies: die at the mix.
- Mix-based dummies: generated by the mix.
- (This was one communication **round**)

System Model. Adversary.

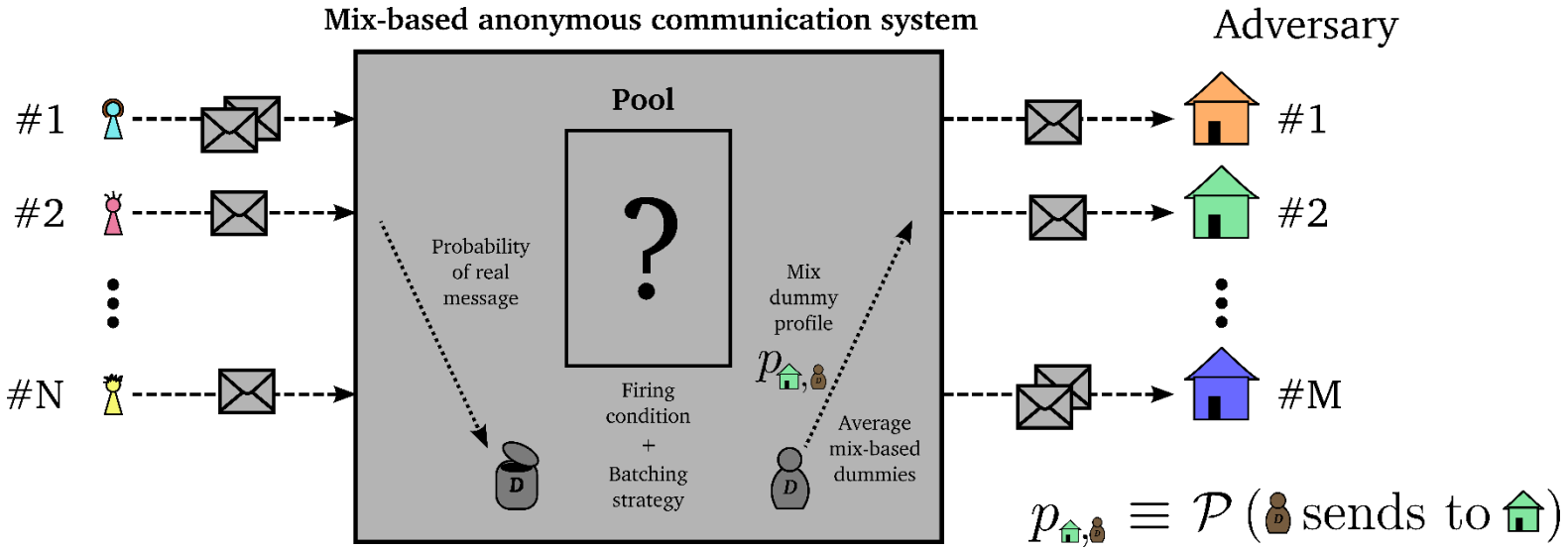
- **Goal:** infer $p_{\text{house},\text{person}} \equiv \mathcal{P}(\text{person sends to house})$
- **She cannot see the content of the messages…**
  - … but she can see the number of messages sent/received.
- **She is aware of the system parameters:**
  - Mix behavior
  - Dummy policy



observes ρ communication rounds

?

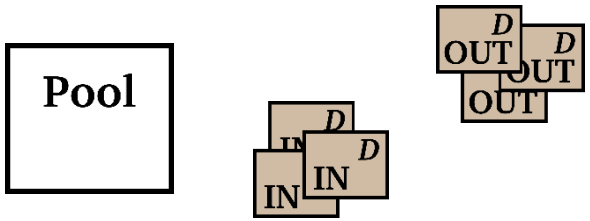
Adversary

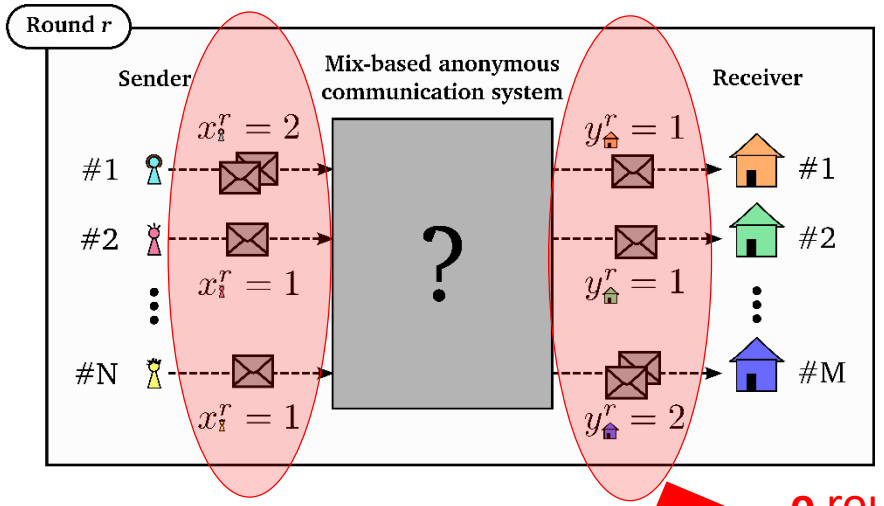


$p_{\text{house},\text{D}} \equiv \mathcal{P}(\text{D sends to house})$

Least-Squares Estimator

- **Goal:** infer all $p_{\text{house},\text{person}}$
- **Information:**





- **Idea:** minimize output error:

$$\hat{\mathbf{p}} = \arg\min \mathrm{E}\left\{||\mathbf{y} - \mathbf{Y}||^2\right\}$$

$$0 \le p_{\text{house},\text{person}} \le 1$$

$$p_{\text{house},\text{person}} + \cdots + p_{\text{house},\text{person}} = 1$$

- Expanding
- Rem. Constraints

$$\hat{\mathbf{p}} = (\hat{\mathbf{H}}_s^T \hat{\mathbf{H}}_s)^{-1} \hat{\mathbf{H}}_s^T \hat{\mathbf{y}}_s$$

ρ rounds

ρ rounds

$\mathbf{y}$

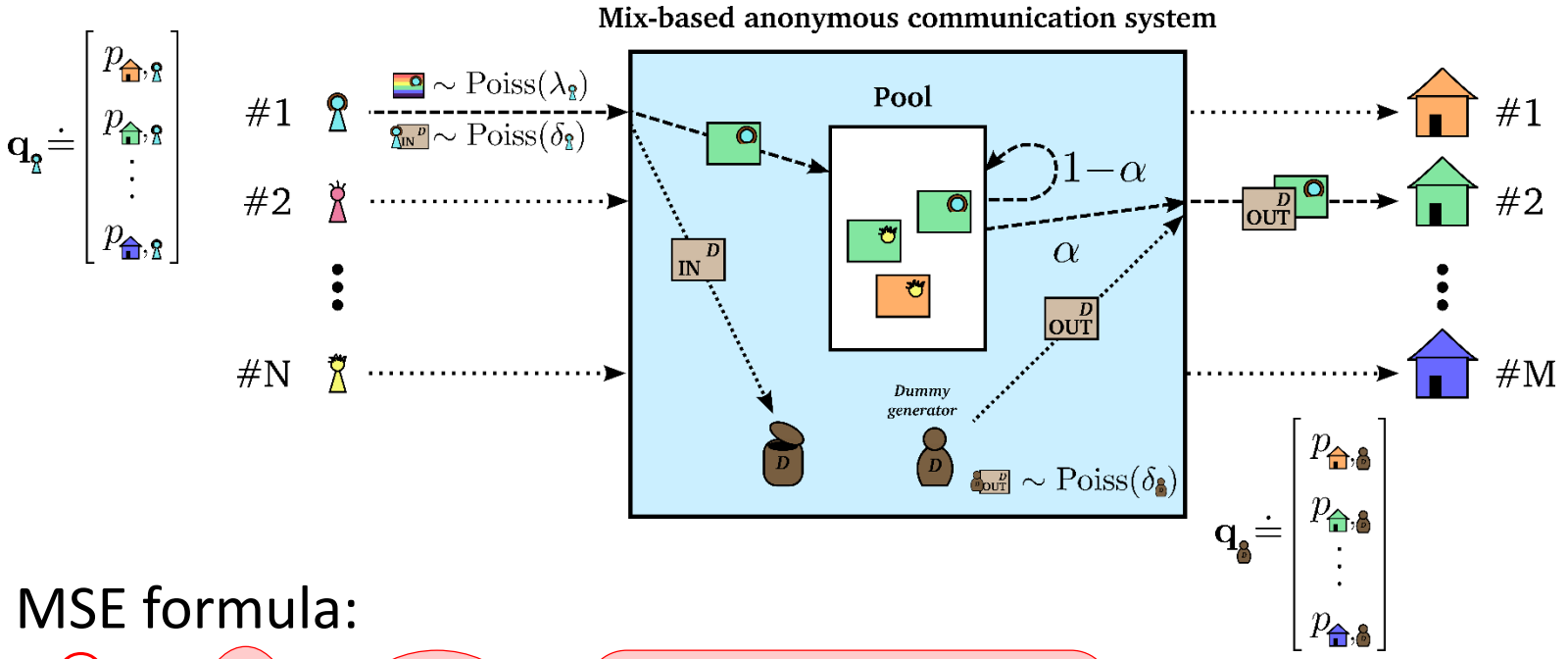
$\mathbf{H}$ → Pool + Dummies, $p_{\text{house},\text{person}}$:probabilities → $\mathbf{Y}$

$$\hat{\mathbf{H}}_s \leftarrow \mathbf{H} + \text{Pool} + \text{D IN D IN IN}$$

$$\hat{\mathbf{y}}_s \leftarrow \mathbf{y} + \text{D OUT D OUT OUT}$$

Performance Analysis. MSE.

- Derive an expression for: $\mathrm{MSE}_{\text{🏠},\text{🧍}} \doteq \left|\hat{p}_{\text{🏠},\text{🧍}} - p_{\text{🏠},\text{🧍}}\right|^2$
- Assumptions:



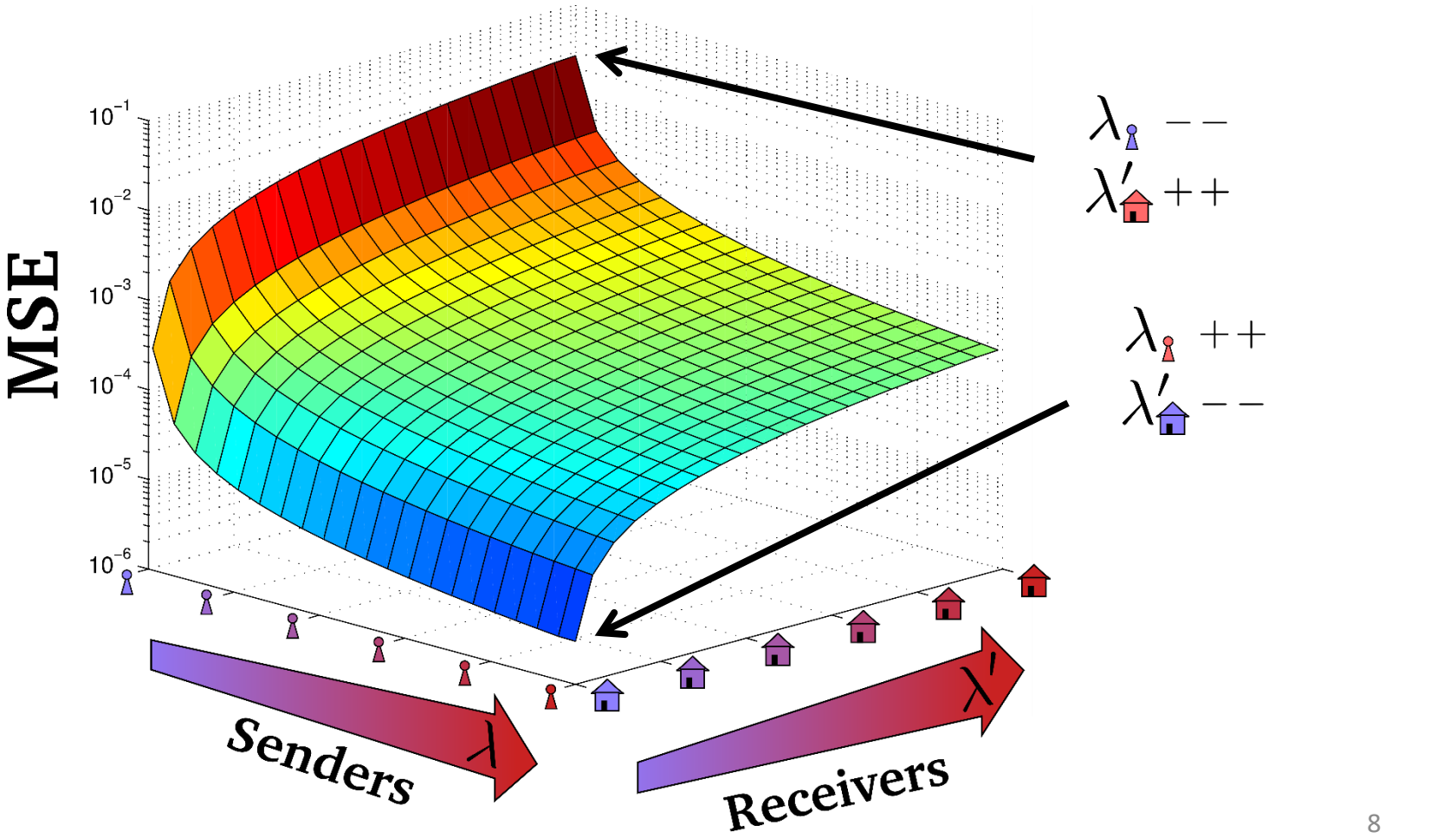
- MSE formula:

$$\mathrm{MSE}_{\text{🏠},\text{🧍}} \approx \frac{1}{\rho} \cdot \frac{2-\alpha}{\alpha} \cdot \frac{1}{\lambda_{\text{🧍}}} \cdot \left(1 + \frac{\delta_{\text{🧍}}}{\lambda_{\text{🧍}}}\right) \cdot \left(\lambda'_{\text{🏠}} + \delta_{\text{🎒}} \cdot p_{\text{🏠},\text{🎒}}\right)$$

Observations | Pool | Sender | Receiver

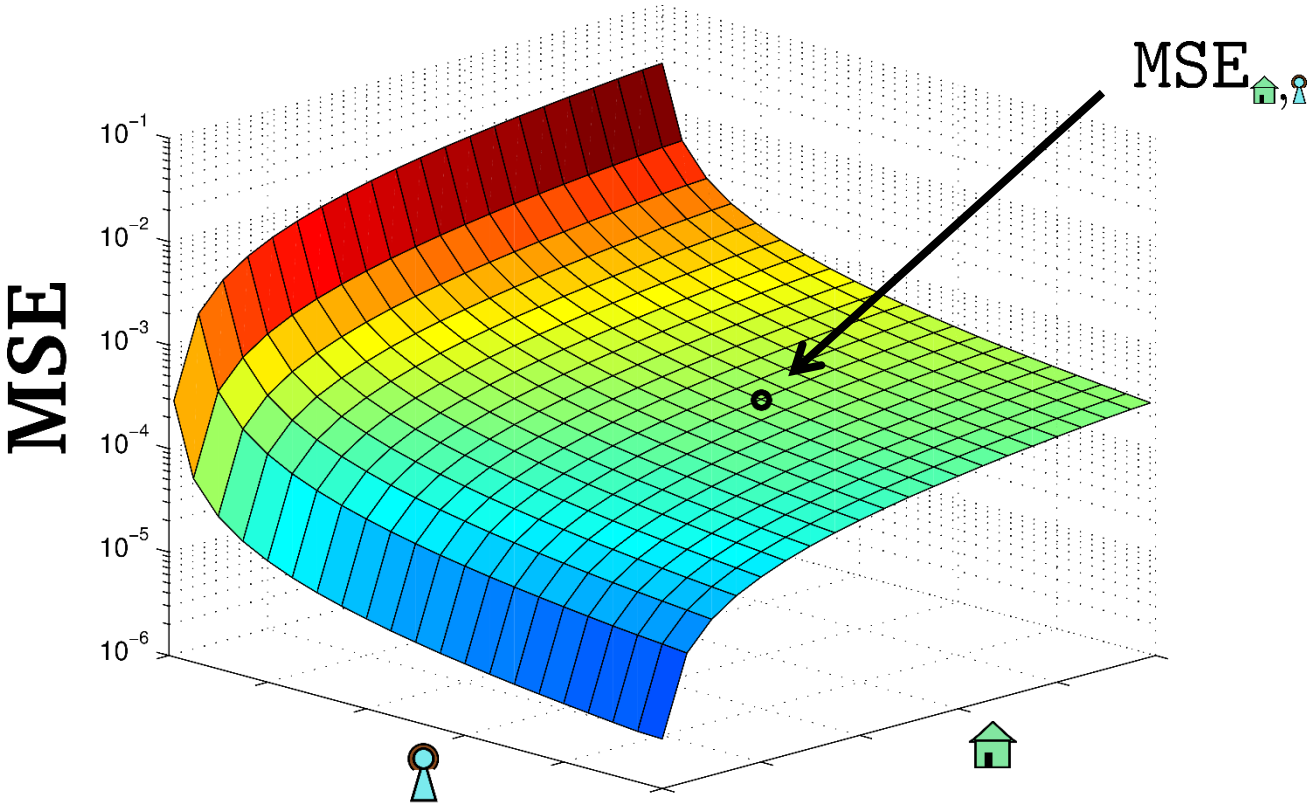
Performance Analysis. Interpretation (I).

$$\mathrm{MSE}_{\text{house},\text{person}} \approx \frac{1}{\rho} \cdot \frac{2-\alpha}{\alpha} \cdot \frac{1}{\lambda_{\text{person}}} \cdot \left(1 + \frac{\delta_{\text{person}}}{\lambda_{\text{person}}}\right) \cdot \left(\lambda'_{\text{house}} + \delta_{\text{person}} \cdot p_{\text{house},\text{person}}\right)$$



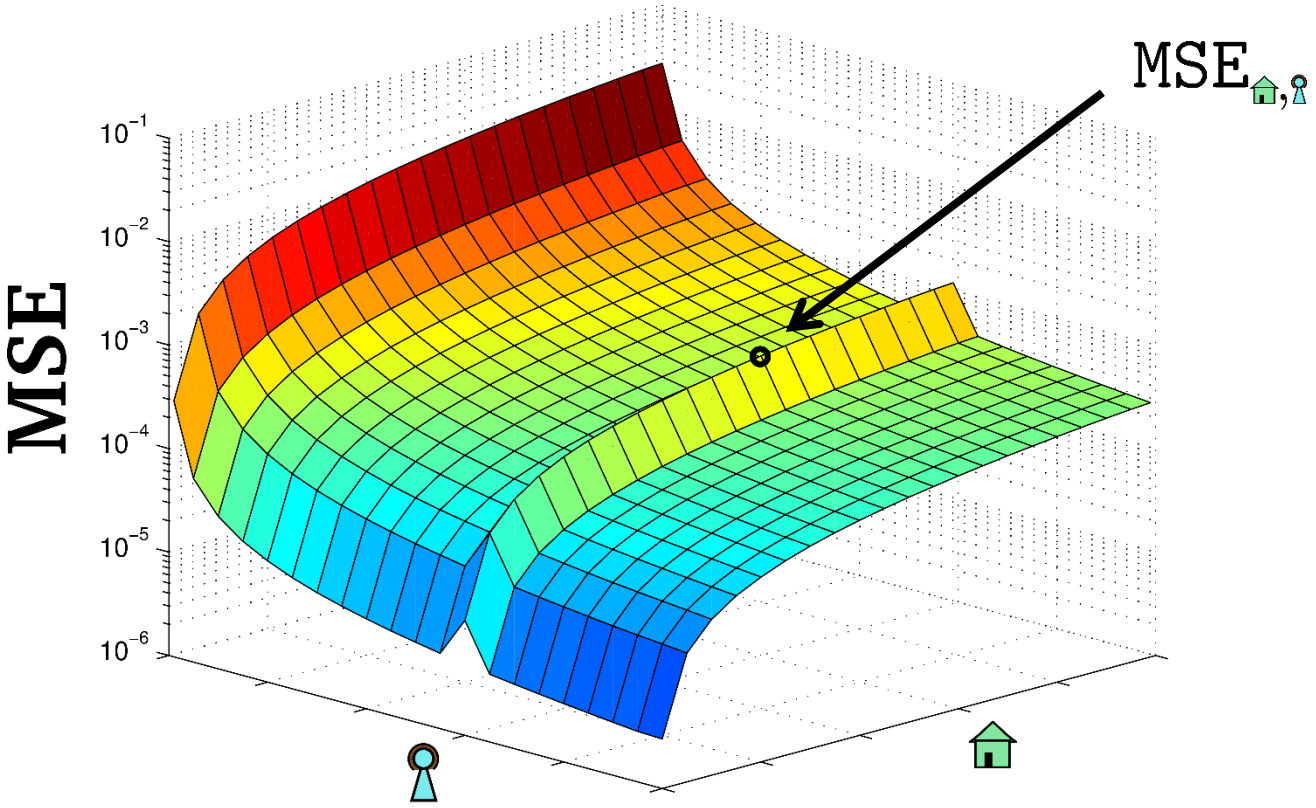
Performance analysis. Interpretation (II).

$$\mathrm{MSE}_{\text{house},\text{key}} \approx \frac{1}{\rho} \cdot \frac{2-\alpha}{\alpha} \cdot \frac{1}{\lambda_{\text{key}}} \cdot \left(1 + \frac{\delta_{\text{key}}}{\lambda_{\text{key}}}\right) \cdot \left(\lambda'_{\text{house}} + \delta_{\text{D}} \cdot p_{\text{house},\text{D}}\right)$$



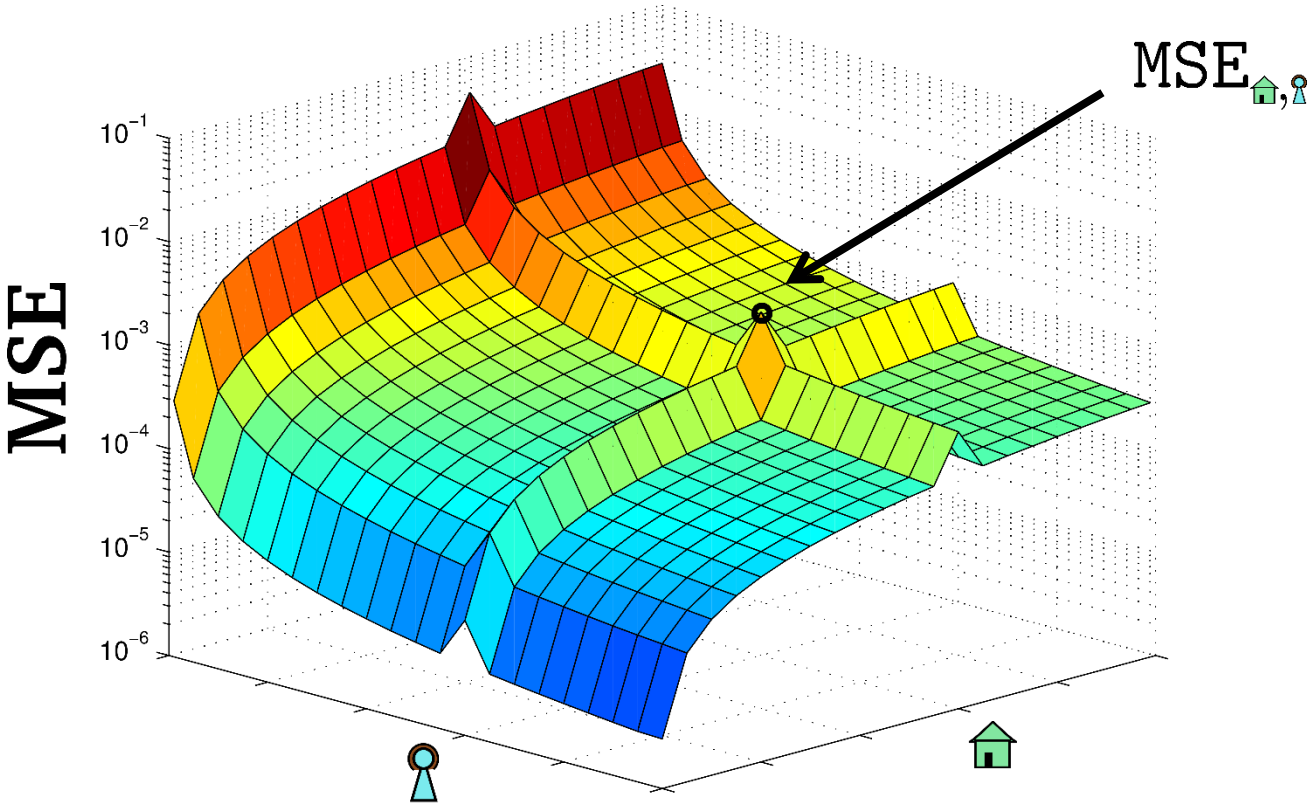
Performance analysis. Interpretation (II).

$$\mathrm{MSE}_{\text{house},\text{key}} \approx \frac{1}{\rho} \cdot \frac{2-\alpha}{\alpha} \cdot \frac{1}{\lambda_{\text{key}}} \cdot \left(1 + \frac{\delta_{\text{key}}}{\lambda_{\text{key}}}\right) \cdot \left(\lambda'_{\text{house}} + \delta_{\text{key}} \cdot p_{\text{house},\text{key}}\right)$$



Performance analysis. Interpretation (II).

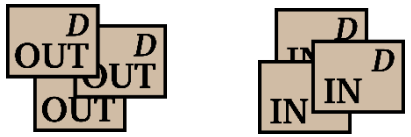
$$\mathrm{MSE}_{\text{house},\text{key}} \approx \frac{1}{\rho} \cdot \frac{2-\alpha}{\alpha} \cdot \frac{1}{\lambda_{\text{key}}} \cdot \left(1 + \frac{\delta_{\text{key}}}{\lambda_{\text{key}}}\right) \cdot \left(\lambda'_{\text{house}} + \delta_{\text{D}} \cdot p_{\text{house},\text{D}}\right)$$



Designing dummy traffic strategies.

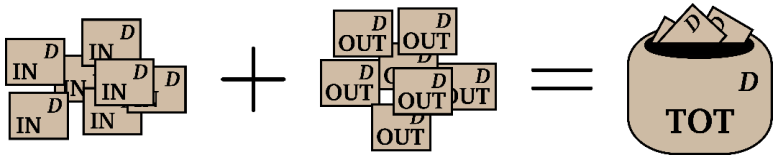
- We understand the effect of the dummies in the MSE.
- Privacy objective

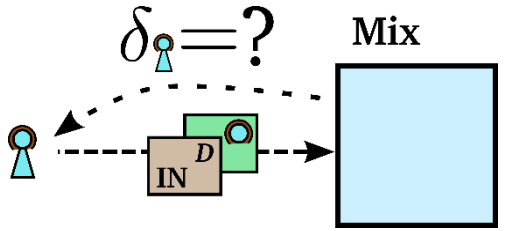
Dummy Strategy!



PICK YOUR FAVOURITE PRIVACY OBJECTIVE!!

- As an example: we pick two objectives.
- Assumptions:
  - Budget of dummies
  - Return channel





Dummy Strategy 1.

- **Objective**: increase the protection of all the relations by a constant factor.

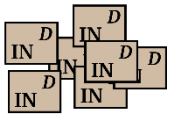
$$\underset{\delta_{\cdot}\,,\,p_{\cdot,\cdot}}{\text{maximize}} \quad \text{MSE}_{\cdot,\cdot} \qquad \forall \cdot,\cdot \in \cdot,\cdot$$

$$s.t. \quad \text{MSE}_{\cdot,\cdot} = \beta \cdot \text{MSE}^{0}_{\cdot,\cdot} \qquad \forall \cdot,\cdot \in \cdot,\cdot$$

Dummy Strategy 1.

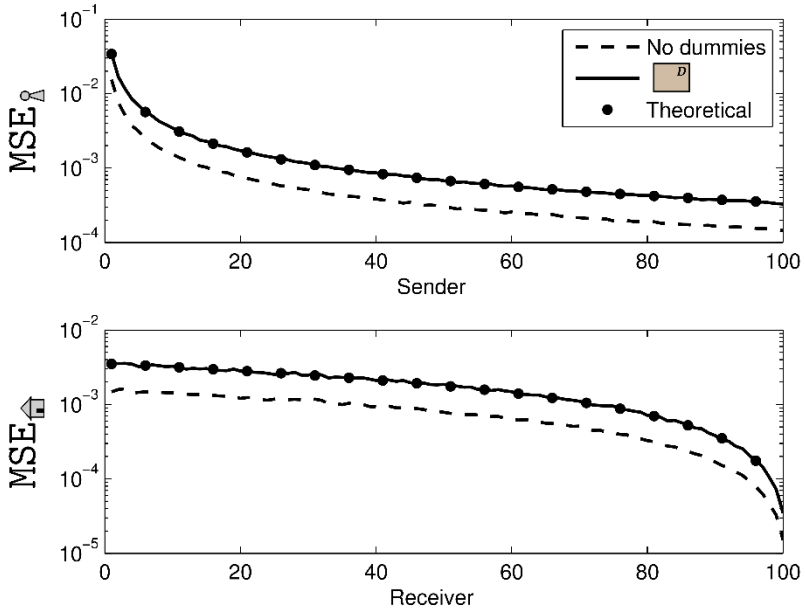
- **Solution**: assign dummies proportionally to message rates.

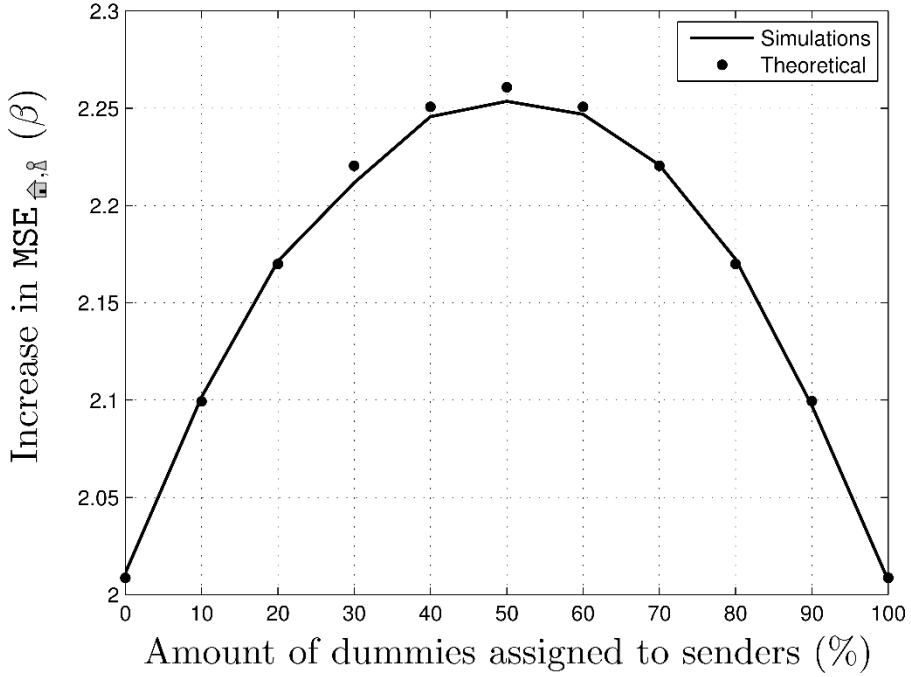
$$\delta_{\text{user}} = \frac{\lambda_{\text{user}}}{\sum_{\text{user} \in \text{group}} \lambda_{\text{user}}} \cdot$$



$$p_{\text{house}, D} = \frac{\lambda'_{\text{house}}}{\sum_{\text{house} \in \text{group}} \lambda'_{\text{house}}}$$

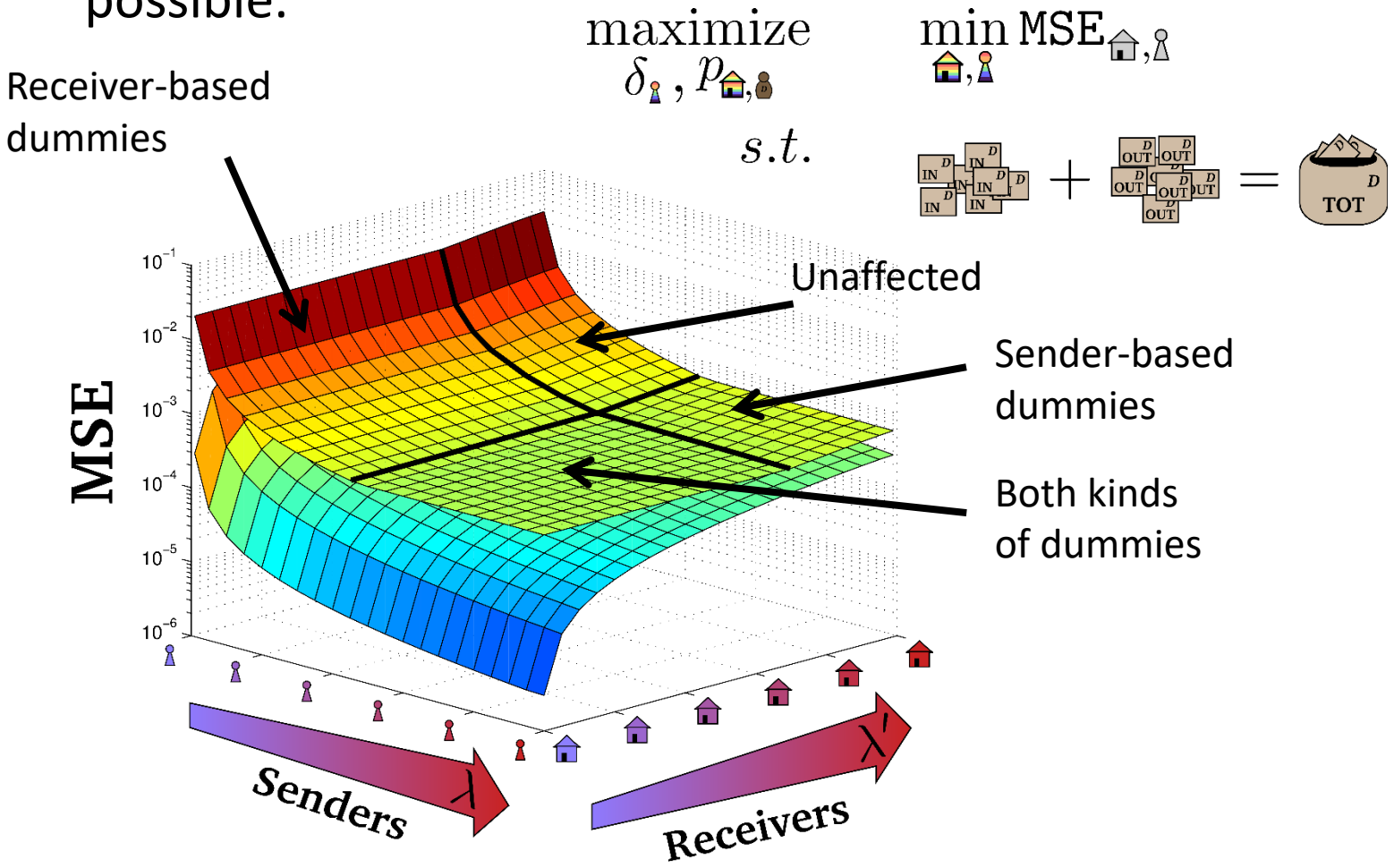
- Experimental results:





Design Strategy 2.

- **Objective**: provide a minimum protection as large as possible.



Dummy Strategy 2.

- **Solution**: waterfilling-like algorithm.

$$\mathcal{A}_i \doteq \{1, 2, \cdots, i\}$$

$$\tilde{\epsilon}_{s,\texttt{MIN}} = \frac{\delta_{\texttt{SEND}} + \sum_{k\in\mathcal{A}} \lambda_k}{\sum_{k\in\mathcal{A}} \lambda_k^2}$$

$$\frac{1}{\lambda_n} \leq \frac{\delta_{\texttt{SEND}} + \sum_{k\in\mathcal{A}_n} \lambda_k}{\sum_{k\in\mathcal{A}_n} \lambda_k^2} \leq \frac{1}{\lambda_{n+1}}$$

$$\delta_i = \lambda_i(\lambda_i \tilde{\epsilon}_{s,\texttt{MIN}} - 1) \text{ if } i \in \mathcal{A}_n$$

$$\mathcal{B}_j \doteq \{1, 2, \cdots, j\}$$

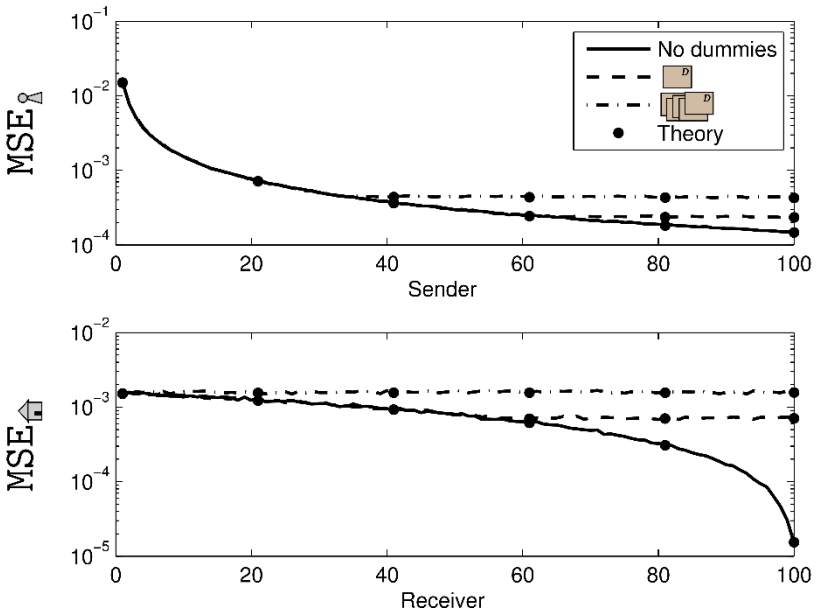
$$\tilde{\epsilon}_{r,\texttt{MIN}} = \frac{\delta_{\texttt{MIX}} + \sum_{j\in\mathcal{B}} \lambda'_j}{|\mathcal{B}|}$$

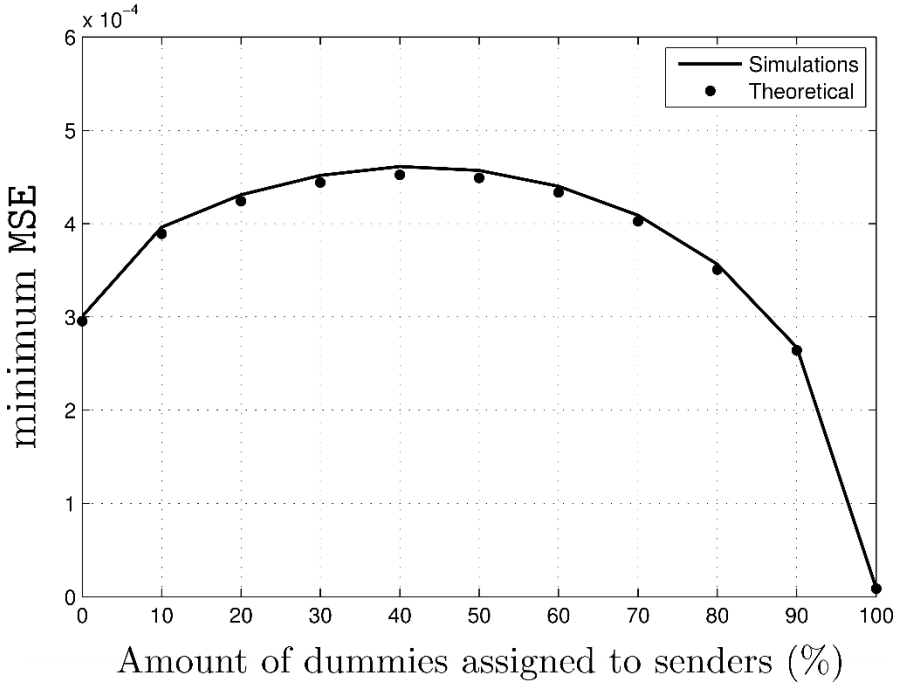
$$\lambda'_n \leq \frac{\delta_{\texttt{MIX}} + \sum_{j\in\mathcal{B}_n} \lambda'_j}{|\mathcal{B}_n|} \leq \lambda'_{n+1}$$

$$p_{j,\texttt{MIX}} = \frac{1}{\delta_{\texttt{MIX}}} \left(\tilde{\epsilon}_{r,\texttt{MIN}} - \lambda'_j\right), \text{ if } j \in \mathcal{B}_n$$

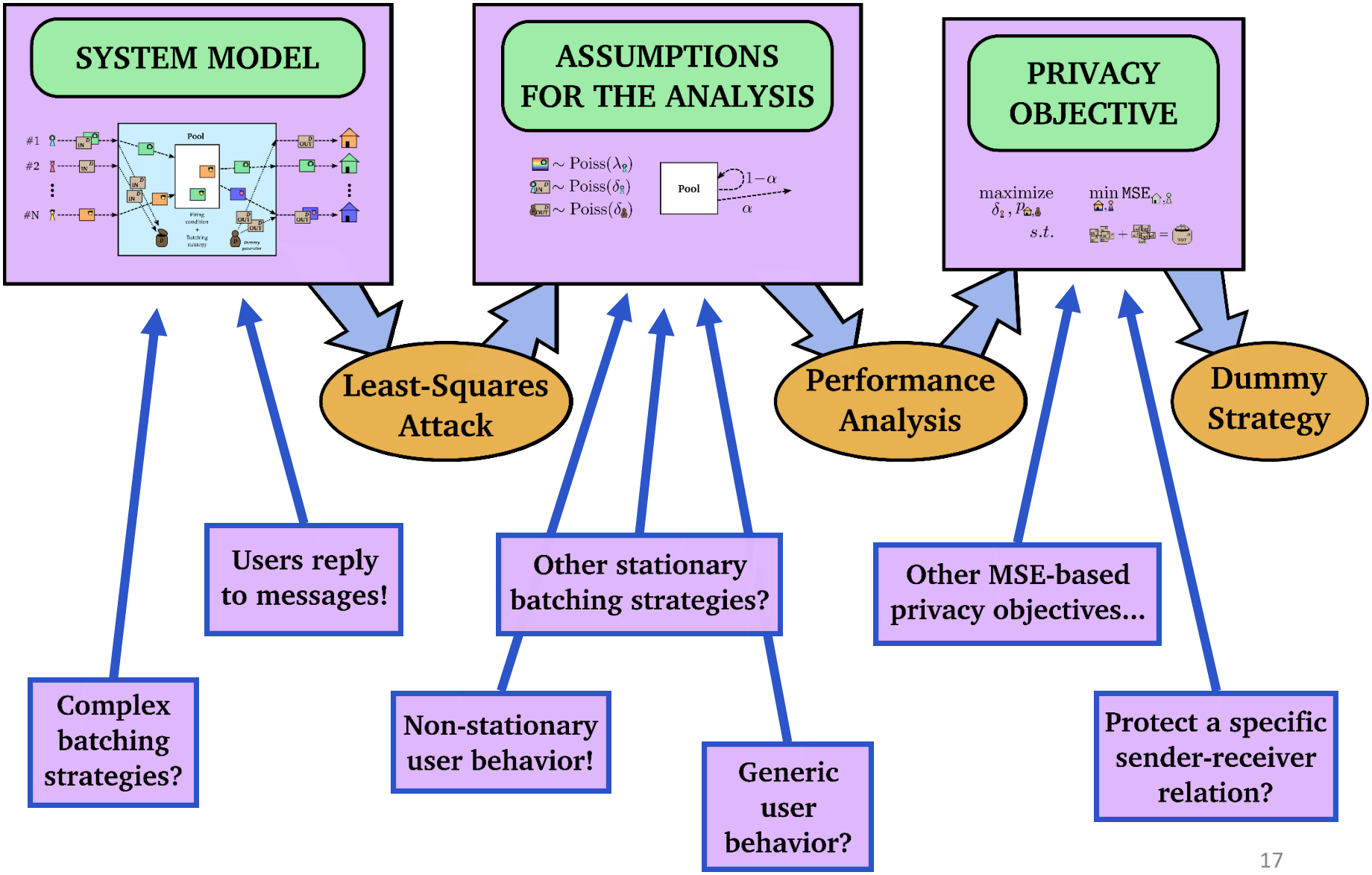
BORING!!!

- Experimental Results





Conclusions. Methodology.



TÁCTICA
ATLANTTIC

AtlantTIC
Research Center for
Information & Communication Technologies

Thanks!

simonoya@gts.uvigo.es

Universidade Vigo

GOBIERNO DE ESPAÑA
MINISTERIO DE ECONOMÍA Y COMPETITIVIDAD

FEDER
Fondo Europeo de Desarrollo Regional
UNIÓN EUROPEA
"Una manera de hacer Europa"