Bayes, not Naïve

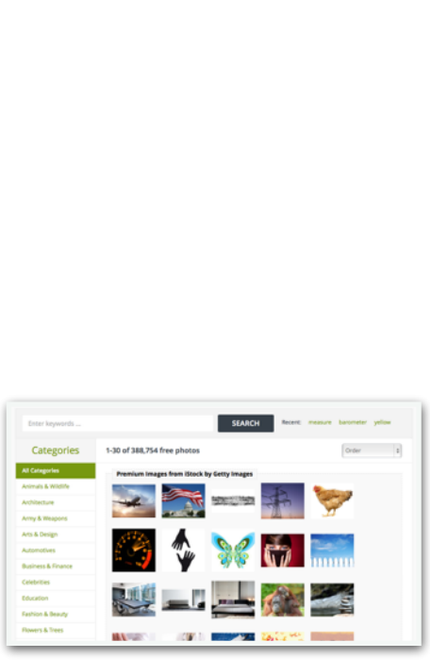
Security Bounds on Website Fingerprinting Defenses

Giovanni Cherubin

Privacy Enhancing Technologies Symposium
Minneapolis, Minnesota, USA
19 July, 2017

@gchers

Website Fingerprinting (WF)

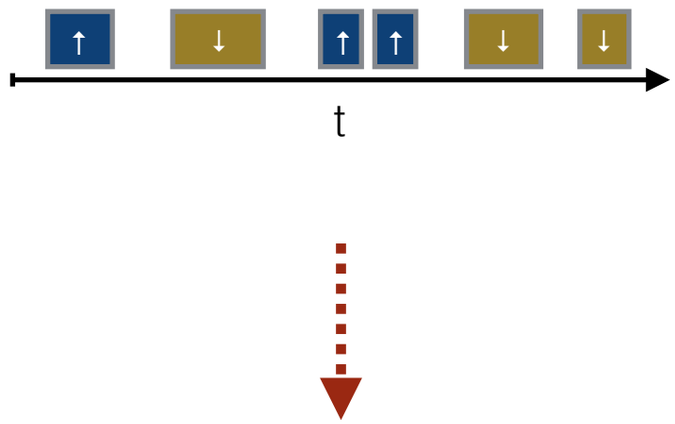


Encrypted Tunnel

Victim

Adversary

Website Fingerprinting (WF)



Adversary

Φ: transmission time, total bandwidth, …

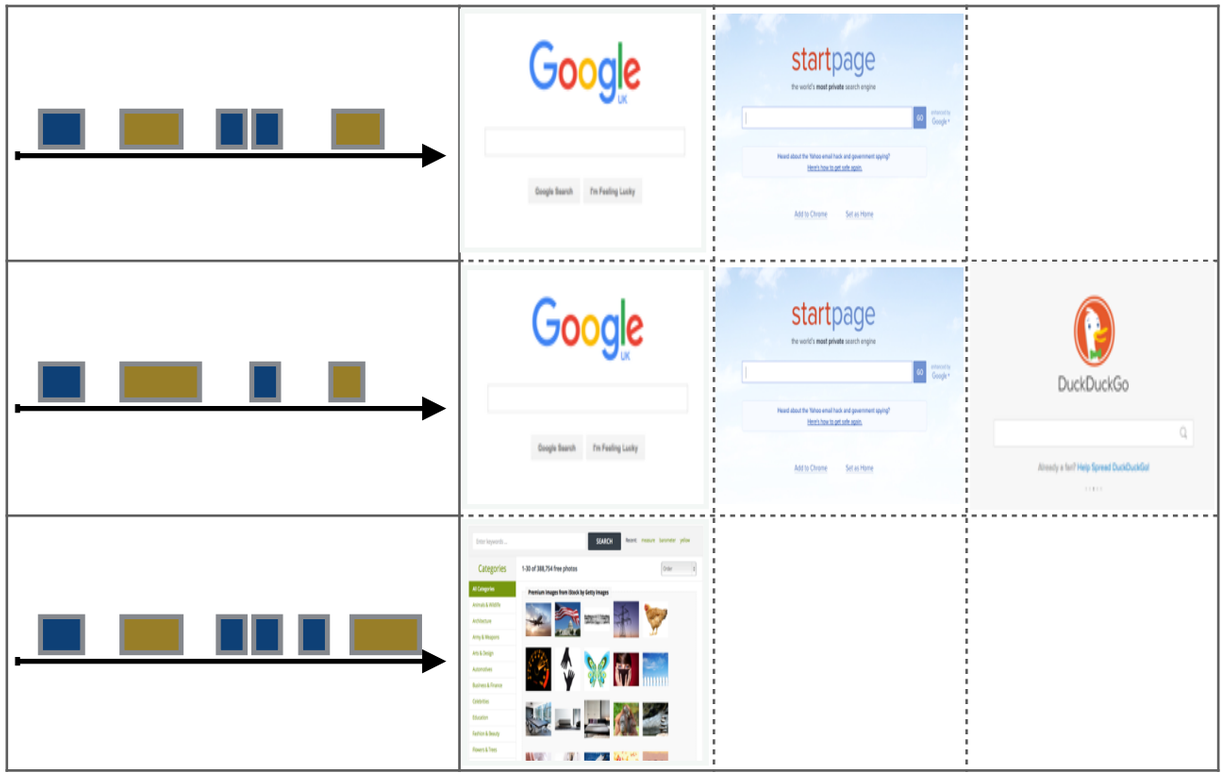
f_{train}: SVM, logistic regression, …

“Lookup-Table” Approach

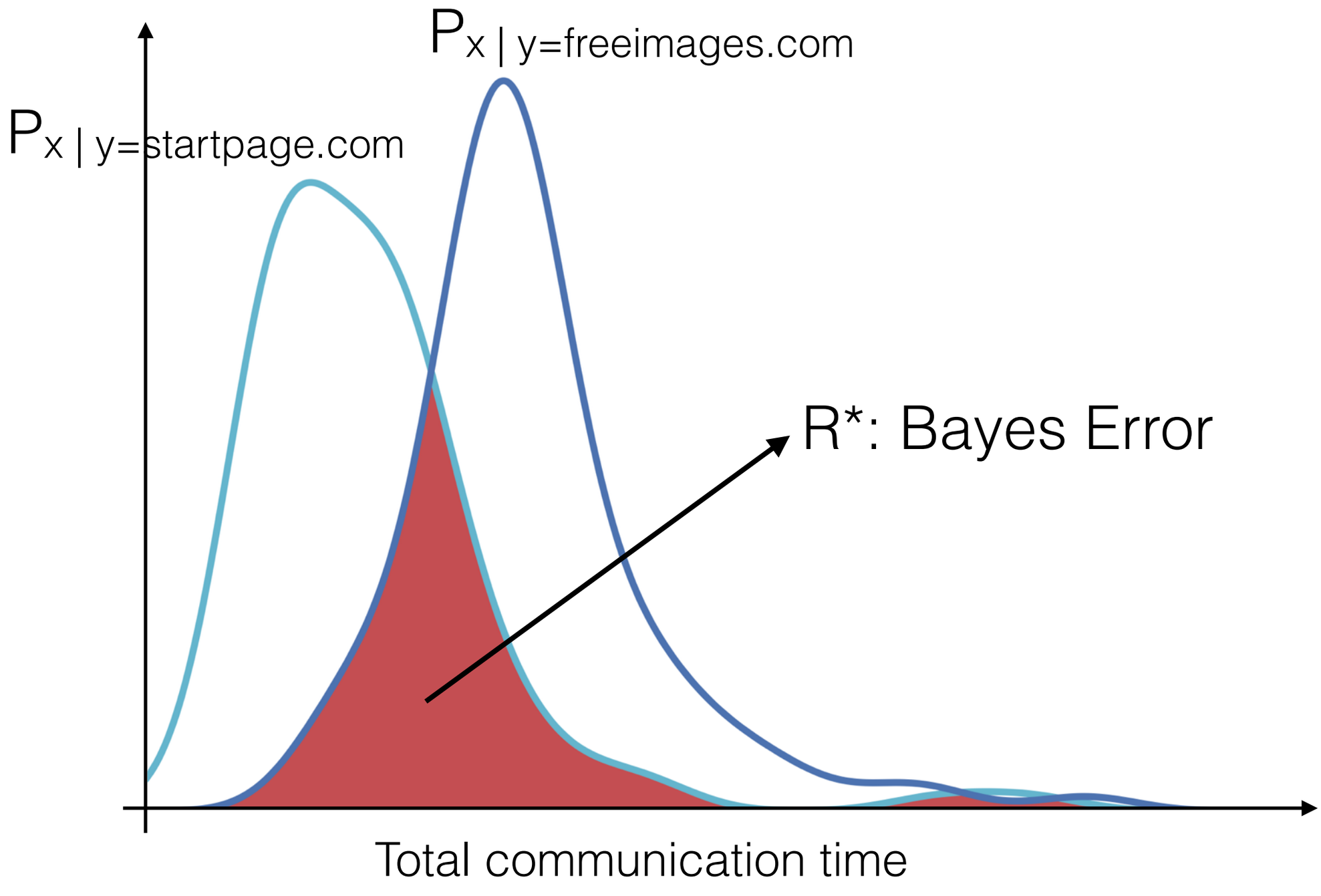
(Cai et al., ’14)

Idealised Adversary: knows exactly what packet sequences each web page may generate. Count the collisions.

Lookup table



Distinguishing Web Pages



“Bayes estimate” approach

(Φ, f_{train})

f ←

R^f : error on new packet sequence

$$\frac{L-1}{L}\left(1-\sqrt{1-\frac{L}{L-1}R^{NN}}\right) \leq R^* \leq R^f$$

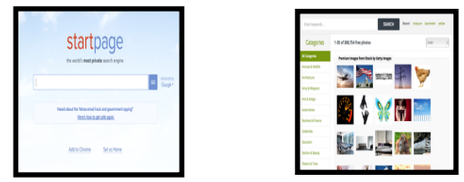
(Cover & Hart, ’67)

(ε,Φ)-privacy

Problem An error estimate $\hat{R}^*$ alone does not convey information about the setting. Random guessing R^G:

?

$R^G = 2/3$



?

$R^G = 1/2$

Define metric (1 - Adv):

$$\varepsilon = \hat{R}^* / R^G$$

(ε,Φ)-privacy

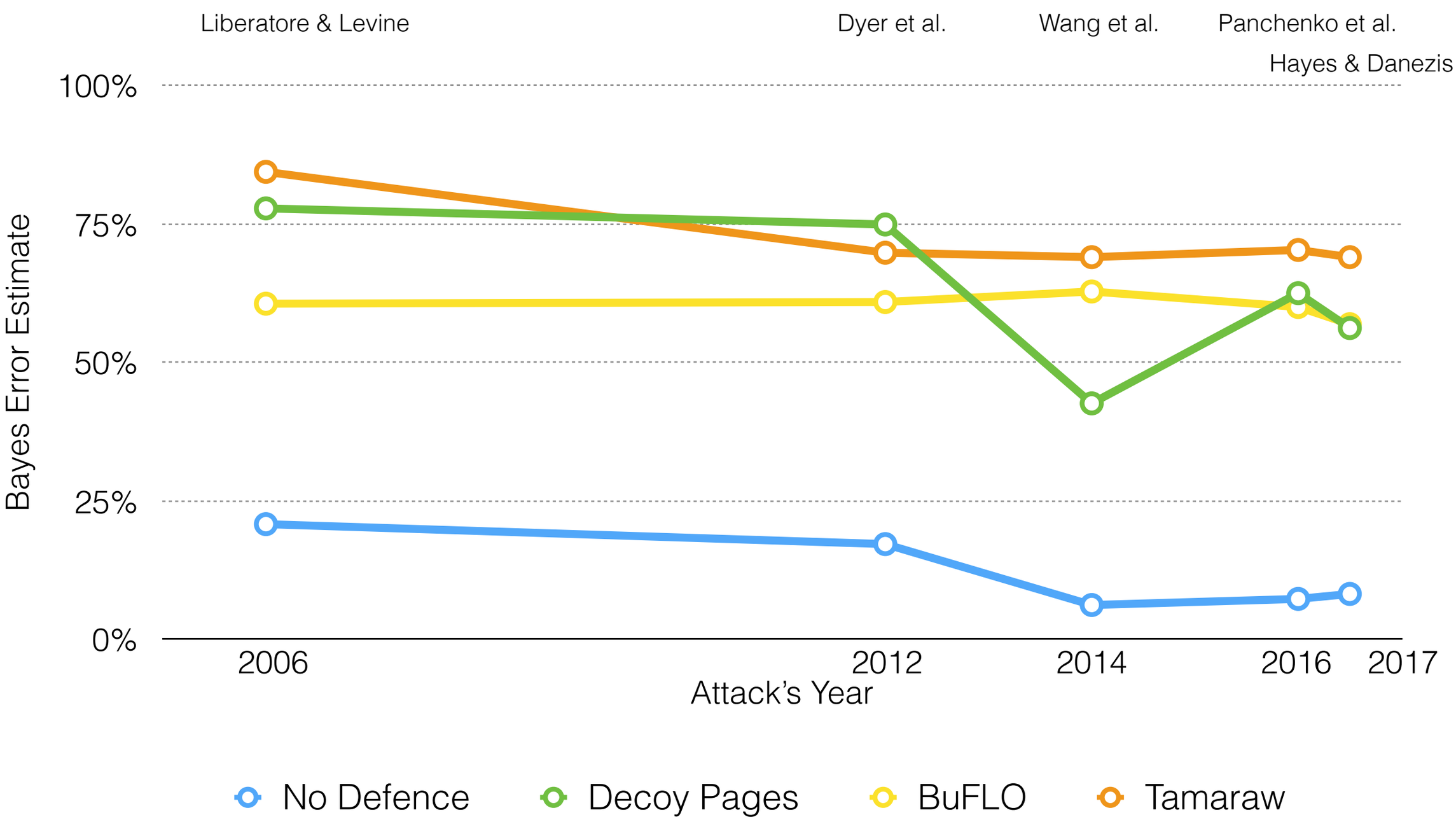
Closed World, WCN+ dataset (Tor traffic)

Defense*	(ε,Φ)-privacy	Packet OH	Time OH
No Defence	(0.06, k-NN)	0%	0%
Decoy Pages	(0.43, k-NN)	134%	59%
WTF-PAD	(0.49, k-FP)	247%	0%
BuFLO	(0.58, k-FP)	110%	79%
CS-BuFLO	(0.63, k-FP)	67%	576%
Tamaraw	(0.70, k-NN)	258%	341%

* Tor's default defense, Randomized Pipelining, is underlying each defense

(How much)

Did Feature Sets Improve?



Summary & Future Work

Blackbox method to derive security bounds for any WF defense and adversary (Φ, ·)

Future Work

- Prove some Φ is complete in some sense (“efficient”): from (ε,Φ)-privacy to ε-privacy
- Other estimates of R*, ensembles
- Other applications of technique: traffic analysis, side channel, generic ML-based attacks

Bayes, not Naïve

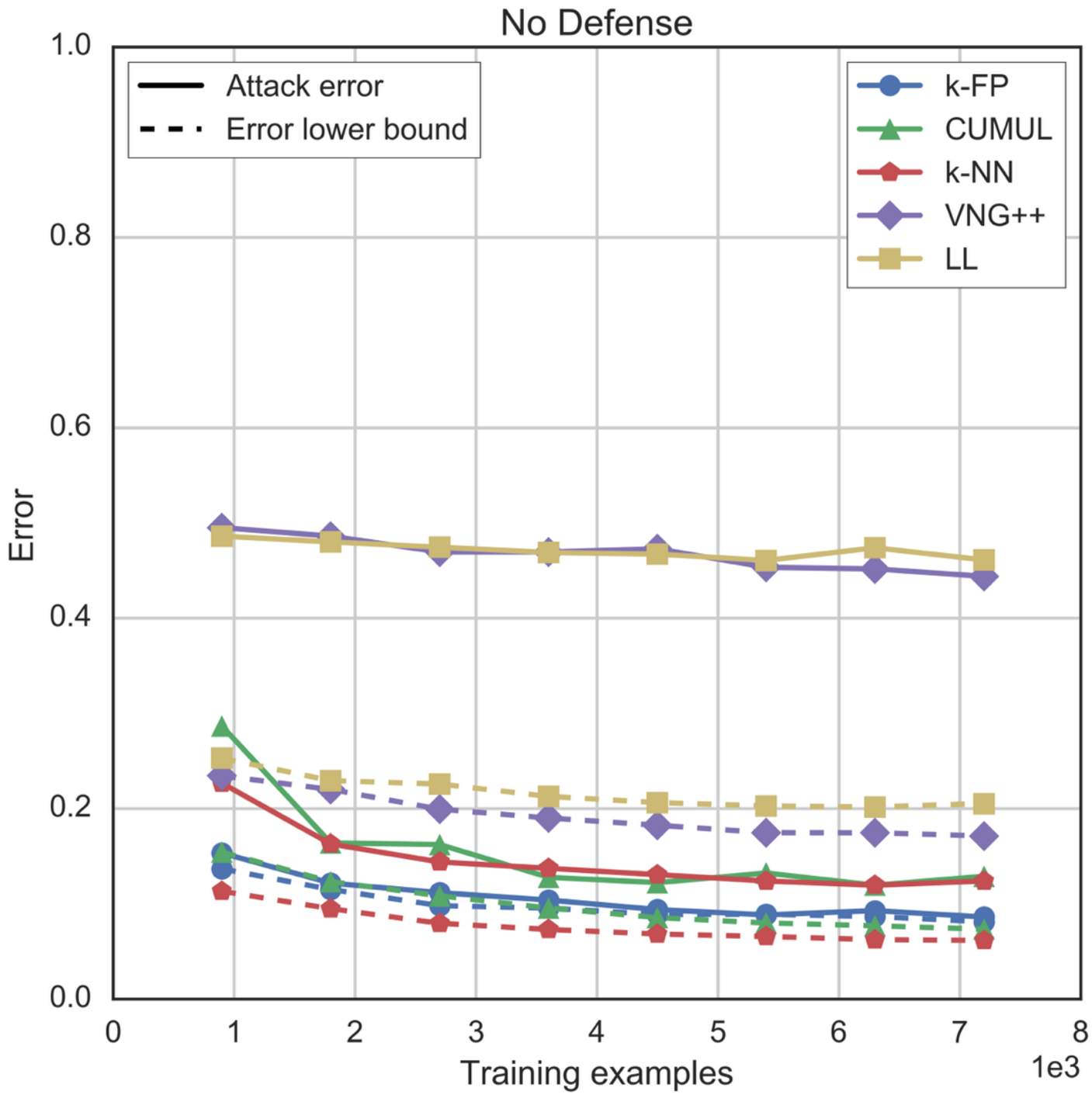
Security Bounds on Website Fingerprinting Defenses

Giovanni Cherubin

Privacy Enhancing Technologies Symposium
Minneapolis, Minnesota, USA
19 July, 2017

@gchers

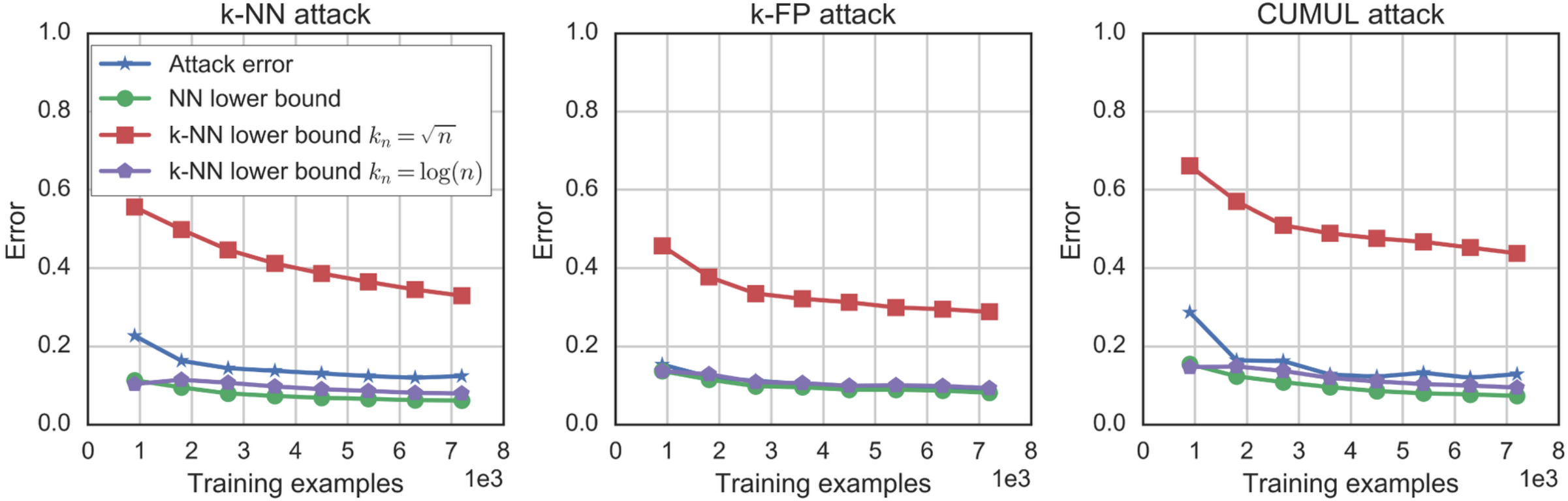
Lower bound convergence



k-NN Bayes Estimate

(Stone, ’77)

Theorem Let $k_n \to \infty$ and $k_n/n \to 0$ as $n \to \infty$, then $R^{k\text{-NN}} \to R^*$



Comparision with Cai et al.

Defence	R* estimate	Cai et al.	Cai et al. (full information)
BuFLO	57%	53%	19%
Tamaraw	69%	91%	11%

(ε,Φ)-privacy

One VS All scenario, WCN+ dataset

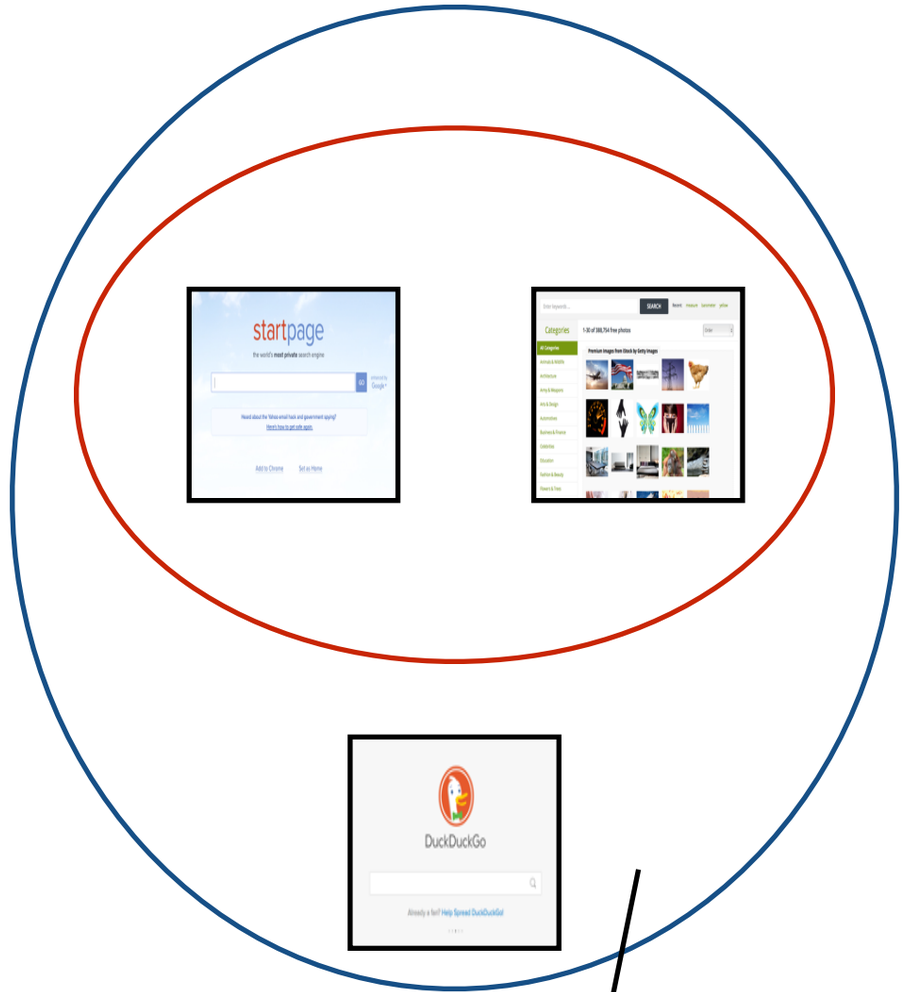
Defence	(ε,Φ)-privacy	Time OH	Packet OH
No Defence	(0.05, k-NN)	0%	0%
Decoy Pages	(0.29, k-NN)	134%	59%
BuFLO	(0.29, k-FP)	110%	79%
Tamaraw	(0.25, k-NN)	258%	341%
CS-BuFLO	(0.16, k-FP)	67%	576%
WTF-PAD	(0.18, CUMUL)	247%	0%

Q: What about priors?

- If true prior probabilities on web pages known, they can be used (i.e., bias the dataset accordingly).

- Ratio of success of one-try adversaries over random guessing maximized by uniform priors (Braun et al., 2009).

Q: Open World?



Adversary knows

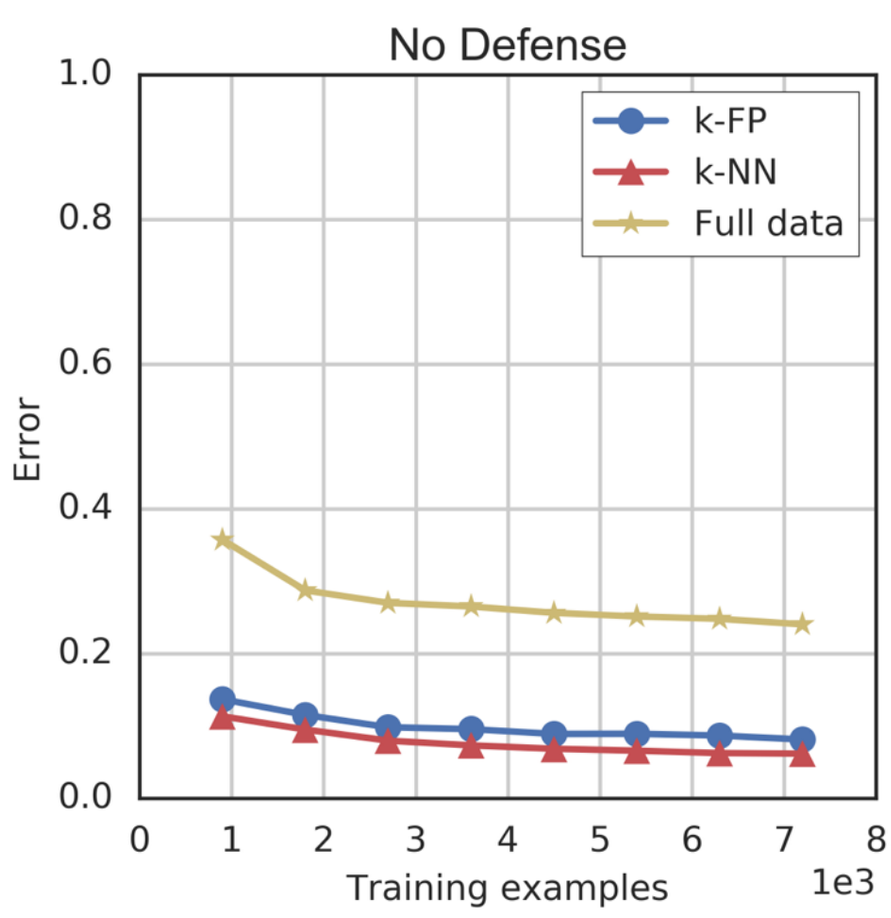
Victim may visit

y = “open”

Q: Bounds on full info?

Theorem For any transformation Φ: P → X,
R*(P) ≤ R*(Φ)

However,



Q: Is the code available?

Yes

https://github.com/gchers/wfes

Bayes, not Naïve

Security Bounds on Website Fingerprinting Defenses

Giovanni Cherubin

Privacy Enhancing Technologies Symposium
Minneapolis, Minnesota, USA
19 July, 2017

@gchers