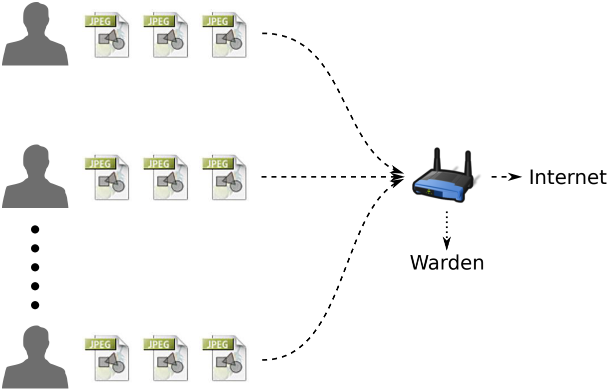
The Challenges of Rich Features in Universal Steganalysis

Tomáš Pevný[a] and Andrew D. Ker[b]

[a]Agent Technology Center, Czech Technical University in Prague, Czech Republic.
[a]Department of Computer Science, Oxford University, England.

7th February 2013

Batch universal steganalysis



Batch universal steganalyzer

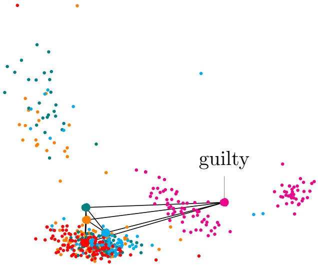
- Extract features.
- Calculate distances between actors (MMD).
- Identify the steganographer(s). local outlier factor (LOF)

Batch universal steganalyzer

- Extract features.
- Calculate distances between actors (MMD).
- Identify the steganographer(s).
  local outlier factor (LOF)

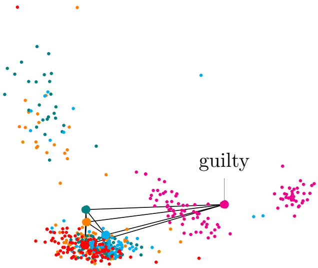
Batch universal steganalyzer

- Extract features.
- Calculate distances between actors (MMD).
- Identify the steganographer(s).
  local outlier factor (LOF)



Batch universal steganalyzer

- Extract features.
- Calculate distances between actors (MMD).
- Identify the steganographer(s).
  local outlier factor (LOF)



The method should work with any stego-sensitive features.

Accuracy with PF274 and CF^* features

	PF274	CF^*
dimension	274	8750
F5	14.6	9.5
nsF5	10.7	23.1
JP Hide&Seek	7.8	16.2
OutGuess	1.9	5.7
Steghide	2.8	4.7

Average rank of one guilty actor (out of 100)
emitting payload 0.1 bpnc

Curse of dimensionality

- Anomaly detection estimates density: more difficult in high dimensions.
- In unsupervised learning cannot discard noise in features.

Curse of dimensionality

Our solution

Supervised dimensionality reduction.

Our aim

Steganographic features should be sensitive to embedding changes, yet insensitive to image content. J. Fridrich, 2004

Dimensionality reduction

Prior art

- Principal component transformation
- Maximum covariance
- Ordinary least square regression

Proposed

- Calibrated least-squares

Principal component transformation (PCT)

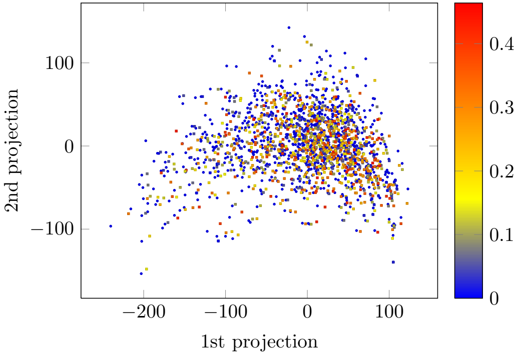
$$\arg\max_{w_k} \operatorname{Var}(\mathbf{X} w_k)$$

subject to

$$w_k \perp w_i,\ i \in \{1, \ldots, k-1\}.$$

$\mathbf{X} \in \mathbb{R}^{n,d}$ — matrix with features

w_i — projections found

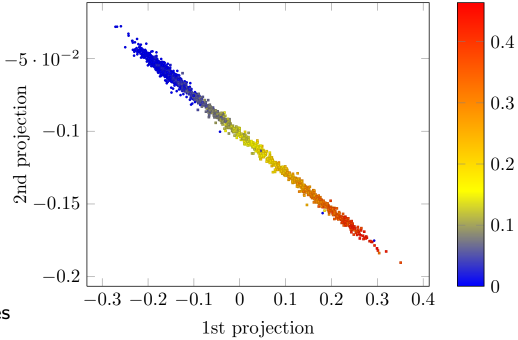


Ordinary least square regression (OLS)

$$\arg\max_{w_k} \mathrm{Cov}(\mathbf{X}^s w_k, \mathbf{Y}^s) - \mathrm{Var}(\mathbf{X}^s w_k)$$

subject to

$$w_k \perp w_i,\ i \in \{1, \ldots, k-1\}.$$



$\mathbf{X}^S \in \mathbb{R}^{n,d}$ — matrix with stego features

$\mathbf{Y}^s \in \mathbb{R}^{n,1}$ — vector with payload

w_i — projections found

Maximum covariance (MCV)

$$\arg\max_{w_k} \operatorname{Cov}(\mathbf{X}^s w_k, \mathbf{Y}^s)$$

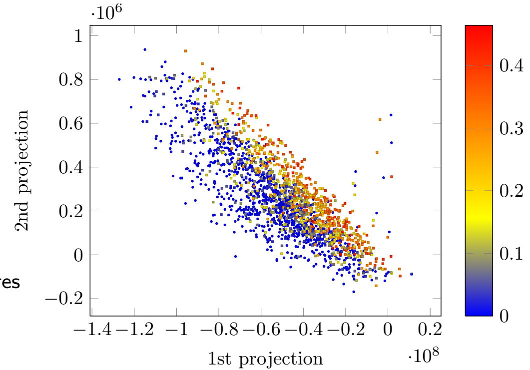
subject to

$$w_k \perp w_i,\ i \in \{1, \ldots, k-1\}.$$

$\mathbf{X}^S \in \mathbb{R}^{n,d}$ — matrix with stego features

$\mathbf{Y}^s \in \mathbb{R}^{n,1}$ — vector with payload

w_i — projections found

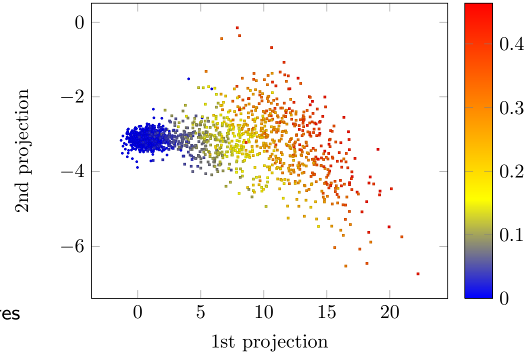


Calibrated least squares (CLS)

$$\arg\max_{w_k} \mathrm{Cov}(\mathbf{X}^s w_k, \mathbf{Y}^s) - \mathrm{Var}(\mathbf{X}^c w_k)$$

subject to

$$w_k \perp w_i,\ i \in \{1, \ldots, k-1\}.$$



$\mathbf{X}^S \in \mathbb{R}^{n,d}$ — matrix with stego features

$\mathbf{Y}^s \in \mathbb{R}^{n,1}$ — vector with payload

$\mathbf{X}^c \in \mathbb{R}^{n,d}$ — matrix with cover features

w_i — projections found

Experimental settings

- 3000 users of leading social network, 100 images from each
  - 1000 users for supervised feature reduction
  - 2000 users used for testing
- Guilty actor emits payload 0.1 bpnc
  - linear (in the paper) or greedy strategy
  - one of following algorithms:
    F5, nsF5, JPHide&Seek (JP), OutGuess (OG), Steghide (SH)
- Steganalyst uses reduced CF^* features.
- Accuracy is measured by average rank of guilty actor.
  - $1.0 =$ perfect, $50.5 =$ random guessing.

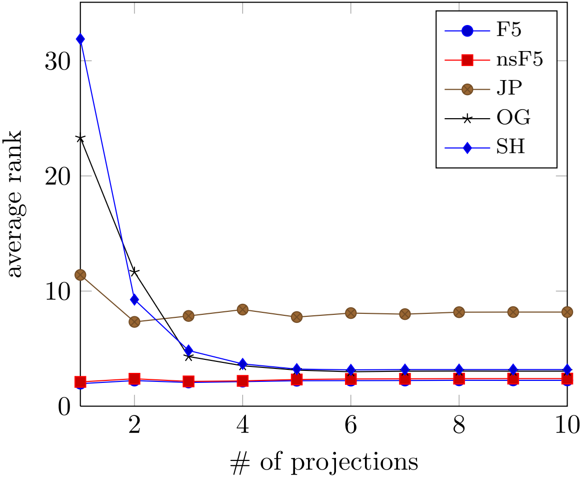
Results

	PCT	MCV	OLS	CLS
F5	40.3	23.4	22.2	1.6
		(4)	(1)	(1)
nsF5	38.0	26.6	5.8	2.1
		(4)	(1)	(1)
JP	38.4	27.2	6.9	1.7
		(5)	(1)	(1)
OG	26.5	31.6	2.4	1.2
		(4)	(1)	(1)
SH	23.0	2.6	1.3	1.1
		(6)	(1)	(1)

Robustness

	PCT	CLS trained on				
		F5	nsF5	JP	OG	SH
F5	40.3	1.6	1.9	8.8	6.6	4.5
		(1)	(1)	(1)	(4)	(3)
nsF5	38.0	1.8	2.1	10.1	10.9	10.5
		(1)	(1)	(1)	(4)	(3)
JP	38.4	8.9	7.2	1.7	15.5	10.5
		(1)	(2)	(1)	(2)	(2)
OG	26.5	3.7	3.0	11.8	1.2	1.1
		(1)	(6)	(2)	(1)	(1)
SH	23.0	5.2	3.2	9.1	1.2	1.1
		(1)	(6)	(2)	(1)	(1)

Optimal number of projections



Conclusion

- High dimensional features are not compatible with unsupervised steganalysis.
- Investigated dimensionality reduction to improve SNR of rich features.
- Validated the approach in universal batch steganalysis.
- The proposed method, CLS, exhibits robustness to embedding method.

F5 phenomenon