Template Attacks vs. Machine Learning Revisited

(and the Curse of Dimensionality in Side-Channel Analysis)

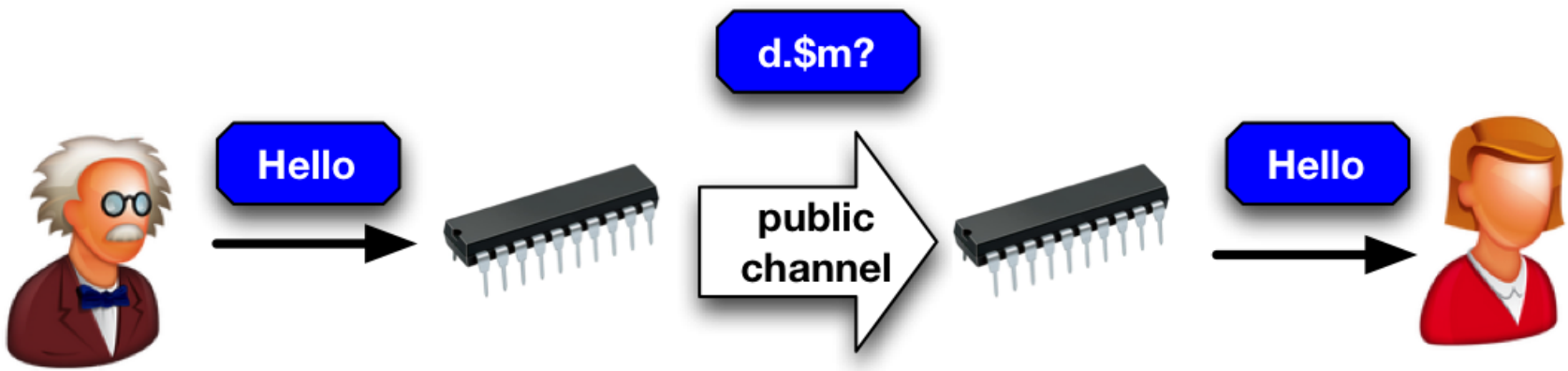
L. Lerman[1,2], R. Poussier[3],
G. Bontempi[2], O. Markowitch[1] and F-X. Standaert[3]

[1]Université libre de Bruxelles
Cryptography and Security Service
[2]Université libre de Bruxelles
Machine Learning Group
[3]Université catholique de Louvain
UCL Crypto Group

Cryptographic device

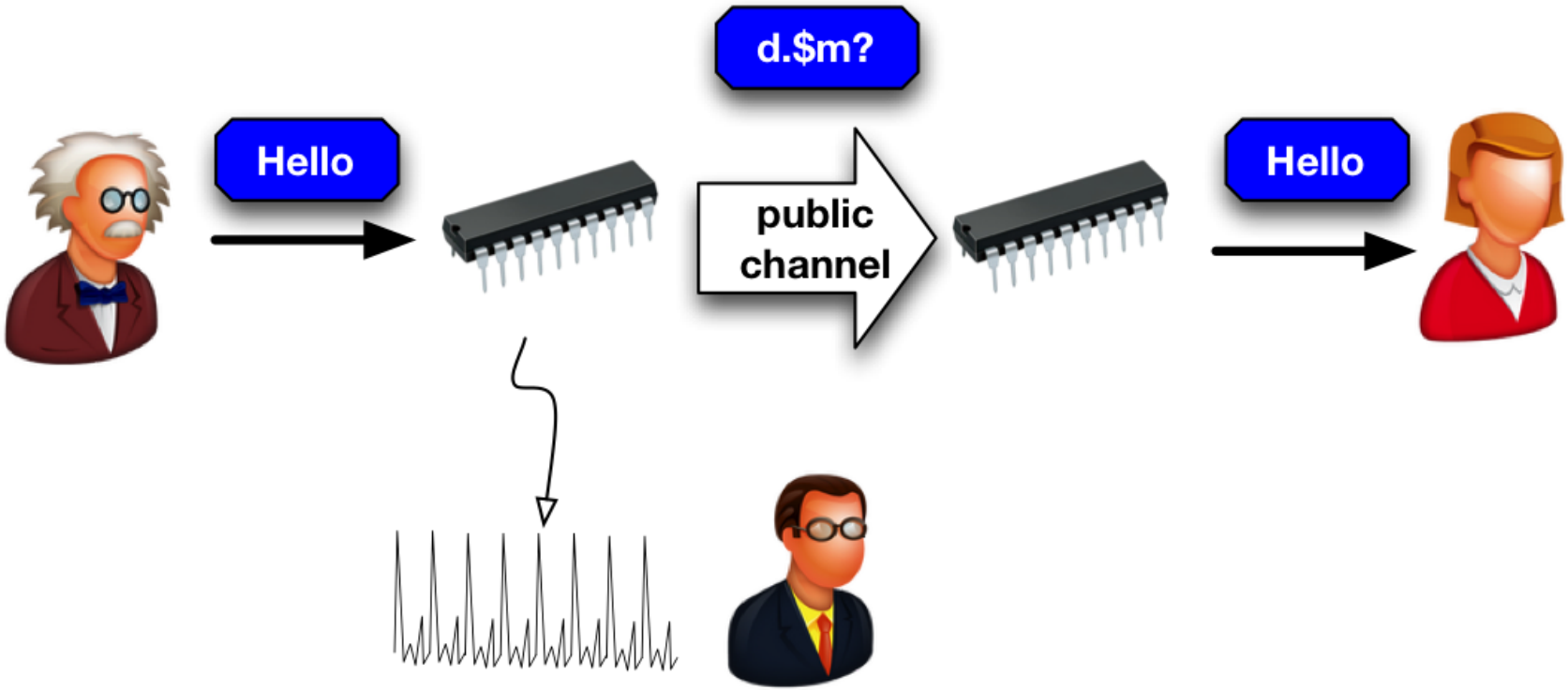
Encryption algorithm are widely used today

Black box model



A set of plaintexts and ciphertexts in order to find the secret key

Grey box model



Physical leakages include
power consumption and electromagnetic emanation

Profiling attack

- Purpose: Evaluating the security level of crypto implementations
- Profiled attacks estimate the worst-case security level in two steps
  - Profiling phase: leakage model estimation
  - Attacking phase: secret key extraction
- Several methods: parametric and non-parametric methods

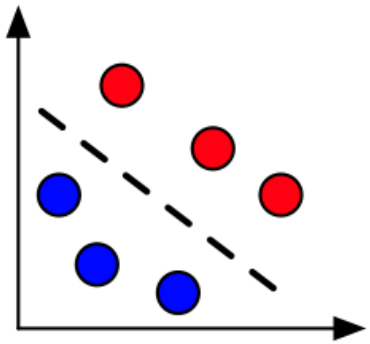
Template attack

- Assumption:
  - $\Pr(l \mid f(x,k)) \sim \mathcal{N}(\mu_{x,k}, \Sigma_{x,k})$
- Profiling phase: estimation of $\Sigma_{x,k}$ and $\mu_{x,k}$
- Attacking phase:
  $\hat{k} = \arg\max_k \hat{\Pr}(l \mid f(x,k)) \times \hat{\Pr}(f(x,k))$
  $\hat{k} = \arg\max_k \prod_{i=1}^{n} \hat{\Pr}(l_i \mid f(x_i,k)) \times \hat{\Pr}(f(x_i,k))$
- Efficient Template Attack:
  - $\Sigma_{x,k} = \Sigma \quad \forall x,k$

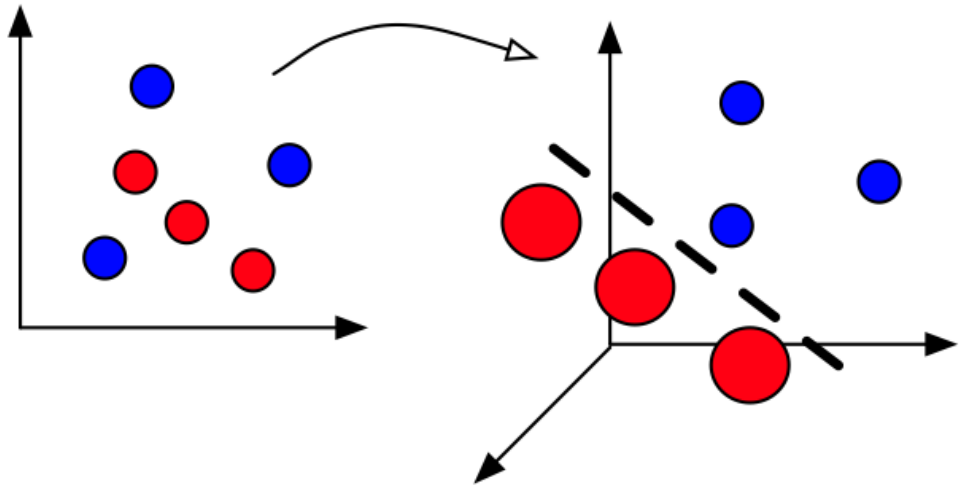
Machine Learning based attacks

- Non-parametric classifiers
  - Support Vector Machine
  - Random Forest

Support Vector Machine



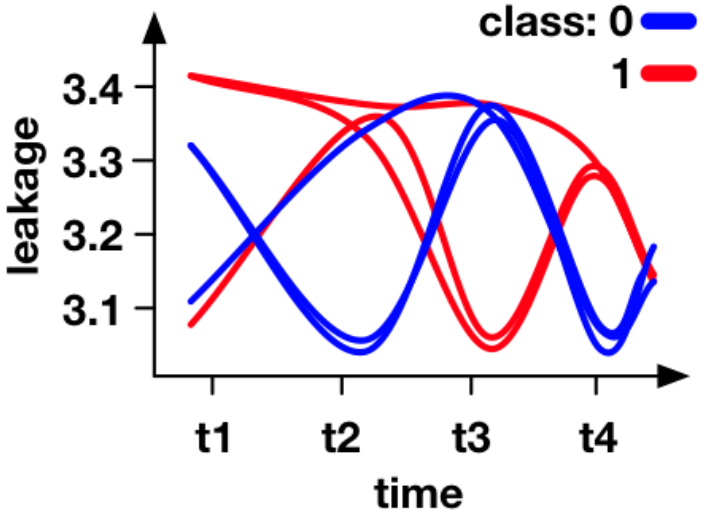
Linear classifier

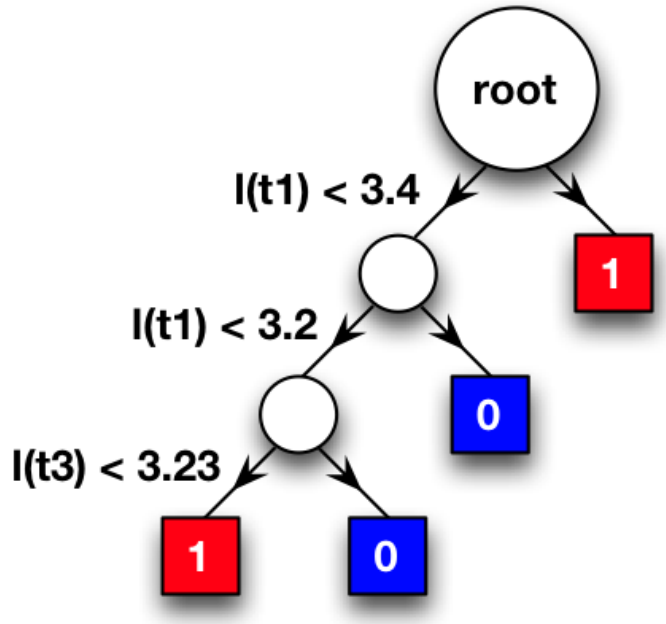


Non-linear classifier

SVM estimates a hyperplane saparating two classes
with the largest possible margin

Random Forest



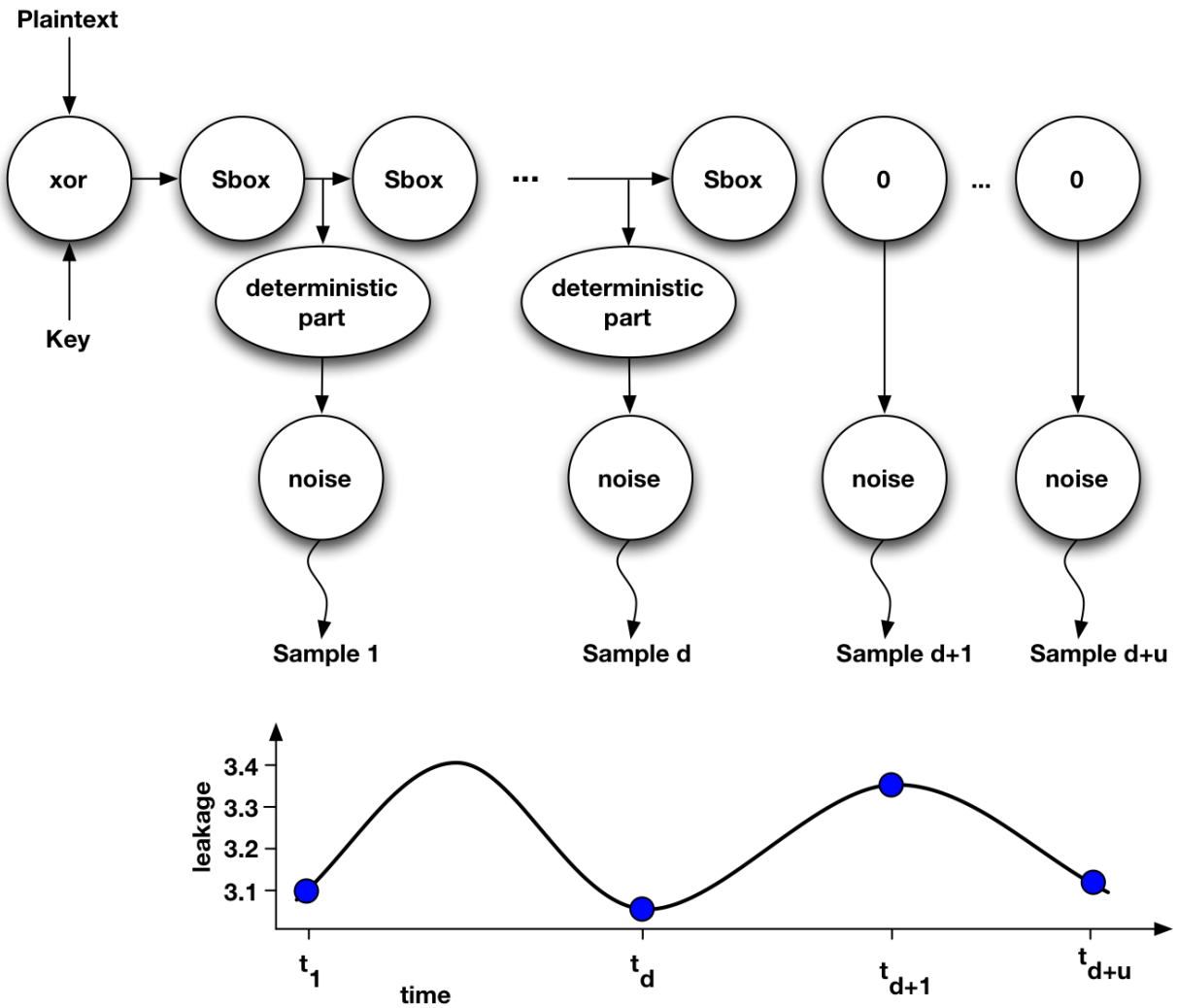


Random forest is a set of decision trees

Contributions

- Previous works claim that Machine Learning based attacks lead to successful key recoveries in experimental studies
- Our purpose: a systematic investigation of the conditions under which ML-based attacks outperform TA by
  - varying the number of traces in the profiling set
  - varying the number of traces in the attacking set
  - varying the number of informative points per trace
  - varying the number of uninformative points per trace
  - varying the signal-to-noise ratio

Leakage simulator



Evaluation metrics

- Information theoretic metric vs security metric
- Mutual Information between the key and the leakage based on a profiling model
  - Probability based profiling attacks (e.g. TA, SA)
- Success rate of a profiling model
  - Scoring and probability based profiling attacks (e.g. RF, TA)

Scenarios

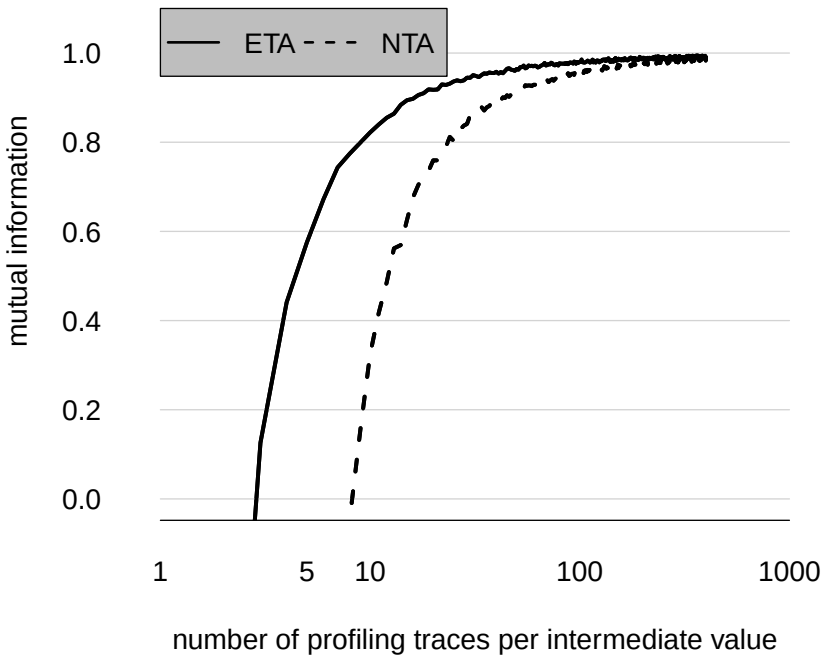
- In our setting, imperfections come from estimation errors
- Future work: imperfect models come from assumption errors
- Three scenarios:
  - Perfect profiling
  - Nearly perfect profiling
  - Imperfect profiling

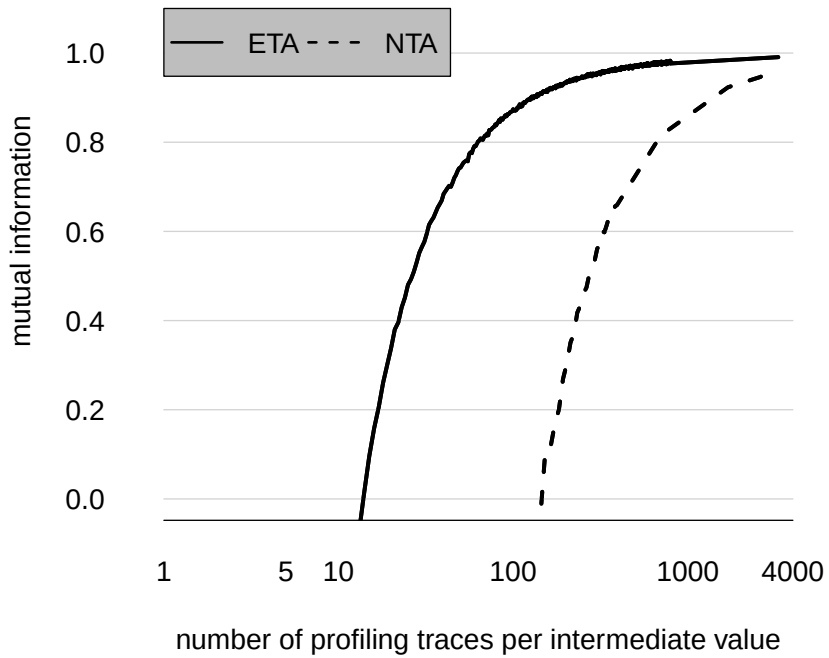
Perfect profiling

- **Proposition:** Let us assume two template attacks with perfect models using two different traces l_1 and l_2 associated to the same plaintext x: l_1 is composed of d samples providing information and $l_2 = [l_1 \parallel \epsilon_1, \epsilon_2, \ldots, \epsilon_u]$ (where the ϵ_i's are independent noise variables). Then the mutual information leakage estimated with their (perfect) leakage models is the same. *(proof in the paper)*
- **In brief:** Noise variables do not impact the mutual information for TA in an asymptotical context.
- **As a result:**
  - Feature selection algorithms are useless
  - ML-based approaches cannot be more efficient in this context

Nearly Perfect profiling

u=0 vs u=15

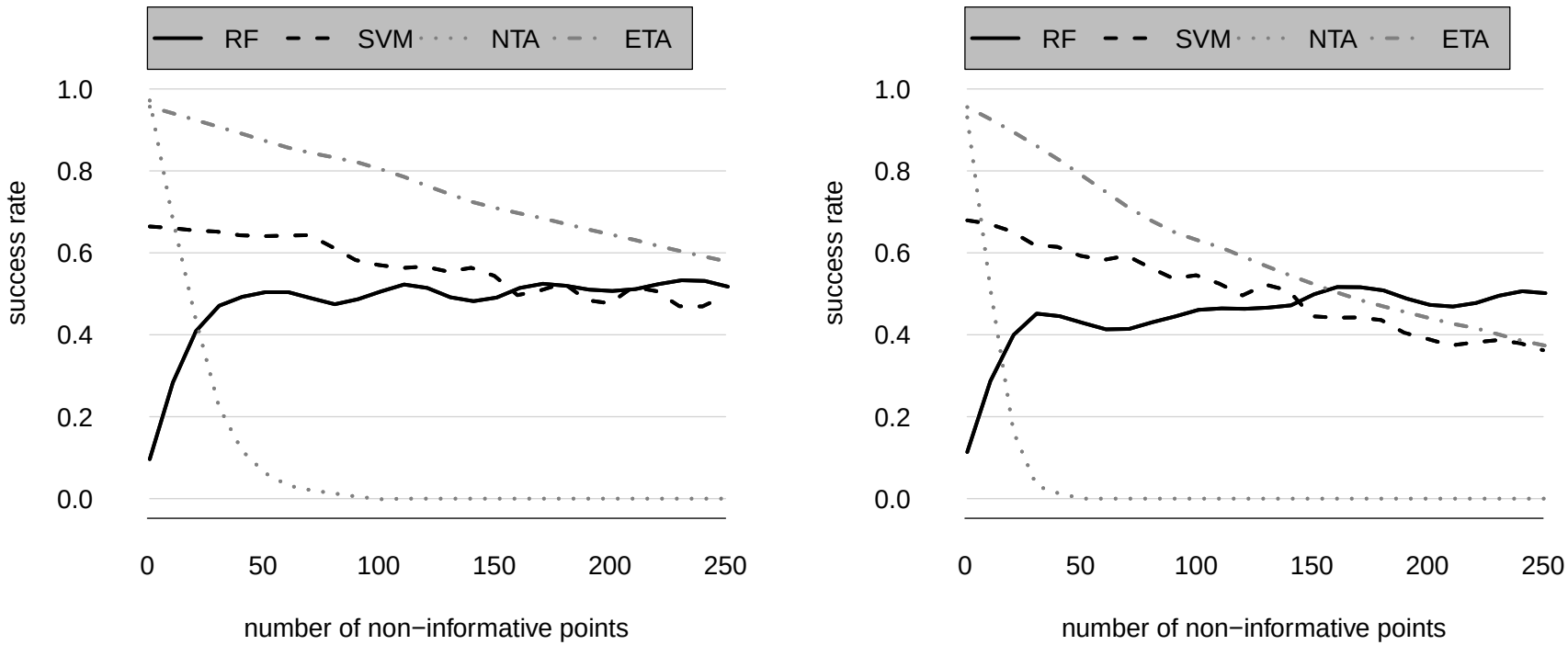




Two informative points, 0 and 15 useless samples, and a SNR of 1

Imperfect profiling

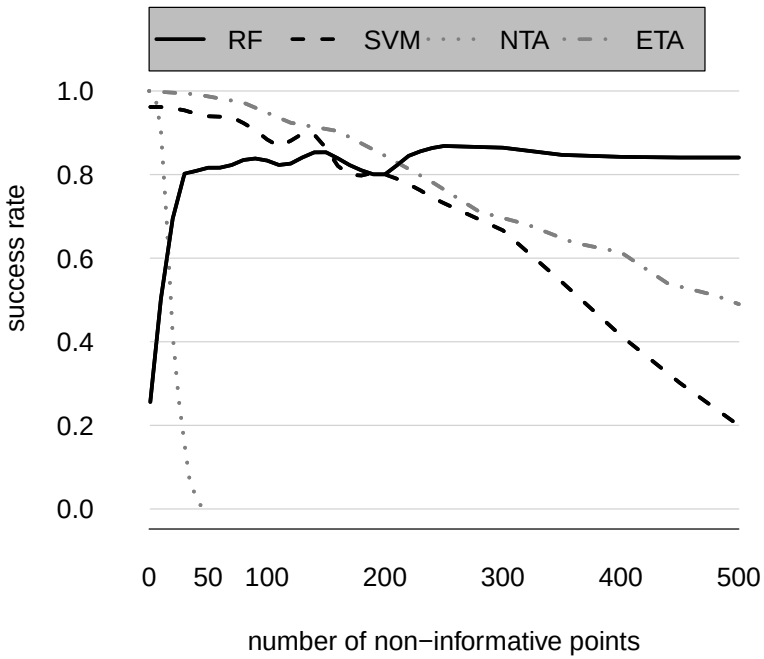
Np=100 vs Np=50



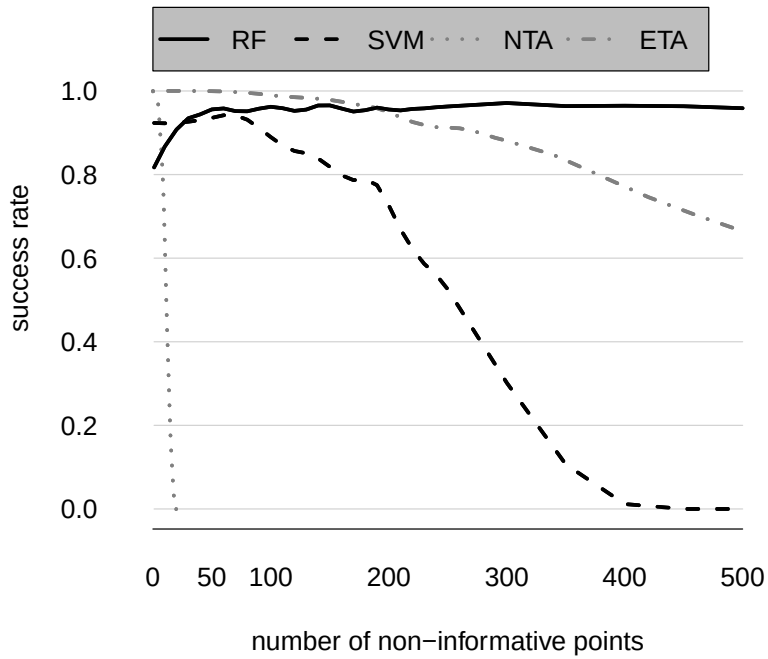
Two informative points, 15 attacking traces, and a SNR of 1

Imperfect profiling

d=2, Np=50, Na=30, SNR=1



d=5, Np=25, Na=15, SNR=1



Discussion & Conclusion

- TA is the method of choice in
  - well understood devices
  - (nearly) perfect profiling
  - low estimation and assumption errors
- ML is the method of choice in
  - high dimensionality contexts
  - low profiling set contexts