How to Form Bags in Batch Steganography

Eli Dworetzky, Jessica Fridrich

IEEE WIFS 2024

BINGHAMTON
UNIVERSITY
STATE UNIVERSITY OF NEW YORK

Batch steganography

Alice / steganographer

- Spreads payload across multiple covers (or **bag** of images)
- Useful when payload cannot fit within one image

Warden / Eve

- Inspects the **entire** bag (pooled steganalysis)
- Applies a single-image detector (SID) to all images
  - E.g. SRNet, rich models, etc.
- "Pools" the SID outputs to decide if the **bag** is cover or stego

Practical question: how to form the bag?

Alice has a **fixed payload** to communicate

1. What kind of images should she use?
2. How many images should she use?

Alice uses source biasing

Source biasing

- Alice samples covers from the cover source with a bias towards harder-to-steganalyze images
- Alice gains security when biasing optimally
- Warden tests for a deviation in cover source by considering the **joint** statistical impact of steganography and biasing

E. Dworetzky, E. Kaziakhmedov, J. Fridrich, "Improving Steganographic Security with Source Biasing", 12th IH&MMSec. Workshop, Vigo, Spain, June 24-26, 2024.

- We will utilize the model proposed in this prior work
- Prior work studied **fixed rate** – we now have a **fixed payload**

Alice's goal

Given a fixed payload, Alice wants to choose bag size n so that

$$P_{\mathrm{D}} \leq \widetilde{P}_{\mathrm{D}} \text{ and } n \leq n_{\max}$$

- $\widetilde{P}_{\mathrm{D}}$ — maximal tolerable detectability by Warden's (pooled) detector
- $n_{\max}$ — bandwidth limit
- Alice's optimal source biasing strength is a function of n

Outline

- Formal setup of batch steganography / pooled steganalysis
- High level overview of the model
- Two benefits of source biasing: **Bias gain** and **bandwidth savings**
- Confirmation by experiments on ALASKA II dataset
    - Bias gain and bandwidth savings observed in practice

Batch stego setup

Alice

- Payload is αC bits, $\alpha > 0$, where C is capacity of each cover
    - Ternary embedding: $C = \log_2 3 \times$ number of pixels
- $\lceil \alpha \rceil$ is the smallest number of images needed to fit the payload
- Independently samples a bag of n covers of fixed size from $\mathcal{X}$, $\mathbf{X} = (X_1, \ldots, X_n)$, $n \geq \alpha$
- Embeds $\alpha_i C$ bits in X_i, $\sum_{i=1}^{n} \alpha_i = \alpha$, $0 \leq \alpha_i \leq 1$
- Her spreading strategy determines the α_i

Warden

- Has a SID $d : \mathcal{X} \to \mathbb{R}$ and a pooler $\pi : \mathbb{R}^n \to \mathbb{R}$
- Given a bag of n images $\mathbf{Y} = (Y_1, \ldots, Y_n)$, Warden's detection statistic is $\pi(d(Y_1), \ldots, d(Y_n))$

Detector-centric approach

We model the effect of embedding and model the source itself through soft outputs of the SID d

- Permits formulating steganalysis and source biasing **jointly** through a single hypothesis test
  - Closed-form ROC of Warden's optimal pooler (LRT)
- Model parameters can be estimated in practice
- We observe a close match between model and experiments on real datasets

Source model and biasing

Alice's cover source $\mathcal{X}$ has only two types of images: Hard & Easy

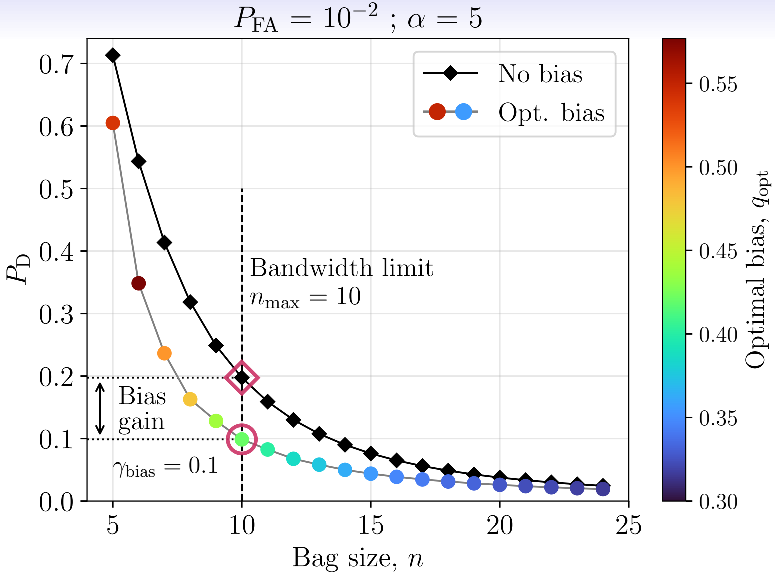
When sampling from $\mathcal{X}$ (no bias), for each i

- X_i is hard with probability p
- X_i is easy with probability $1 - p$

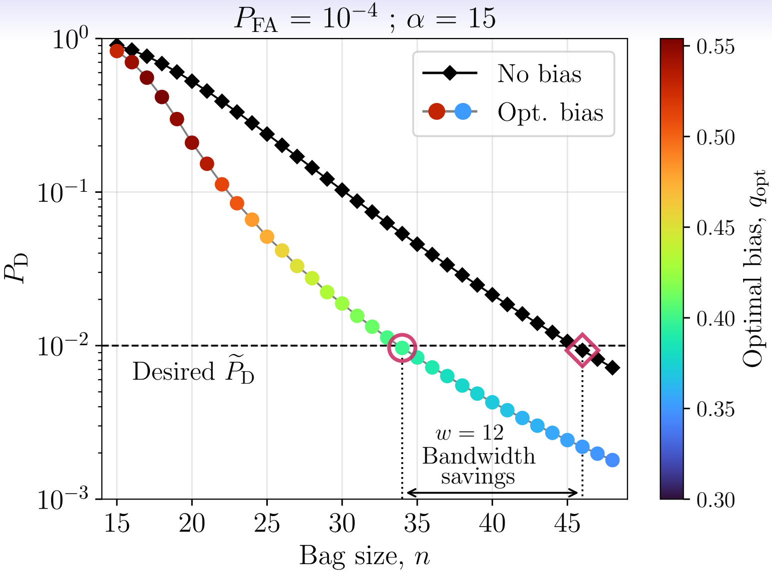
Alice's bias

- Selects hard images with probability $q \geq p$

Insight from the model: bias gain



Insight: bandwidth savings



Biasing and spreading in practice

Given bag $(X_1, \ldots, X_n)$

- Alice estimates the difficulty of each image by seeing how her own detector $d^{(\mathrm{A})}$ reacts to embedding

Biasing:

- Done by inverse transform sampling modified with a parametric model with parameter q
- $q = 1$ corresponds to unbiased sampling, $q > 1$ biased

Greedy spreader:

- Orders images by difficulty
- Starting with the most difficult, she embeds fully with HILL one by one
- α images will be embedded, $n - \alpha$ will be empty

Experiments on ALASKA II

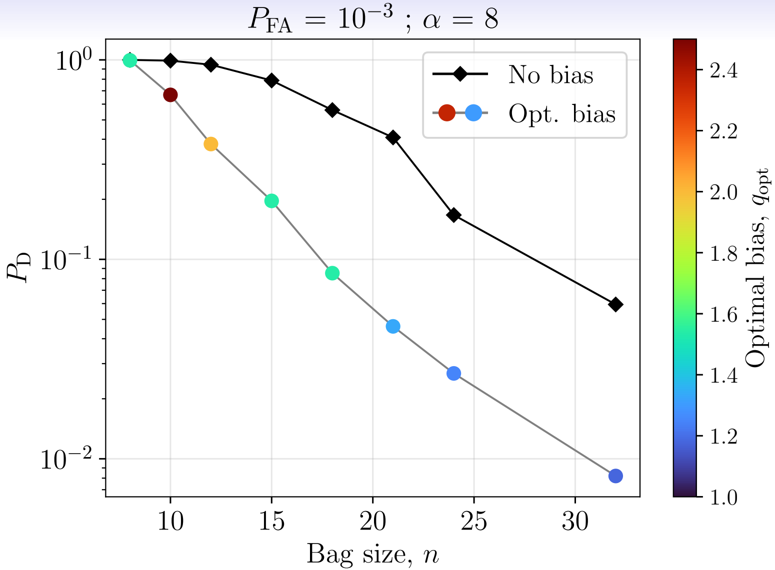
ALASKA II (75k grayscale images) divided into four subsets

- 23k used for training Alice's detector
- 23k used for training Warden's detector and 10k for pooler
- 19k used for evaluation
- Both detectors SRNets, JIN pre-trained, refined on HILL with uniform payloads on $[0, \log_2 3]$

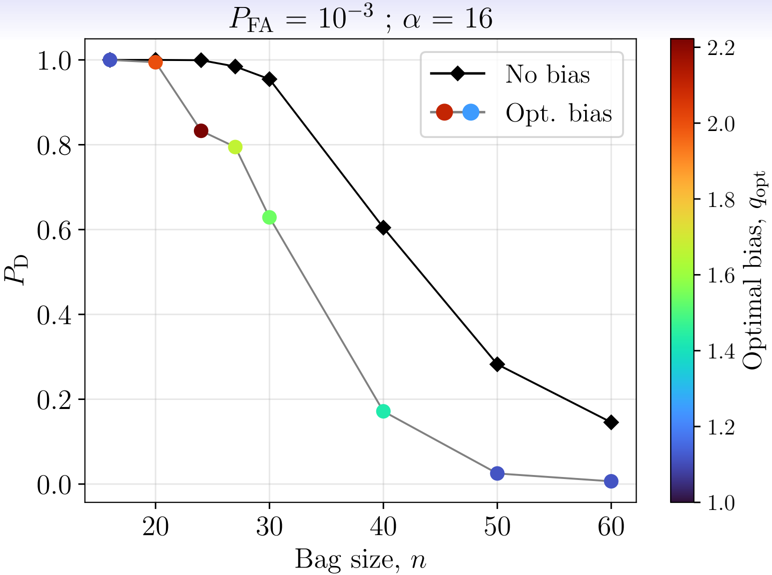
Warden's Pooler

- Trained as random forest on $2n + 2$ dim. feature extracted from bag $(X_1, \ldots, X_n)$

Experiments on ALASKA II



Experiments on ALASKA II



Conclusions

Selecting covers with an optimal bias, Alice benefits in terms of

- Lower detection probability P_{D} at the same bag size n
- Smaller bag size n at the same P_{D}

What kind of images should Alice use?

- Difficult cover images...but not a suspiciously high ratio of them

How many images should Alice use? Depending on her preferences

- As many as she can within her bandwidth constraint
- Just enough to achieve a desired detectability if bandwidth is costly
- ...or somewhere in between