CENTER FOR CYBERSECURITY

http://www.rit.edu/cybersecurity

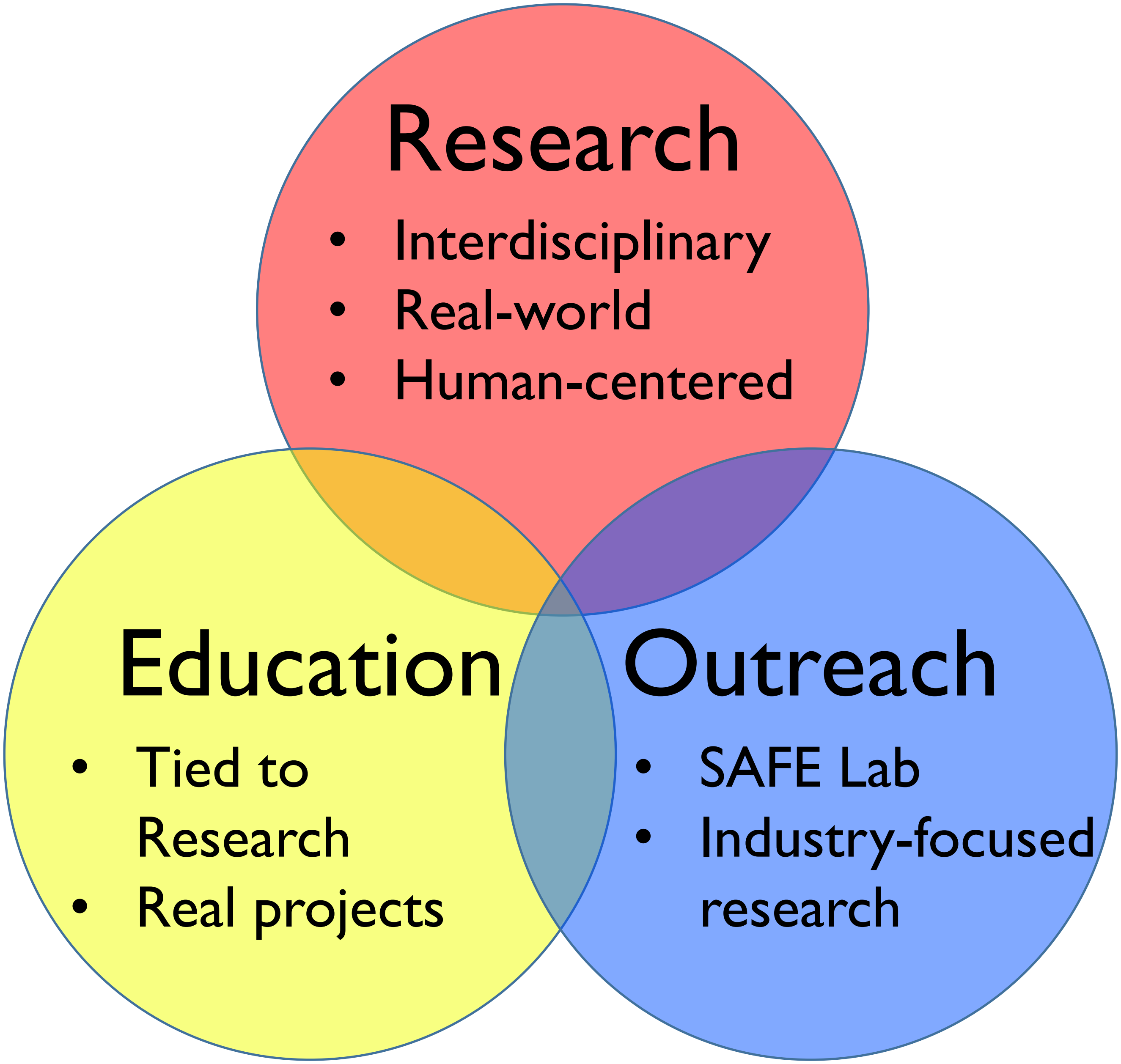
Matthew Wright, PhD

Director of the Center for Cybersecurity

Professor of Computing Security

Rochester Institute of Technology

Center Mission



Research

- Interdisciplinary
- Real-world
- Human-centered

Education

- Tied to Research
- Real projects

Outreach

- SAFE Lab
- Industry-focused research

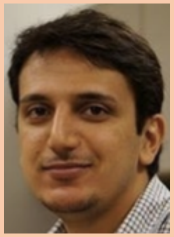
Security Analytics

- Prediction of attacks
  - Modeling attacker behavior
  - Simulation to predict outcomes

Katie McConky

S. Jay Yang

- Discovering Architectural Weaknesses
  - Finding & characterizing design flaws
  - Working w/ MITRE's CWE



Mehdi Mirakhorli

- Mining for Software Vulnerabilities
  - Understanding how software vulnerabilities happen
  - Metrics

Andy Meneely

Crypto & Trusted Hardware

- ML on Encrypted Data
  - Applying homomorphic encryption
  - Fully secure in the cloud

Peizhao Hu

- Trusted Computing
  - Cache-based attacks in SGX
  - Defenses

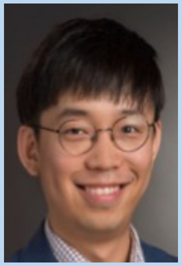
Ziming Zhao

- Crypto Hardware
  - FPGA implementations
  - Power analysis attacks

Marcin Lukowiak

Network Security

- Measuring Internet Security
  - DNSSEC Deployment
  - Certificate Authorities

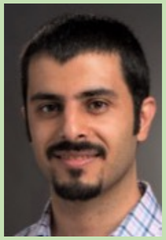


Tijay Chung

- Software-Defined Networks
  - SDN Firewalls
  - SDN Honeynets

Ziming Zhao

- Wireless Security
  - Full-frame Encryption
  - Securing PHY-layer attributes



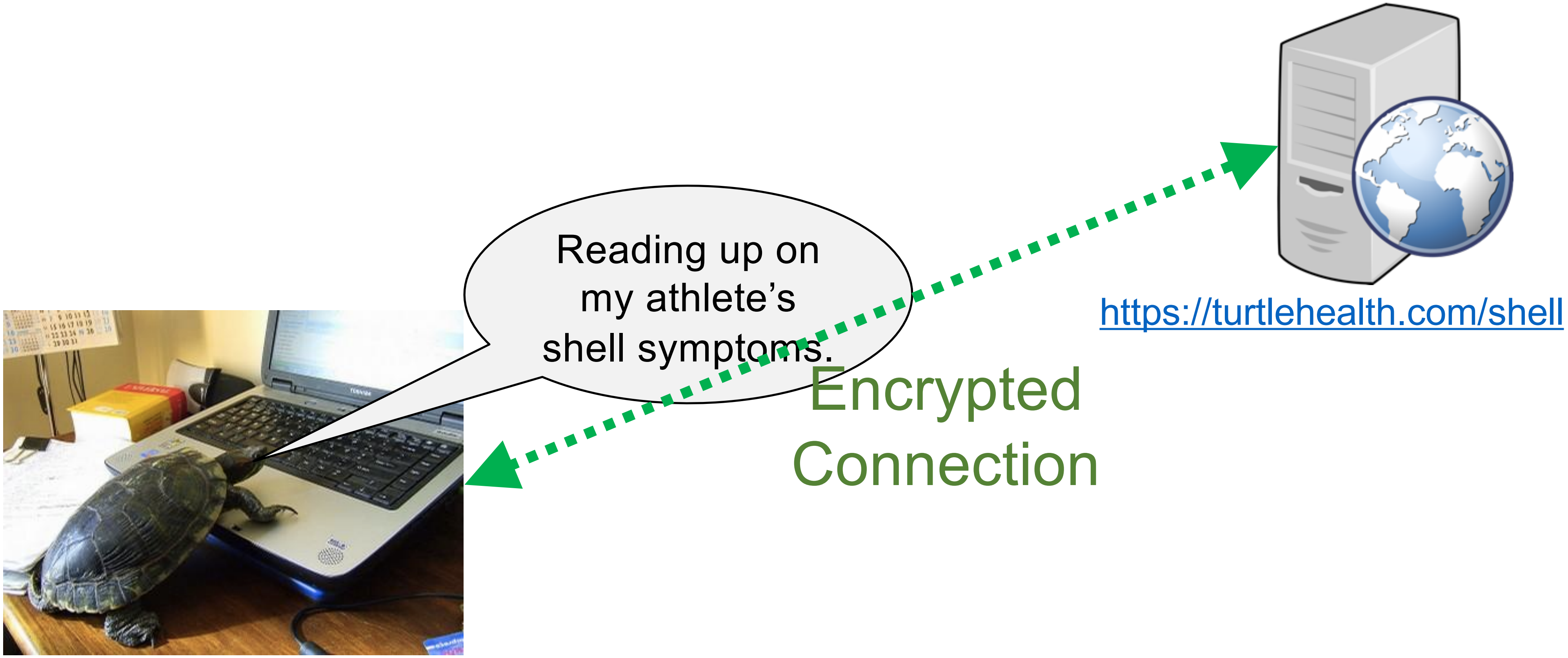
Hanif Rahbari

CENTER FOR CYBERSECURITY

http://www.rit.edu/cybersecurity

How Attackers Can Read Your Encrypted Traffic … and What to Do About It

Encrypted Traffic



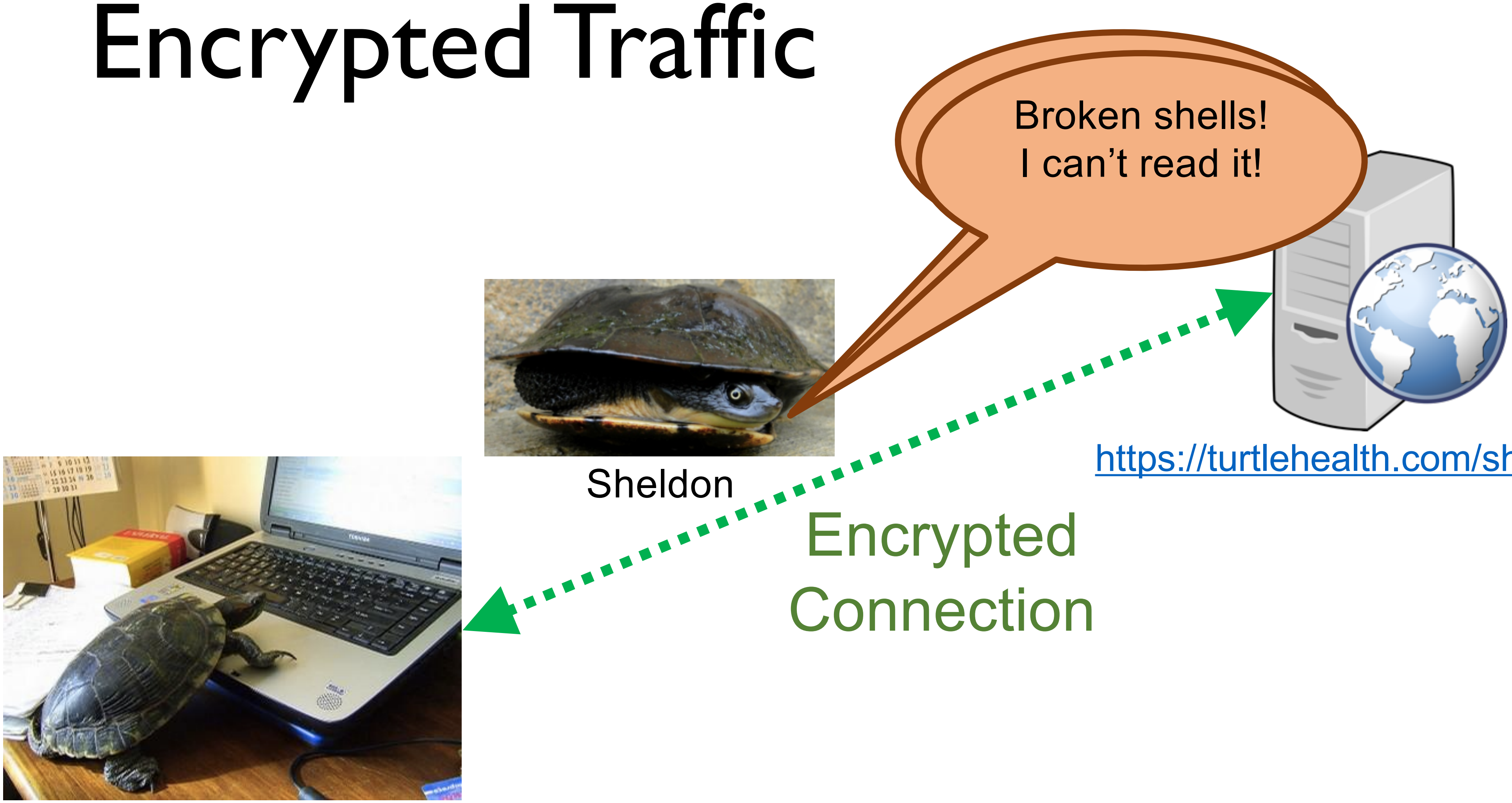
Reading up on my athlete's shell symptoms.

https://turtlehealth.com/shell

Encrypted Connection

Shelly

Encrypted Traffic



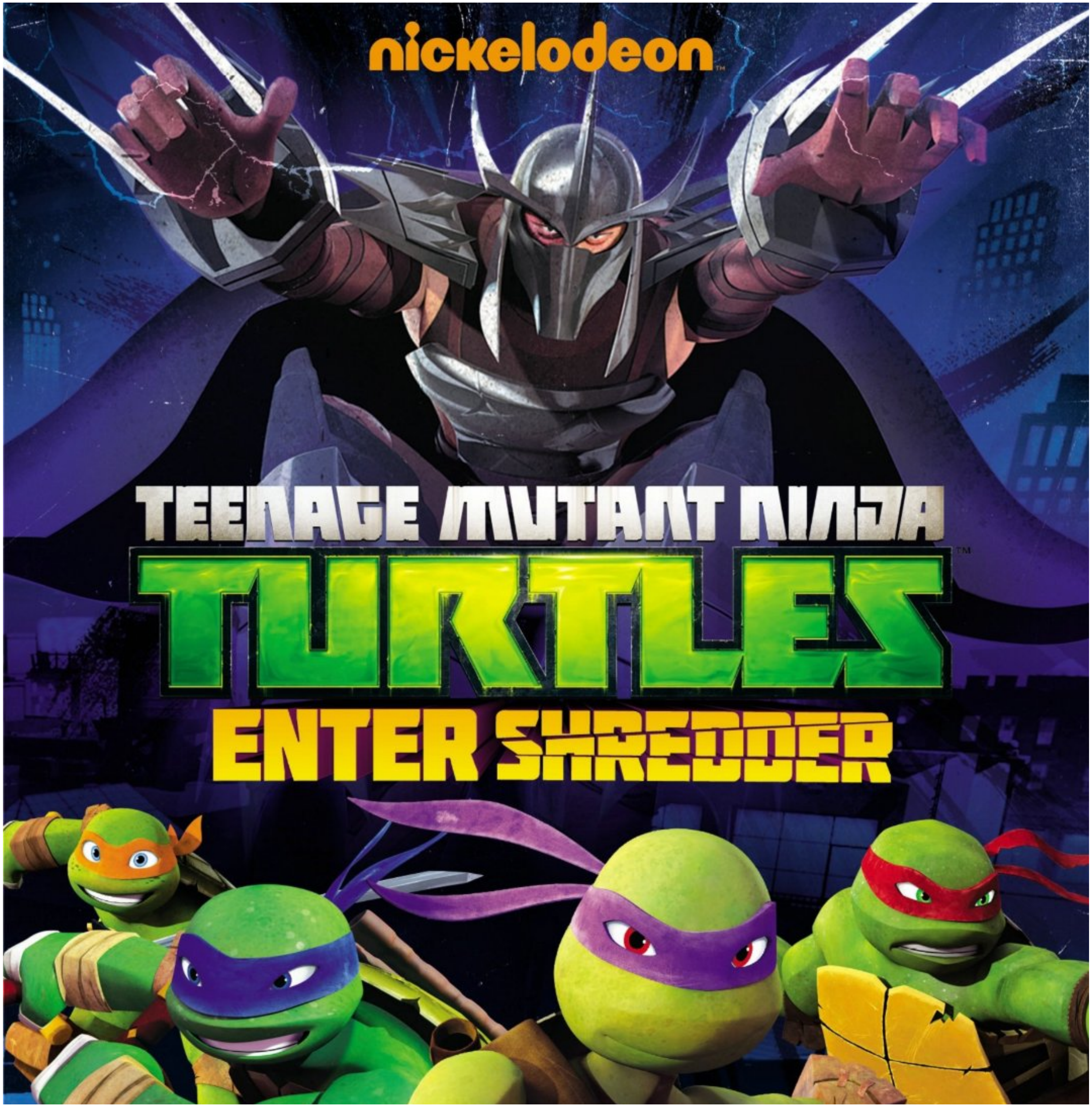
Broken shells!
I can’t read it!

Sheldon

https://turtlehealth.com/shell

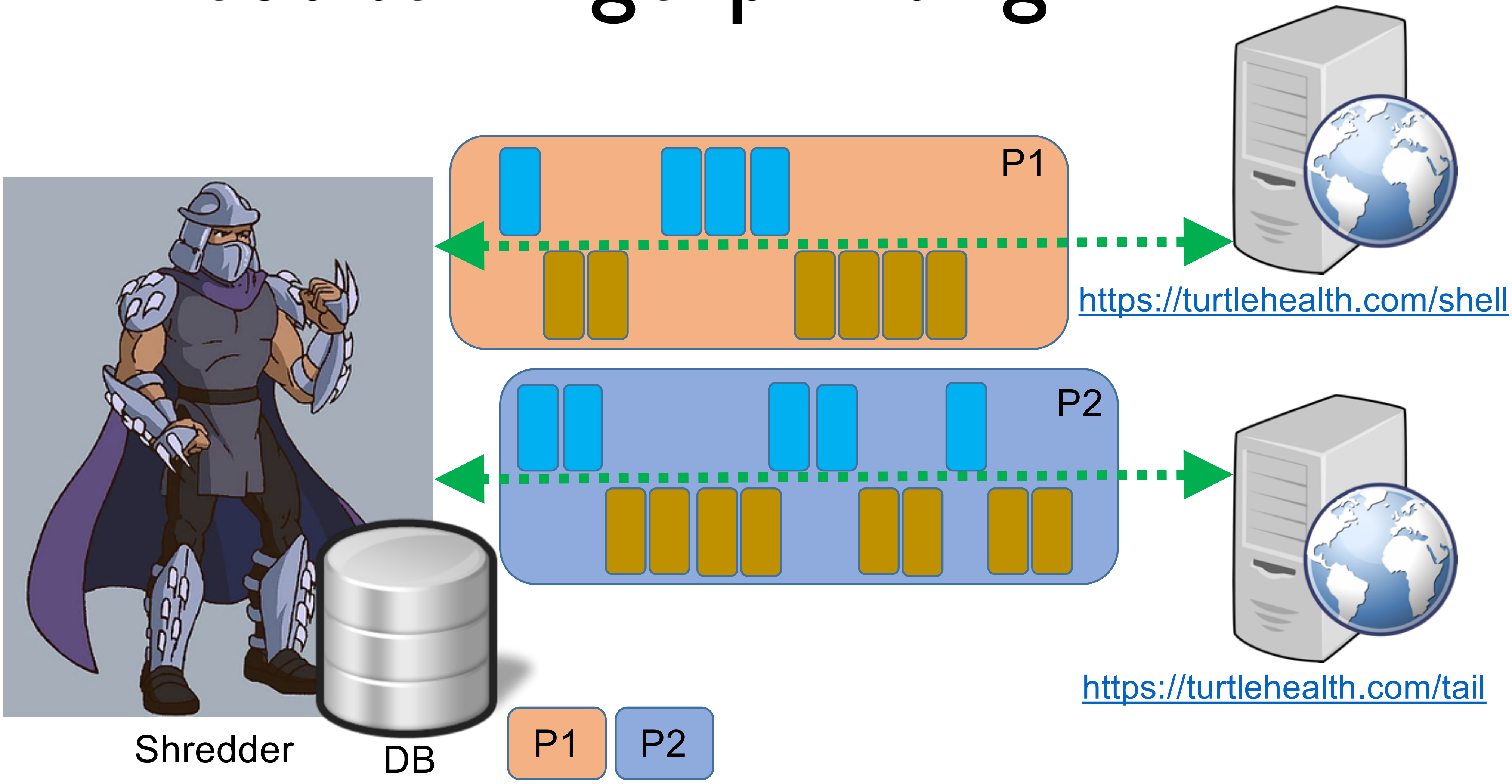
Encrypted
Connection

Shelly



http://www.nickandmore.com/wordpress/wp-content/uploads/2013/08/cover.jpg

Website Fingerprinting



P1

https://turtlehealth.com/shell

P2

https://turtlehealth.com/tail

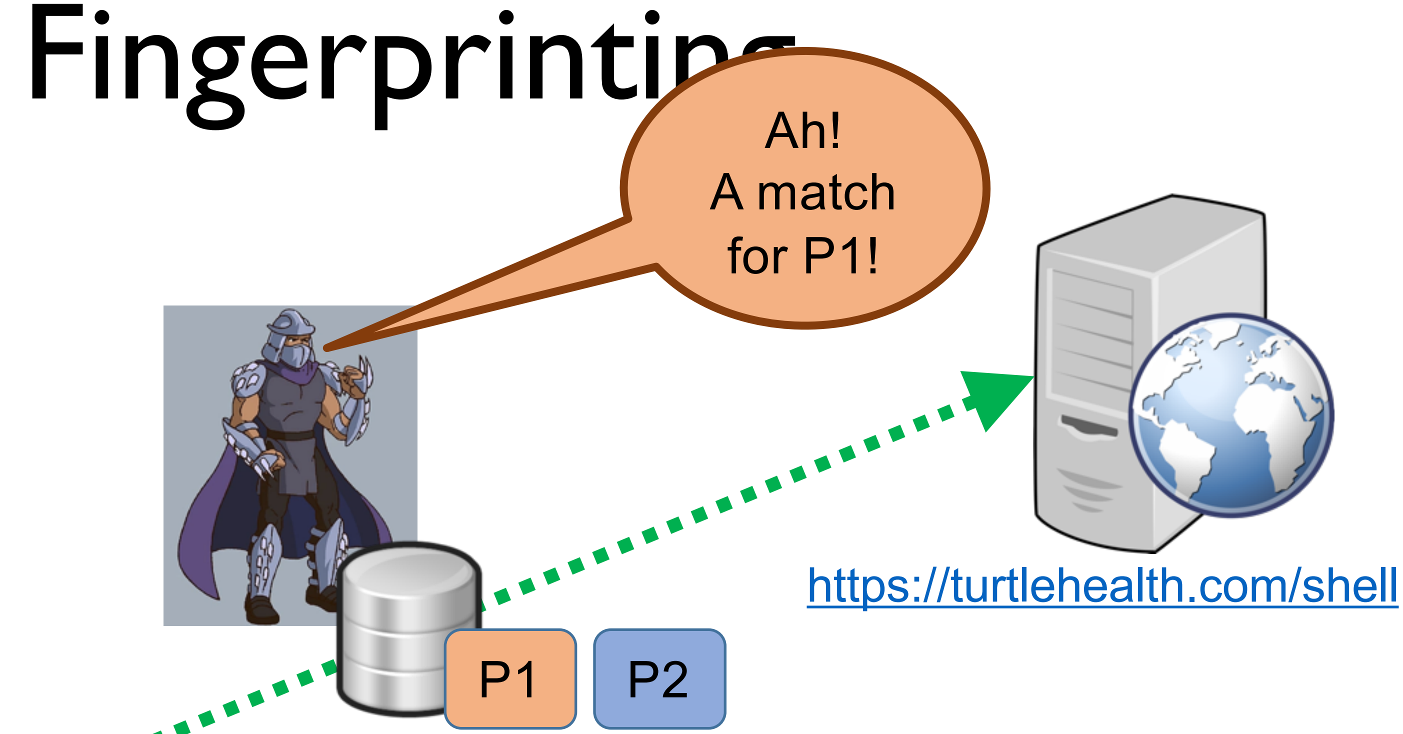
Shredder

DB

P1

P2

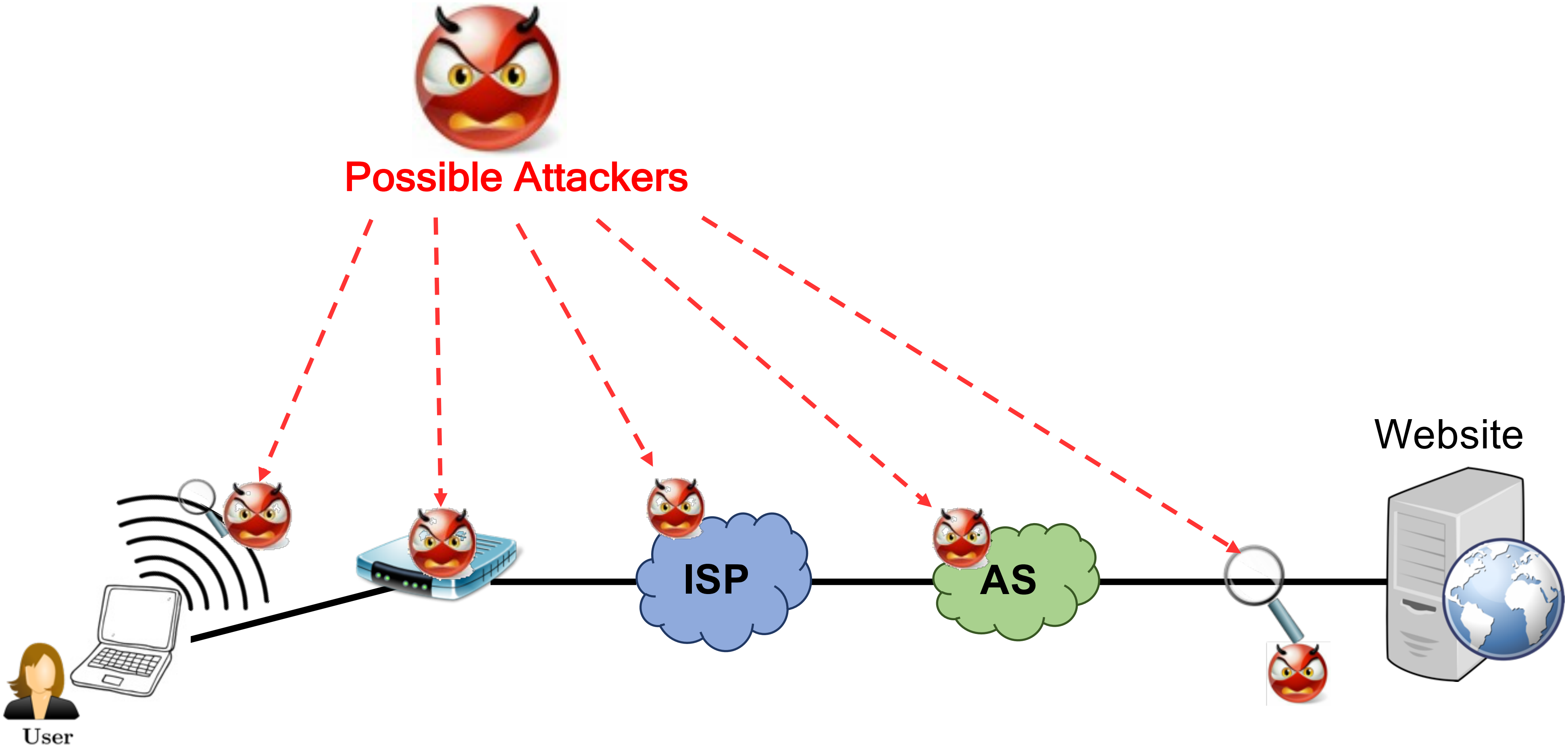
Website Fingerprinting



Shelly

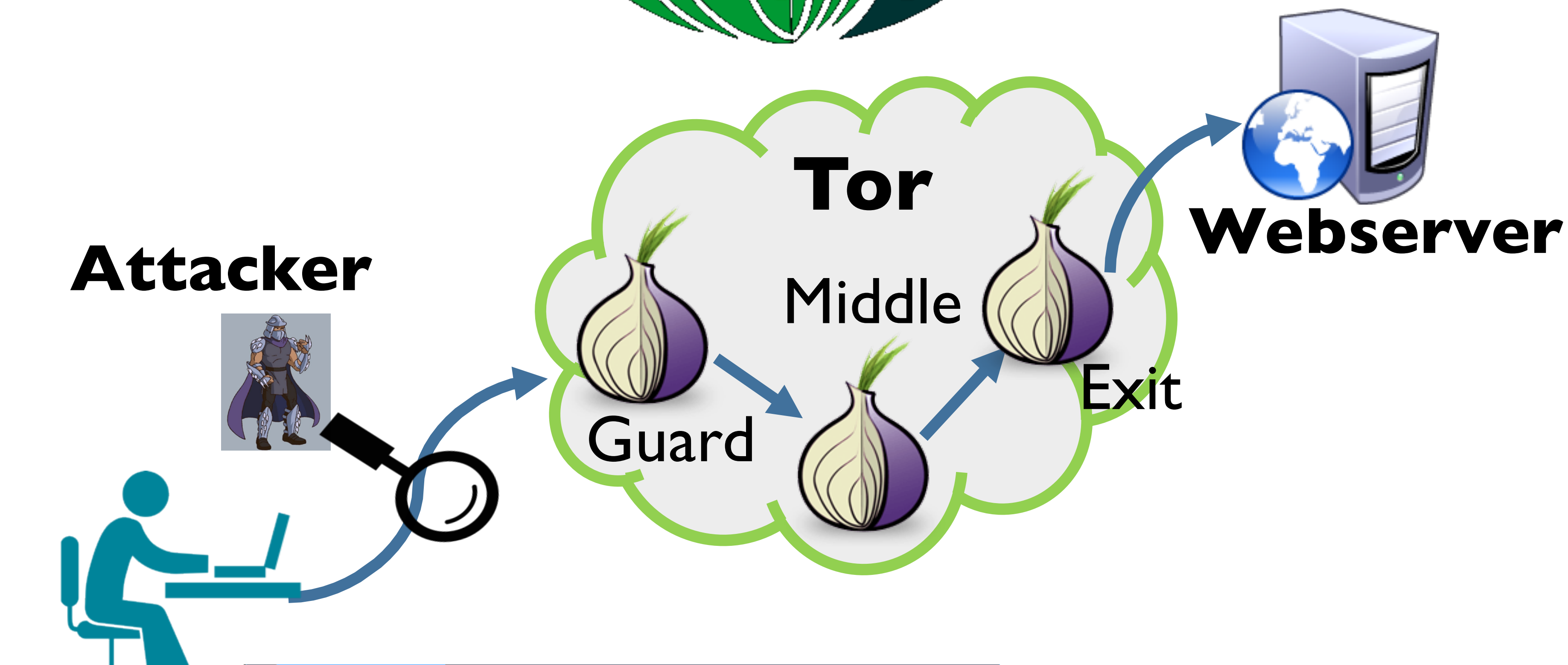
90%+ Accuracy

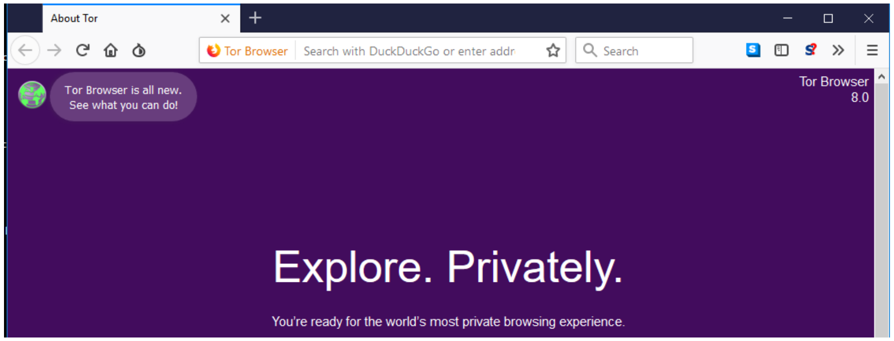
Website Fingerprinting Threat Model



Tor®

TorProject.org





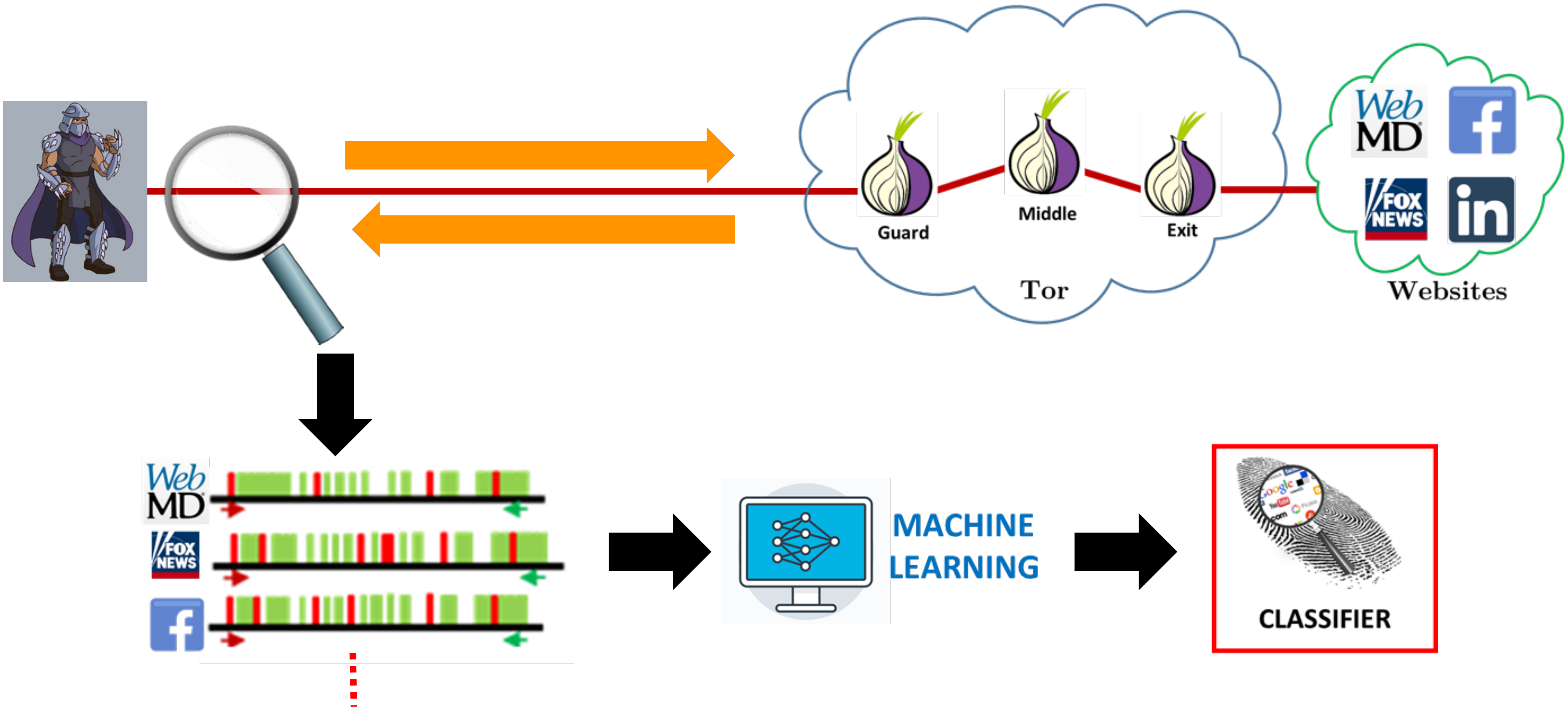
Tor Browser is all new. See what you can do!

Tor Browser 8.0

Explore. Privately.

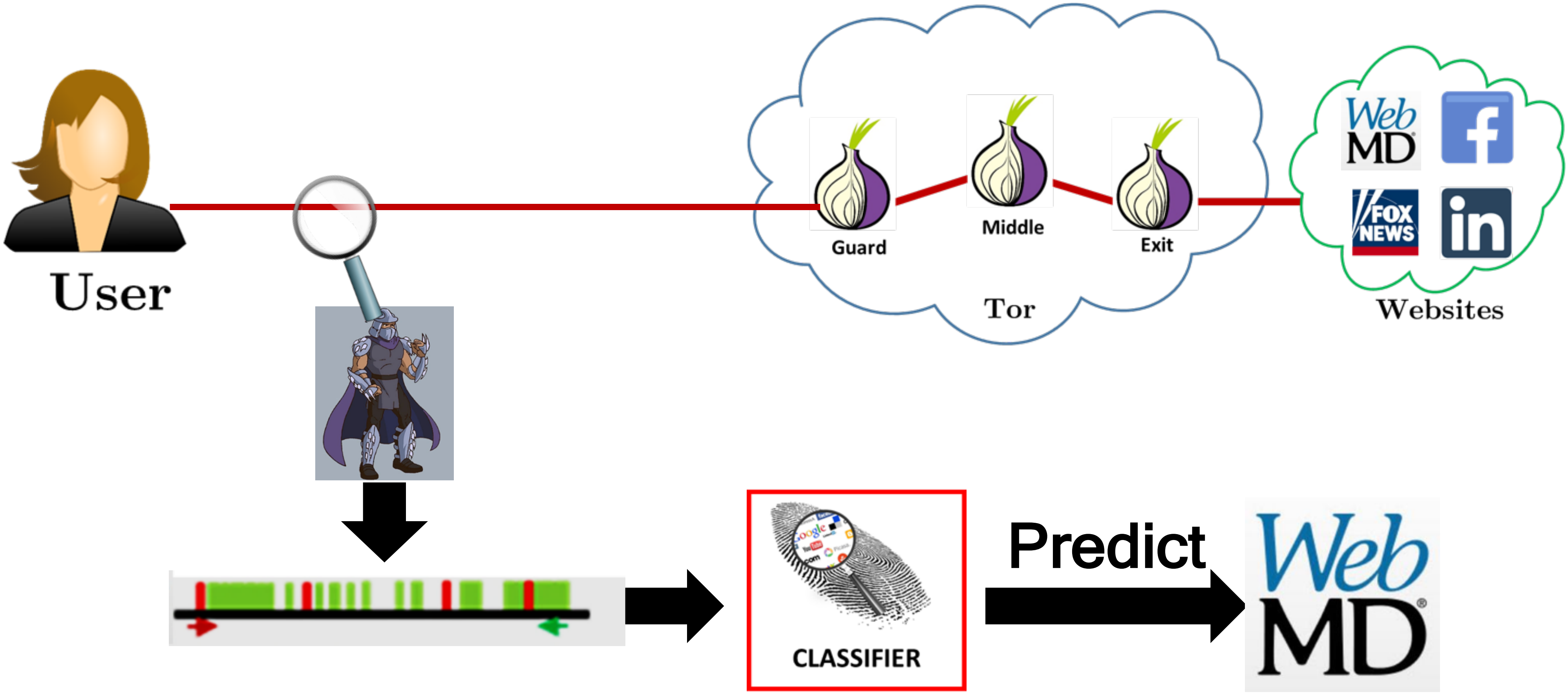
You're ready for the world's most private browsing experience.

Website Fingerprinting in Tor



Train the classifier

Website Fingerprinting in Tor



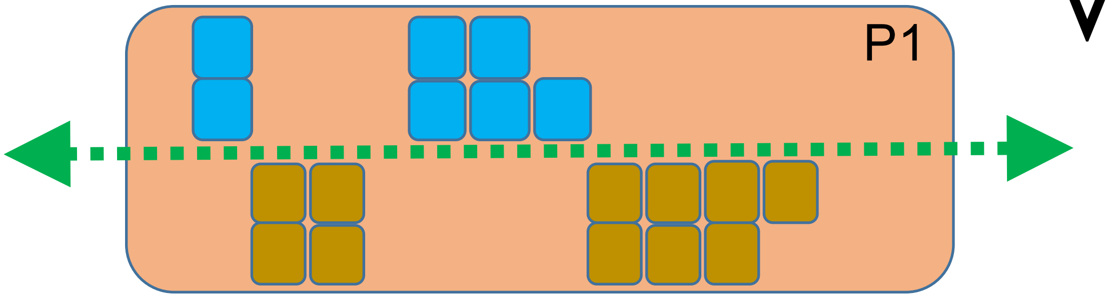
Perform the attack

Heh!
Nice try
☺

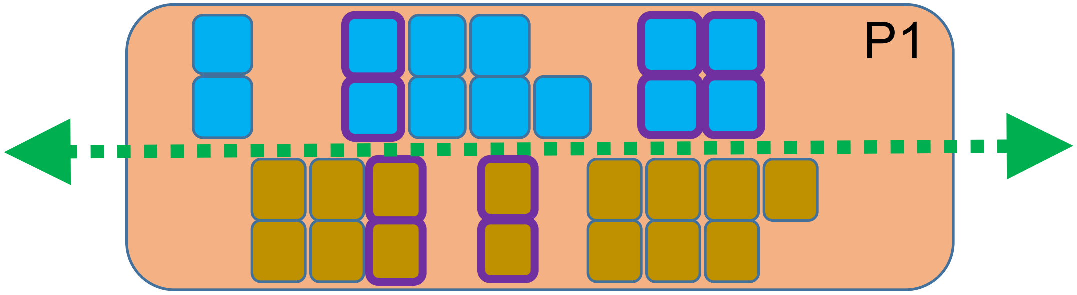
90%+ Accuracy*

* For ~100 sites, not pages

Adaptive Padding



Tor (unpadded)



Tor w/ Adaptive Padding

WTF-PAD

- AP for Tor
- 90% accuracy → 17%
- 54-64% bandwidth overhead
- Minimal added delay

Transition to Practice

- Working with Tor to deploy this

WTF!?!

CENTER FOR CYBERSECURITY

Questions?

R·I·T

Deep Fingerprinting

Undermining Website Fingerprinting Defenses with Deep Learning

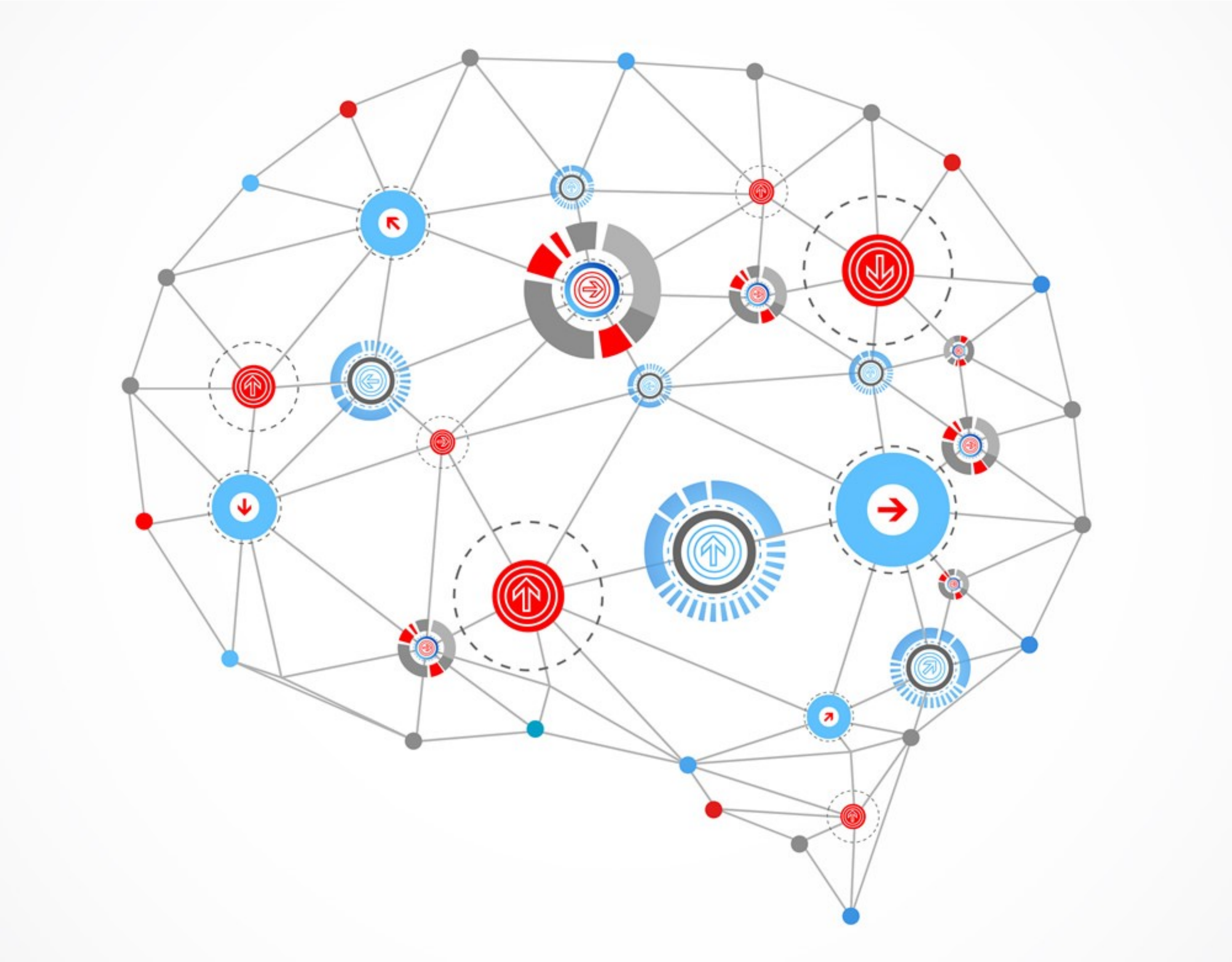
Payap Sirinam	Rochester Institute of Technology
Mohsen Imani	University of Texas at Arlington
Marc Juarez	imec-COSIC KU Leuven, Belgium
Matthew Wright	Rochester Institute of Technology

Payap

Mohsen

Marc

Deep Learning

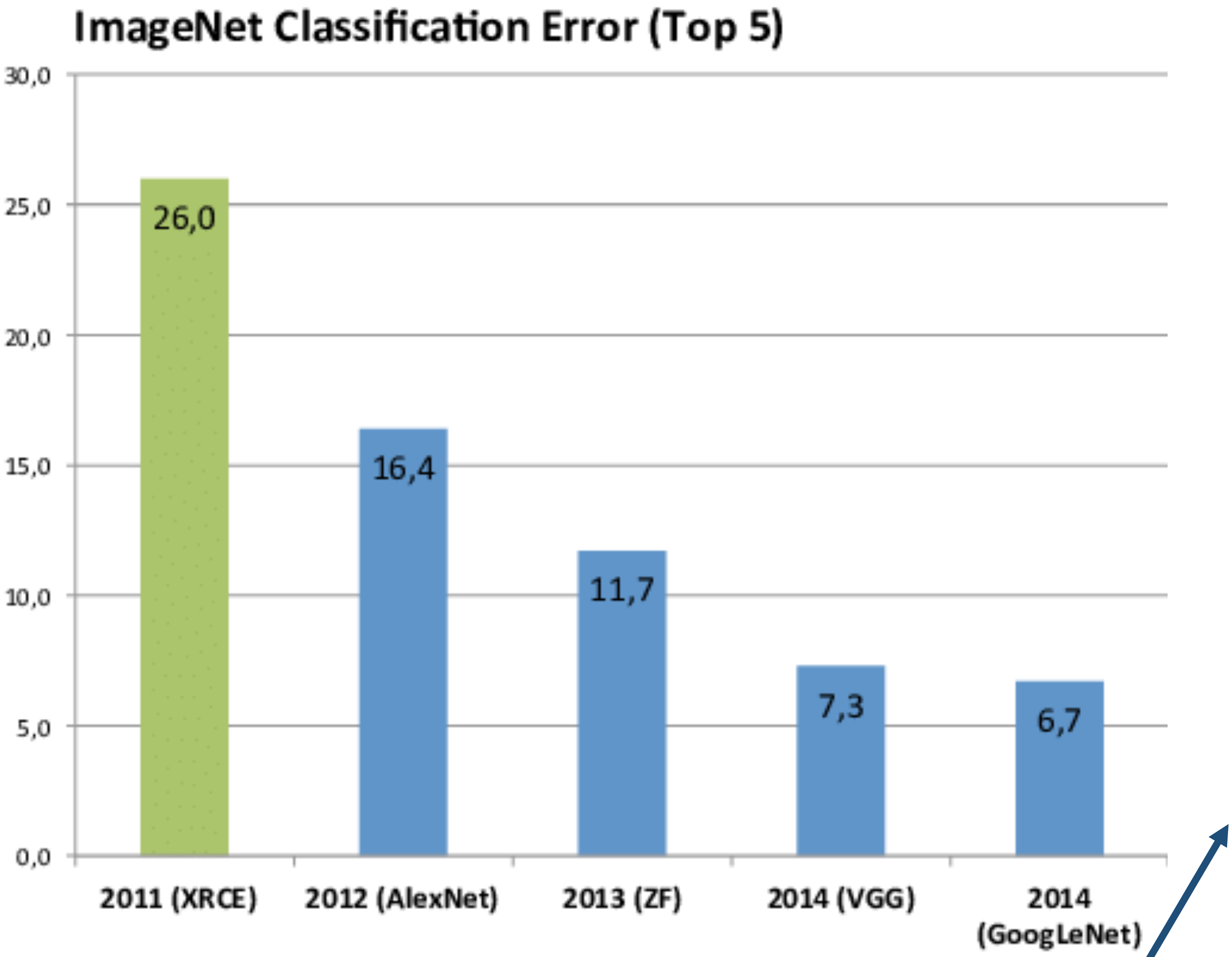


https://codeburst.io/deep-learning-what-why-dd77d432f182

ILSVRC: 1.2M images, 1.2K categories

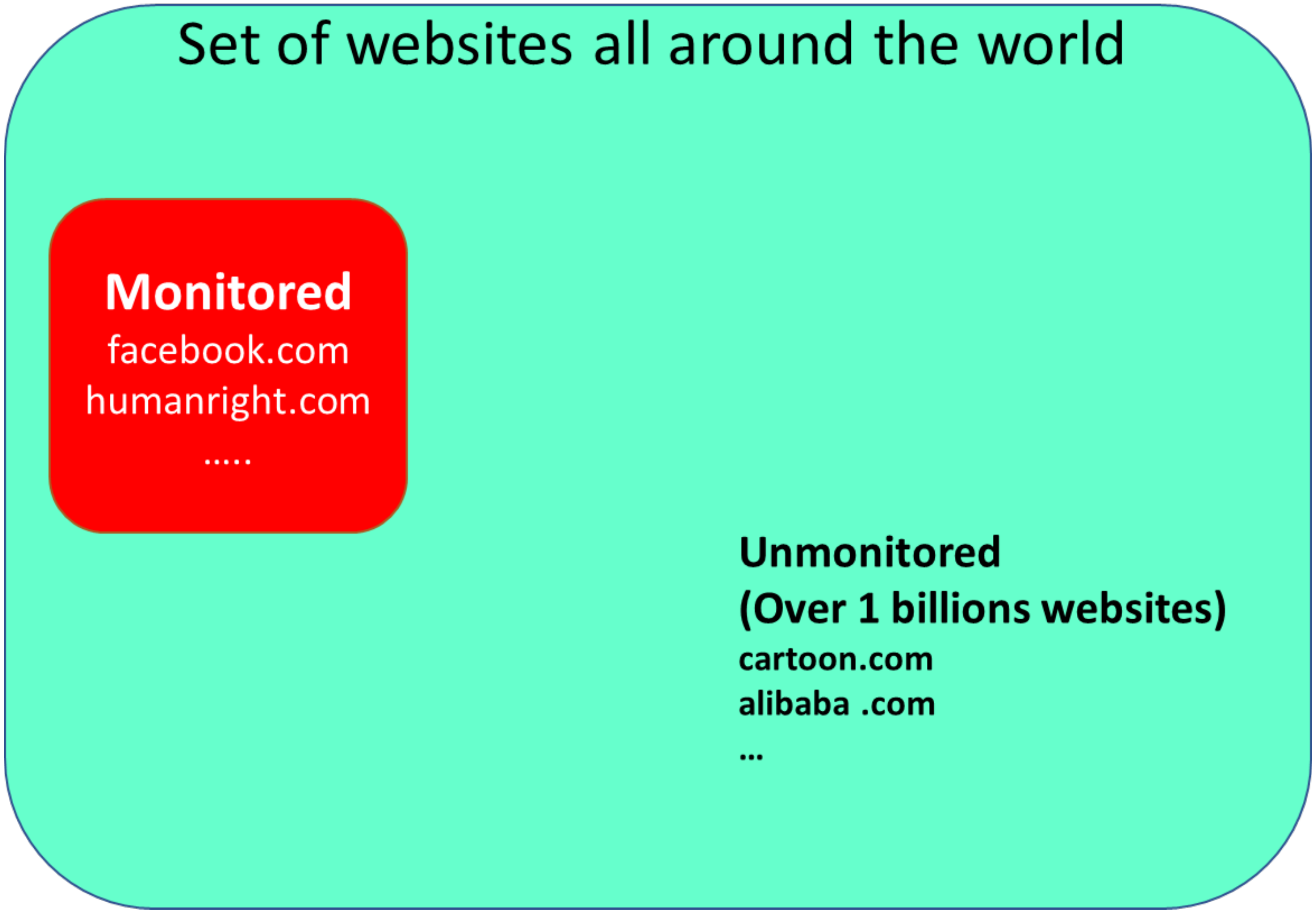
120 Breeds

http://arcticicekennels.tripod.com/puppies.html


ImageNet Classification Error (Top 5)
30,0
25,0
20,0
15,0
10,0
5,0
0,0
26,0
16,4
11,7
7,3
6,7
2011 (XRCE)
2012 (AlexNet)
2013 (ZF)
2014 (VGG)
2014 (GoogLeNet)

Trained!

Website Fingerprinting in Tor



Set of websites all around the world

Monitored
facebook.com
humanright.com
.....

Unmonitored
(Over 1 billions websites)
cartoon.com
alibaba .com
...

Monitored- vs Unmonitored Websites

Website Fingerprinting in Tor

Closed- vs Open World Scenarios

Monitored

facebook.com

humanright.com

…..

Closed-World Scenario

- Users only visit monitored websites

- Identify which website ?

- **Accuracy** of the attack

- Unrealistic [*JAA14*]

- Classifier performance evaluation

[JAA14] Juarez et al. A critical evaluation of website fingerprinting attacks., CCS 2014

Website Fingerprinting in Tor

Closed- vs Open World Scenarios

Open-World Scenario

- Users can visit any website in the world (> billions)

- Recognizing monitored or unmonitored

- More realistic and more difficult

- **Precision and Recall** [*JAA14* , *PLZ16*]

[JAA14] Juarez et al. A critical evaluation of website fingerprinting attacks., CCS 2014
[PLZ16] Panchenko et al. Website fingerprinting at internet scale., NDSS 2016

Website Fingerprinting Attacks & Defenses

Website Fingerprinting Attacks & Defenses

WF Attacks using Hand-crafted Features

- Feature engineering
- 3 state-of-the-art
  - *k*-NN [*WCN14*]
  - CUMUL [*PLZ16*]
  - *k*-FP [*HD16*]
- 90+% Accuracy

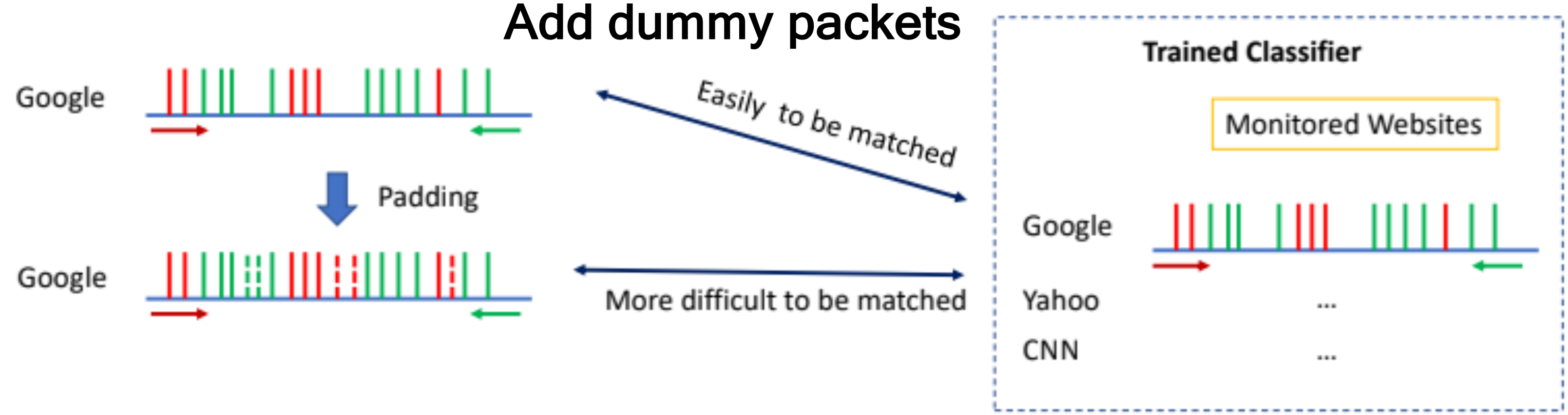
[WCN14] Wang et al. Effective attacks and provable defenses for website fingerprinting., USENIX 2014
[PLZ16] Panchenko et al. Website fingerprinting at internet scale., NDSS 2016
[HD16] Hayes and Danezis. k-Fingerprinting: A robust scalable website fingerprinting technique., USENIX 2016.

Website Fingerprinting Attacks & Defenses

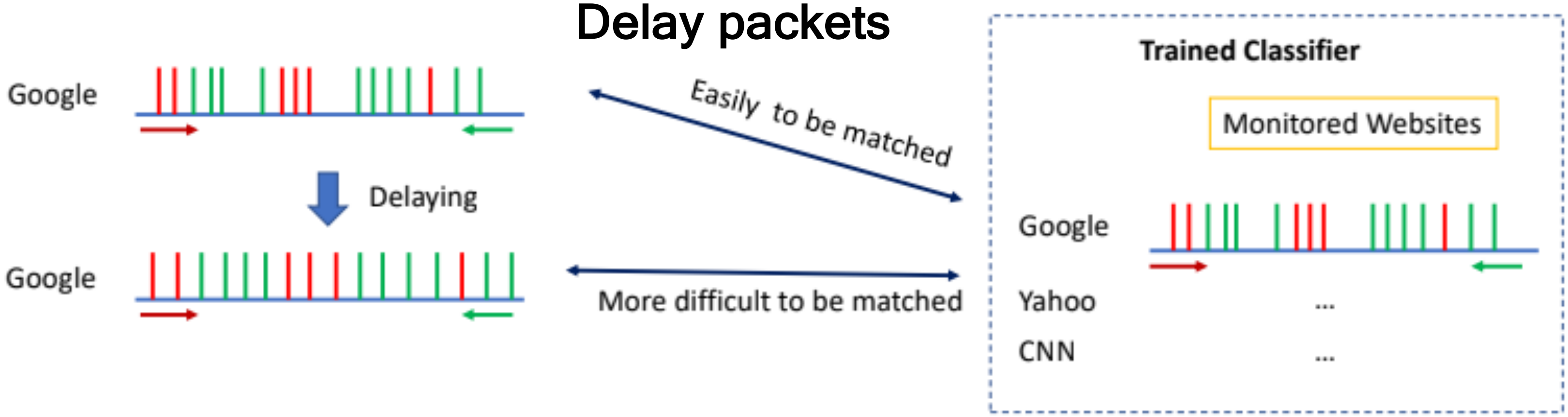
WF Defenses

- Basic mechanisms

Add dummy packets



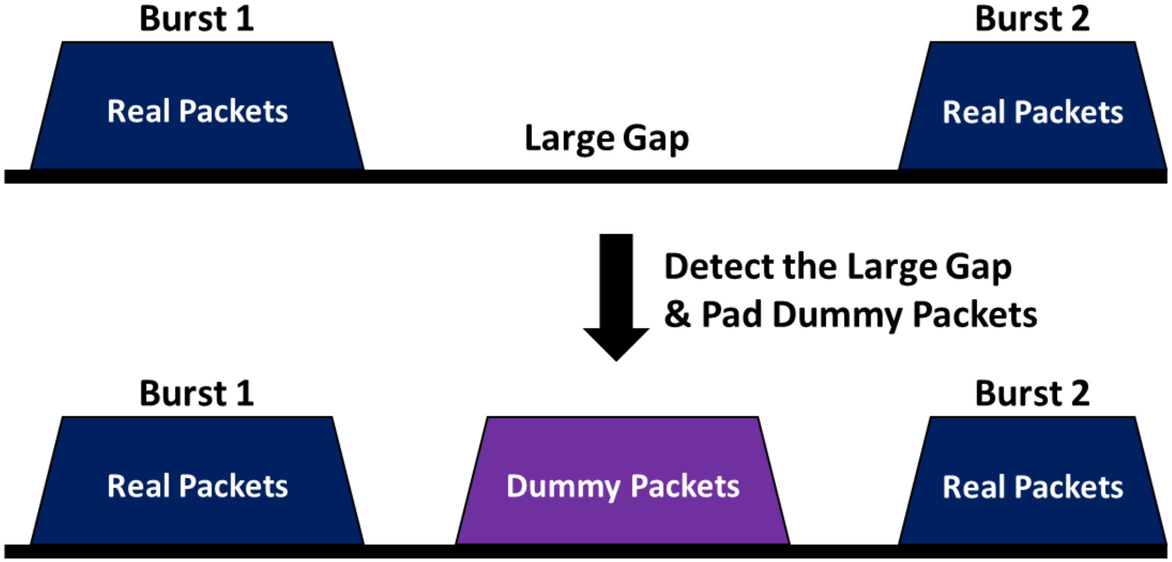
Delay packets



Website Fingerprinting Attacks & Defenses

Lightweight WF Defenses

- WTF-PAD [*JIP16*]



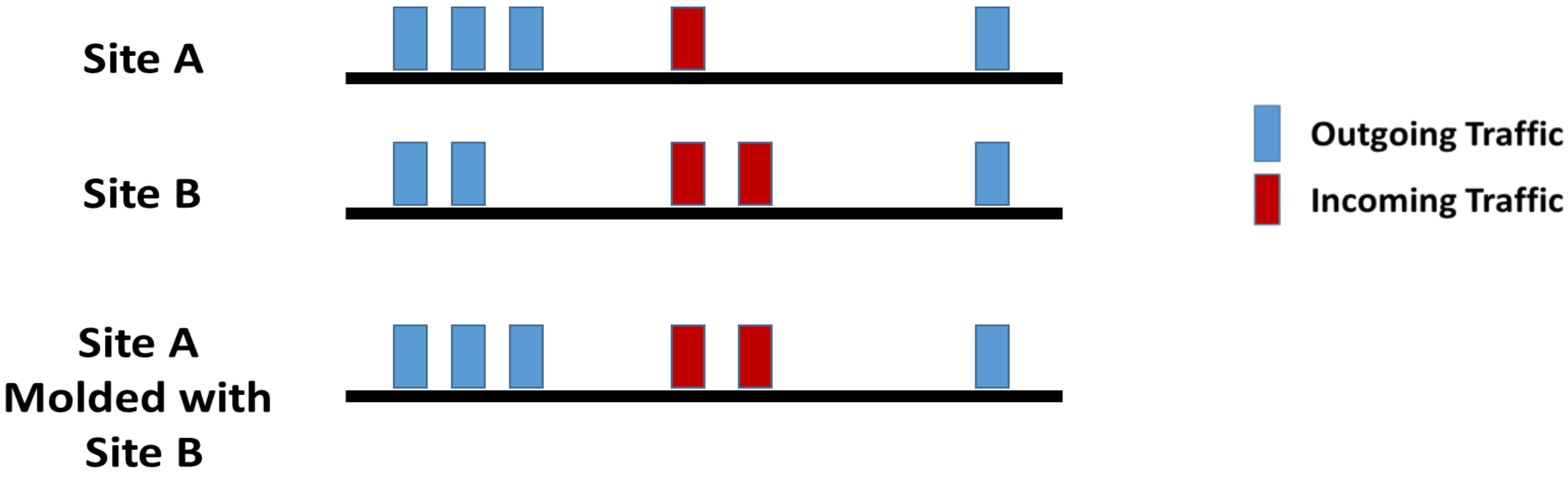
- Moderate bandwidth e.g. 54% + Low delay
- Reduce accuracy < 20%
- Main candidate to be deployed in Tor. [*PER15*]

[JIP16] Juarez et al. Toward an efficient website fingerprinting defense., ESORIC2016.
[PER15] Mike Perry. Padding negotiation. Tor protocol specification., 2015.

Website Fingerprinting Attacks & Defenses

Lightweight WF Defenses

- Walkie-Talkie (W-T) [*WG17*]



  - 31% extra bandwidth overhead & 34% extra latency overhead
  - Reduce accuracy < 30%

[WG17] Wang and Goldberg. Walkie-talkie: An efficient defense against passive website fingerprinting attacks. USENIX 2017

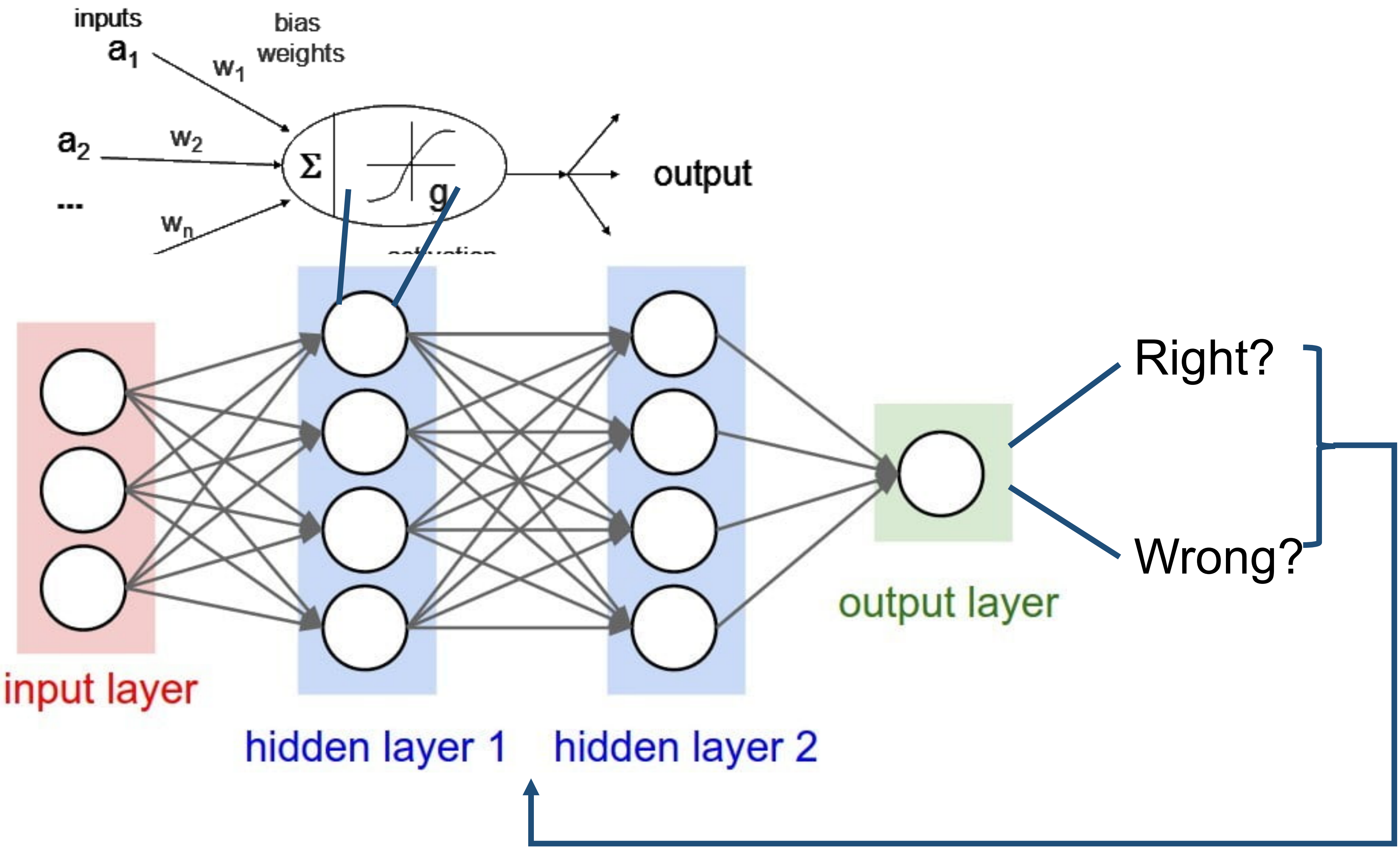
Website Fingerprinting Attacks & Defenses

WF Attacks using Deep Learning

- Rimmer et al. work [*RPJ18*]
  - Automated feature engineering
  - 3 DL vs 1 Hand-crafted
    - SDAE, CNN, LSTM vs CUMUL
  - CNN, SDAE and CUMUL consistently perform best
    - 95-97% Accuracy

[RPJ18] Rimmer et al. Automated website fingerprinting through deep learning., NDSS2018

Neural Networks (in 1 slide)

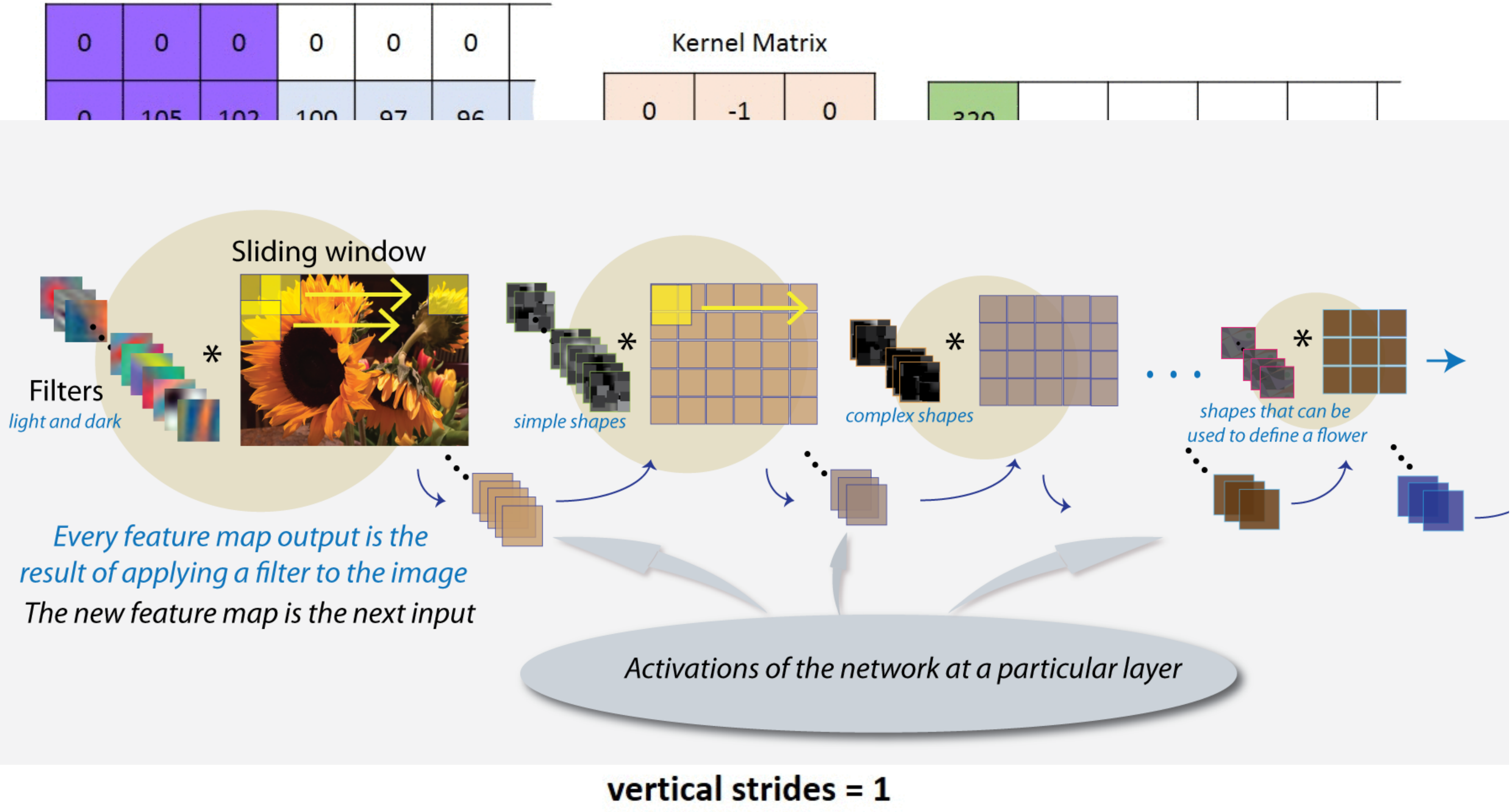


Right?

Wrong?

https://stats.stackexchange.com/questions/188277/activation-function-for-first-layer-nodes-in-an-ann
https://www.digitaltrends.com/cool-tech/what-is-an-artificial-neural-network/

CNNs (in 1 slide)



Every feature map output is the result of applying a filter to the image

The new feature map is the next input

Activations of the network at a particular layer

https://stats.stackexchange.com/questions/188277/activation-function-for-first-layer-nodes-in-an-ann
https://www.digitaltrends.com/cool-tech/what-is-an-artificial-neural-network/

Website Fingerprinting Attacks & Defenses

Goals

- Prior work
  - CNN model → early-proposed architecture
- Improvement of CNN in the literature

~55% Accuracy

AlexNet (2012)

~71% Accuracy

VGG19 (2014)

~80% Accuracy

Inception V4 (2016)

Canziani et al. An Analysis of Deep Neural Network Models for Practical Applications., arXiv:1605.07678

Website Fingerprinting Attacks & Defenses

Key Challenges

- No evaluation against WF defenses

Original

CNN Model

Effective
e.g. ~80 Accuracy

Distorted

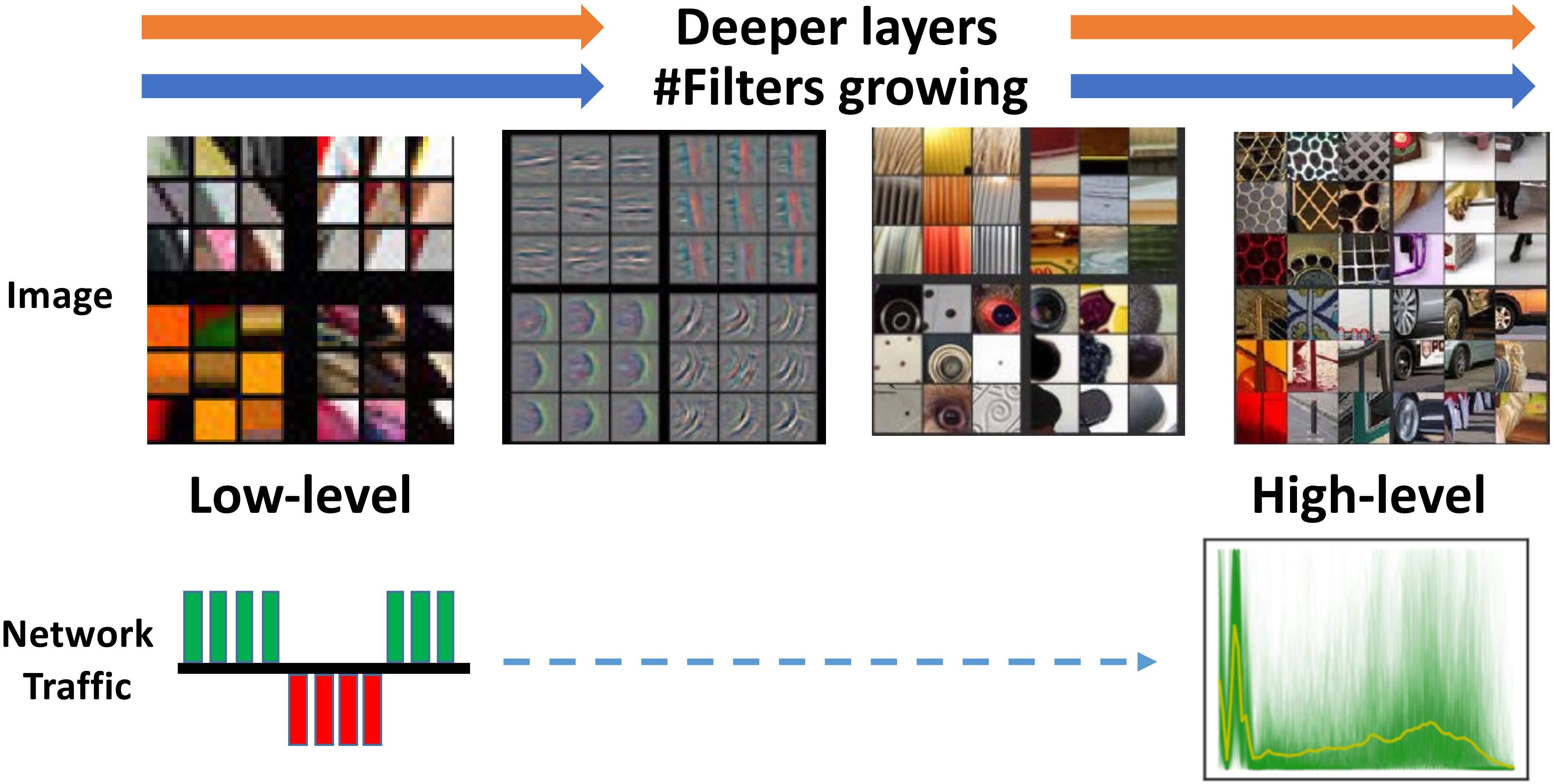
CNN Model

Effective?

Deep Fingerprinting

Deep Fingerprinting

DF Model: Improved Design of CNN



Deeper layers

#Filters growing

Image

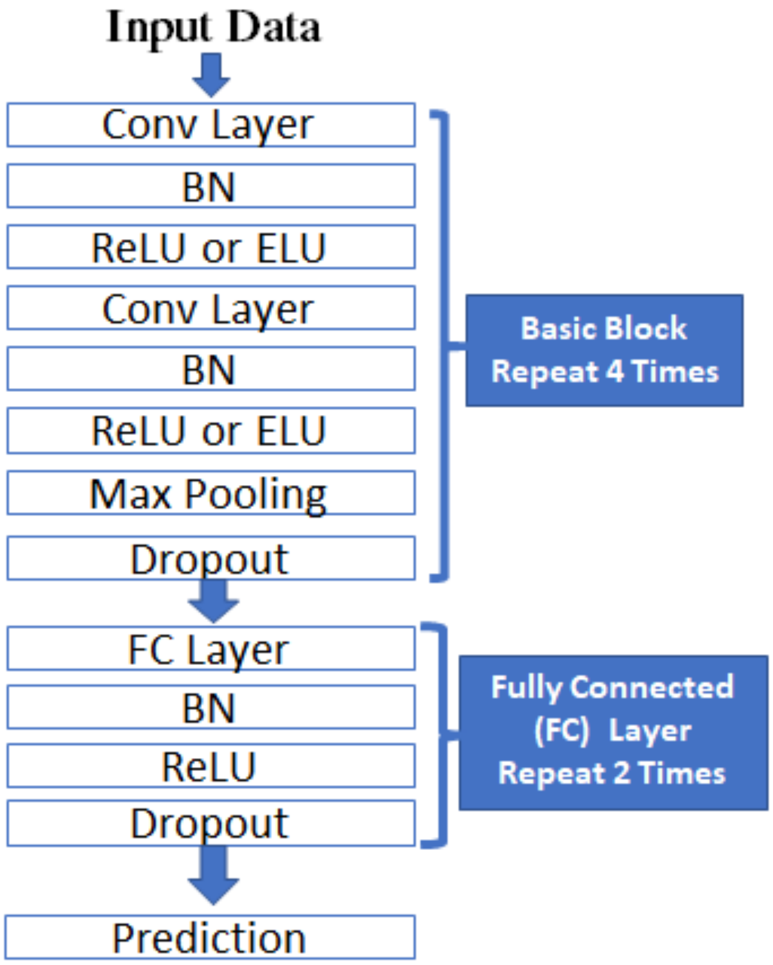
Low-level

High-level

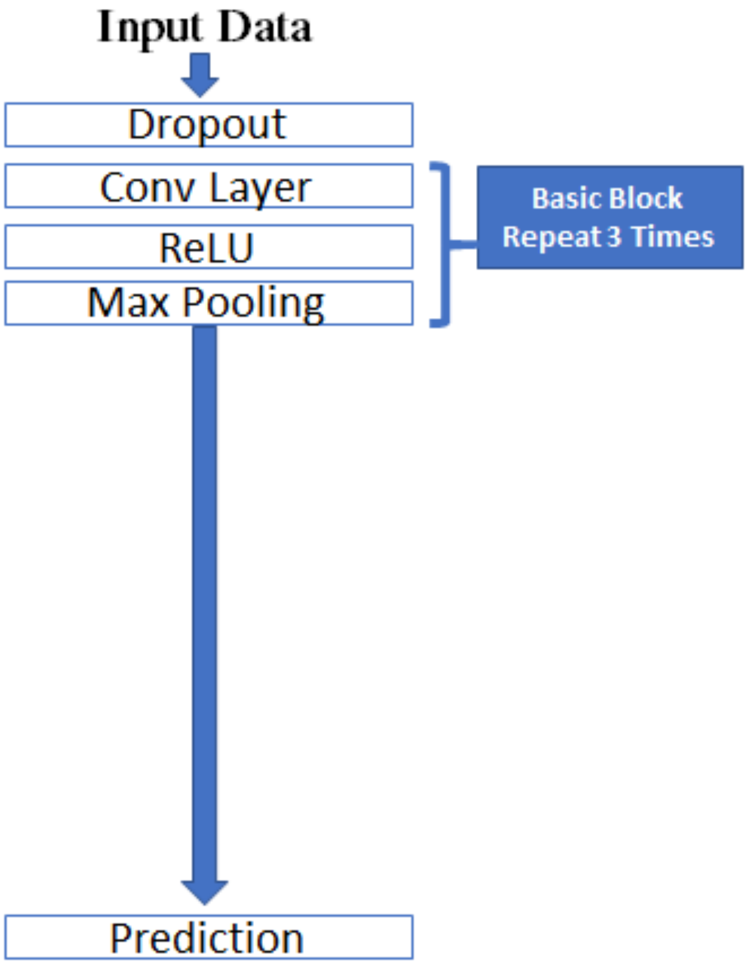
Network Traffic

Zeiler and Fergus. "Visualizing and understanding convolutional networks". ECCV, 2014.

Deep Fingerprinting

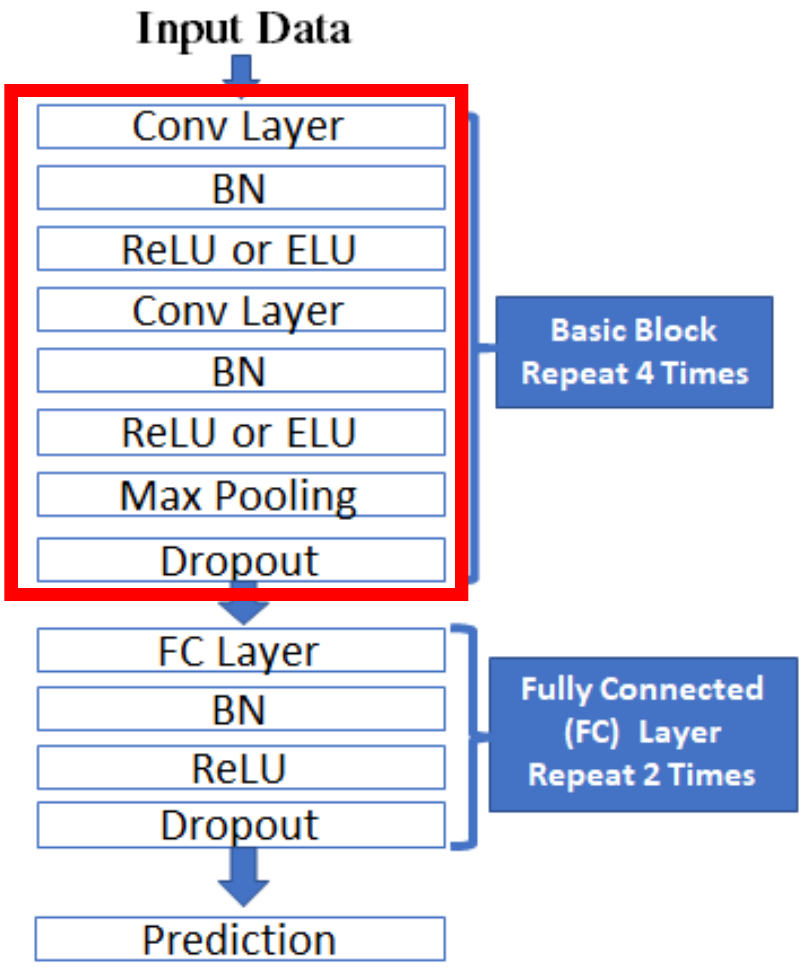


DF Model
(Our)

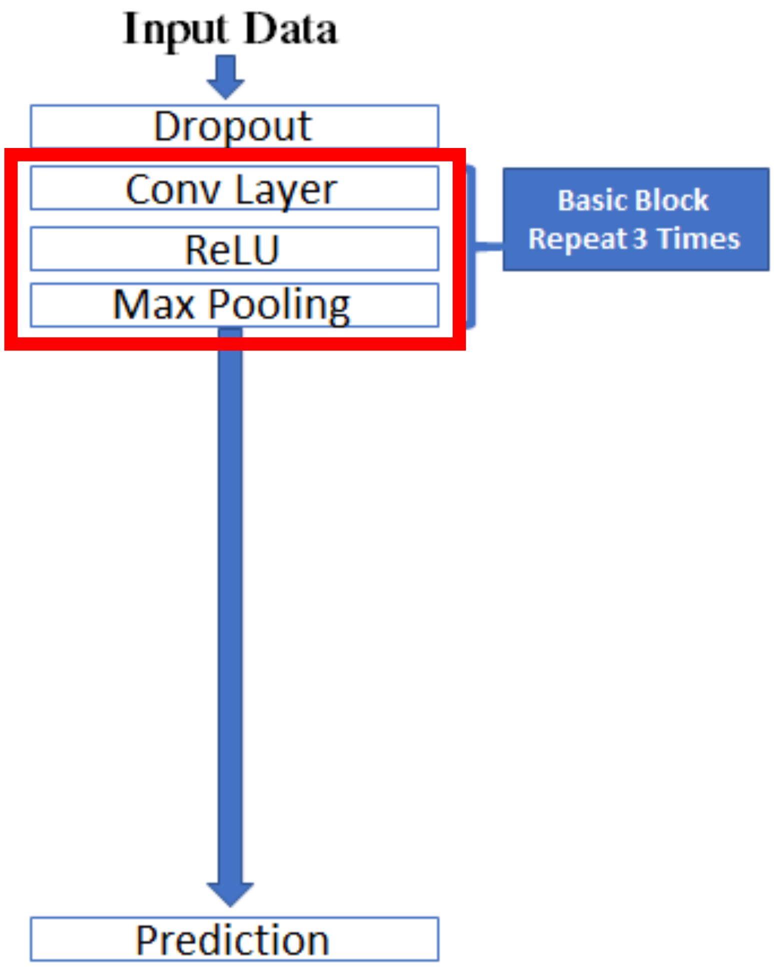


AWF Model
(Rimmer et al.)

Deep Fingerprinting

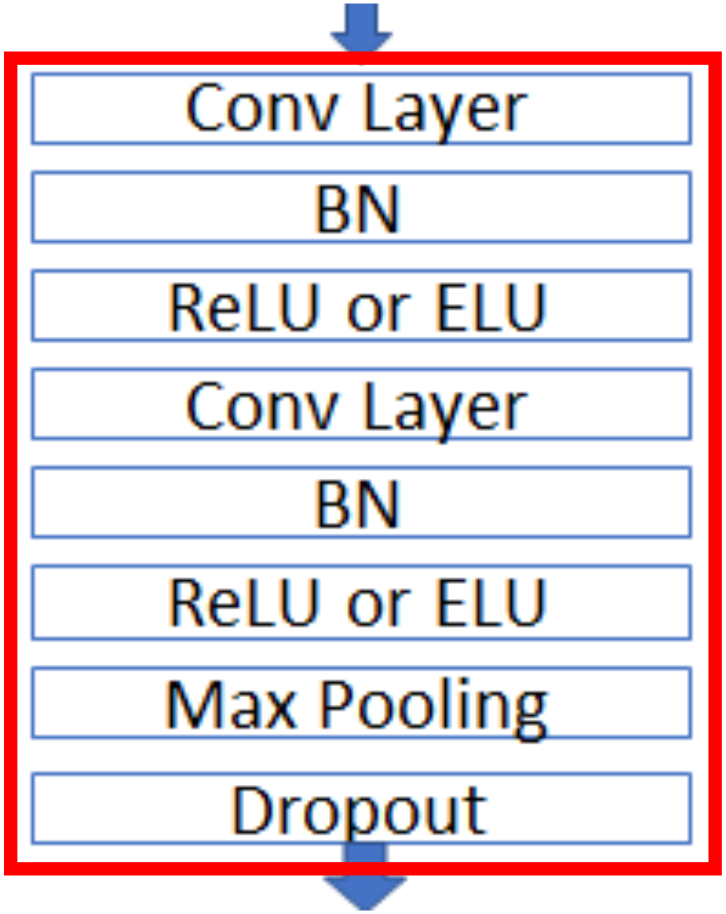


DF Model
(Our)



AWF Model
(Rimmer et al.)

Deep Fingerprinting



Conv Layer

BN

ReLU or ELU

Conv Layer

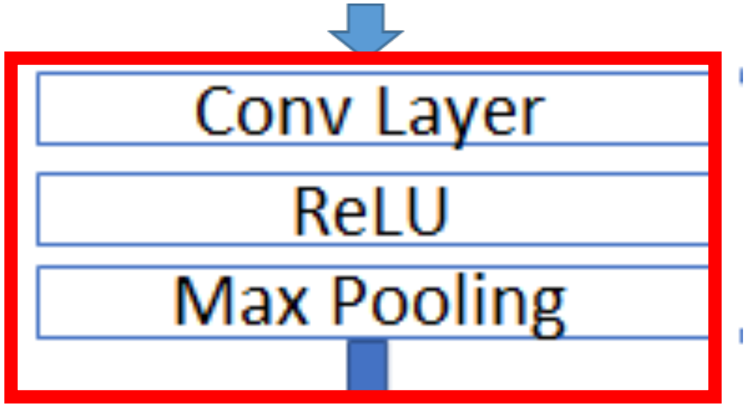
BN

ReLU or ELU

Max Pooling

Dropout

DF Model (Our)



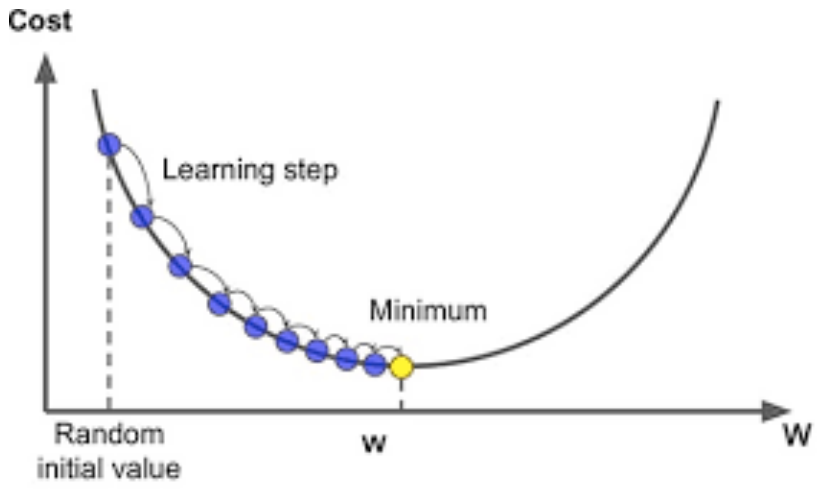
Conv Layer

ReLU

Max Pooling

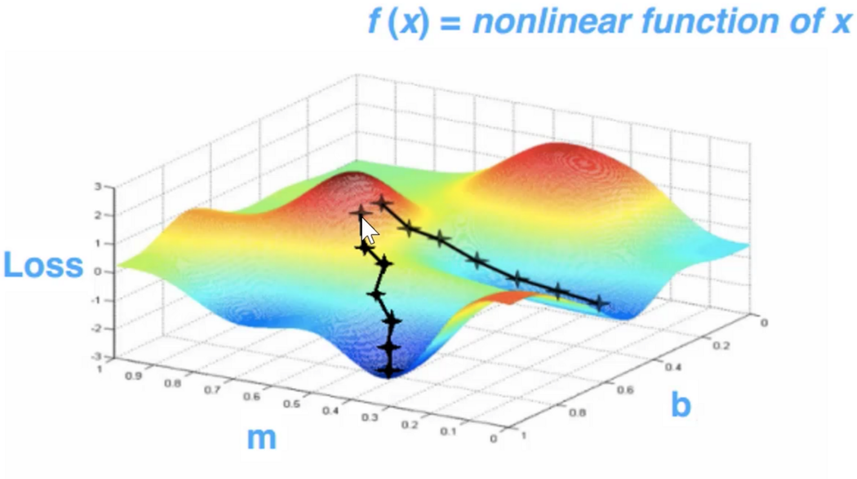
AWF Model (Rimmer et al.)

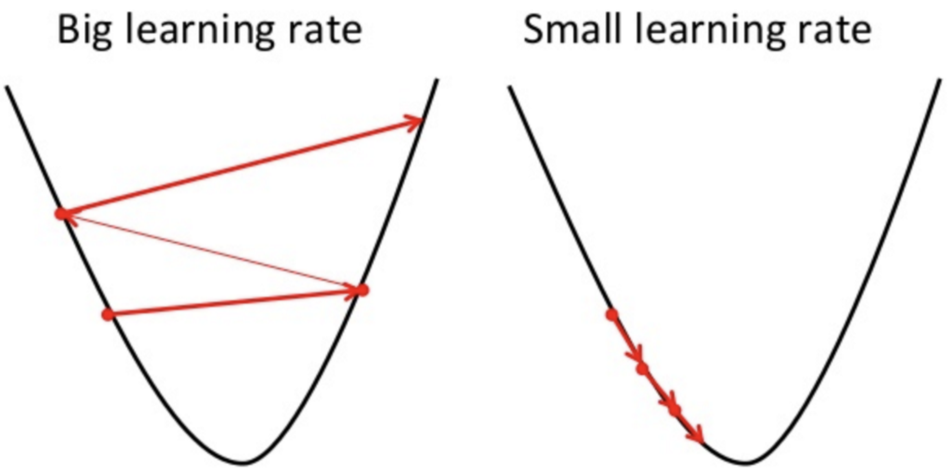
Batch Normalization



Gradient Descent

Gradient Descent





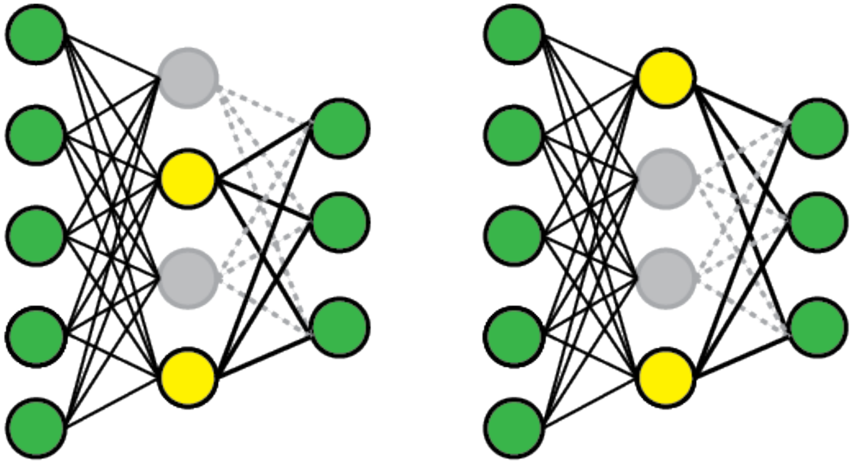
BN: 1 ft. max

https://saugatbhattarai.com.np/what-is-gradient-descent-in-machine-learning/
https://towardsdatascience.com/gradient-descent-in-a-nutshell-eaf8c18212f0
https://medium.com/@julian.harris/stochastic-gradient-descent-in-plain-english-9e6c10cdba97

Dropout

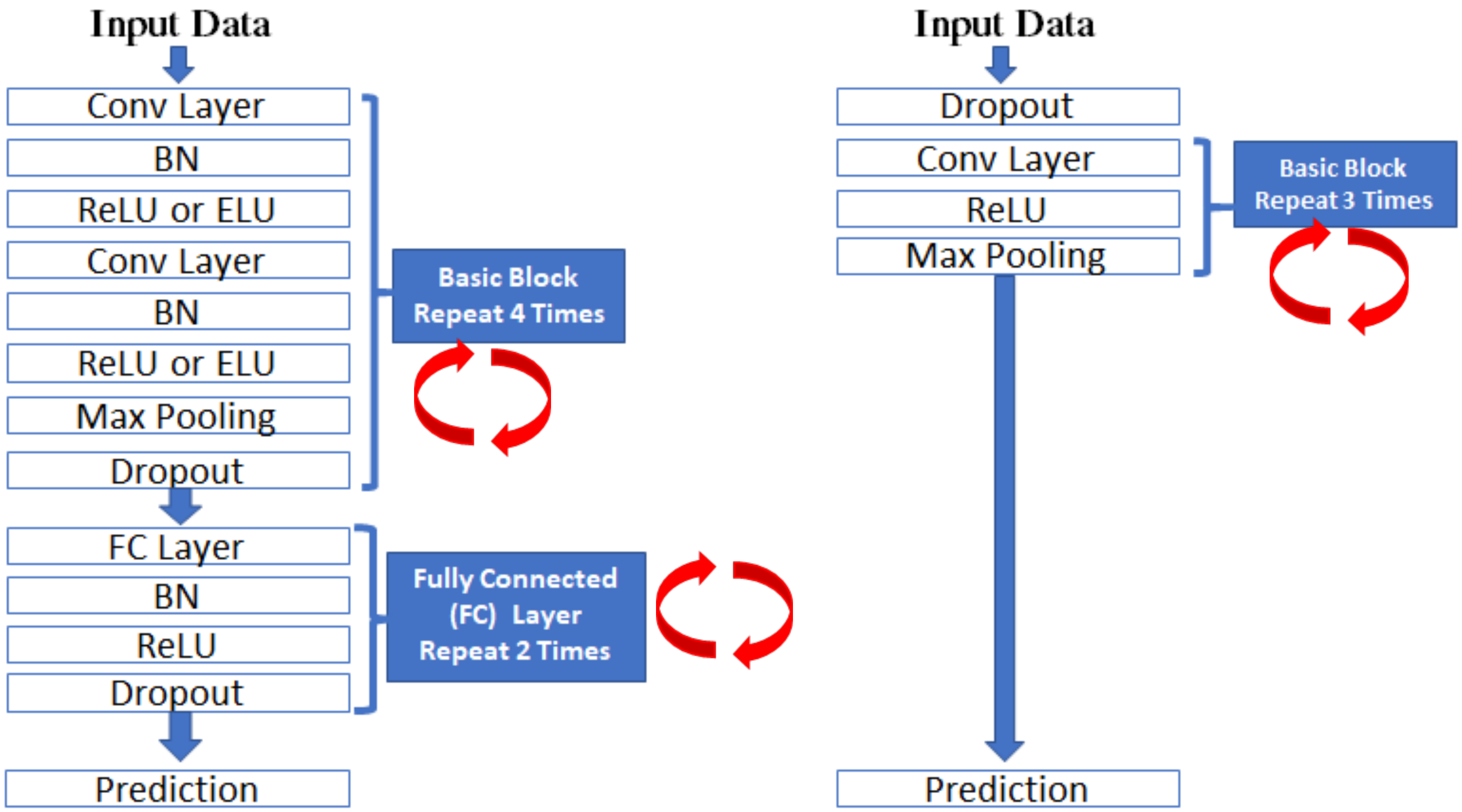
Train

Test



https://stats.stackexchange.com/questions/201569/difference-between-dropout-and-dropconnect

Deep Fingerprinting



DF Model
(Our)

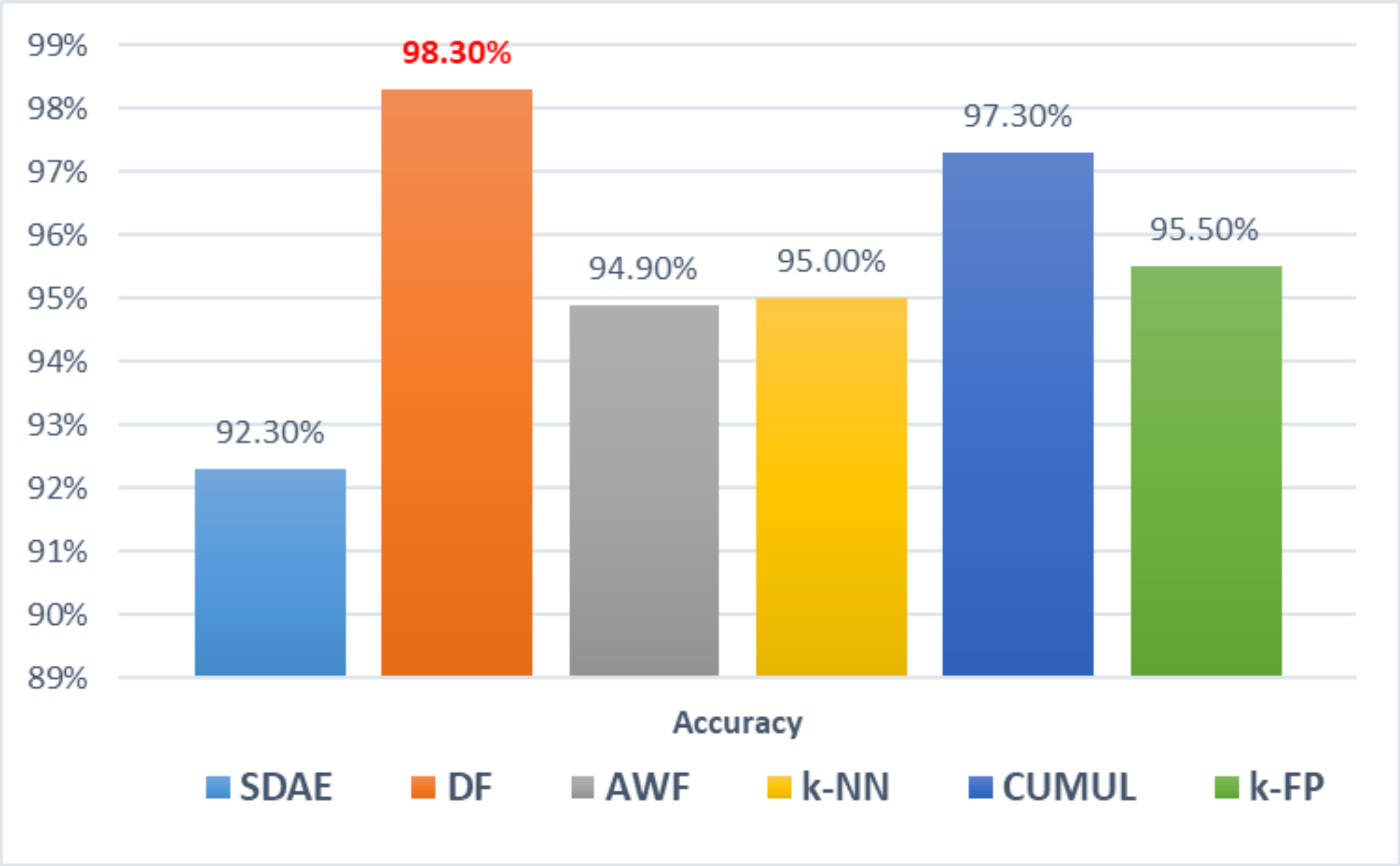
~3X deeper

AWF Model
(Rimmer et al.)

Deep Fingerprinting

Experimental Evaluation

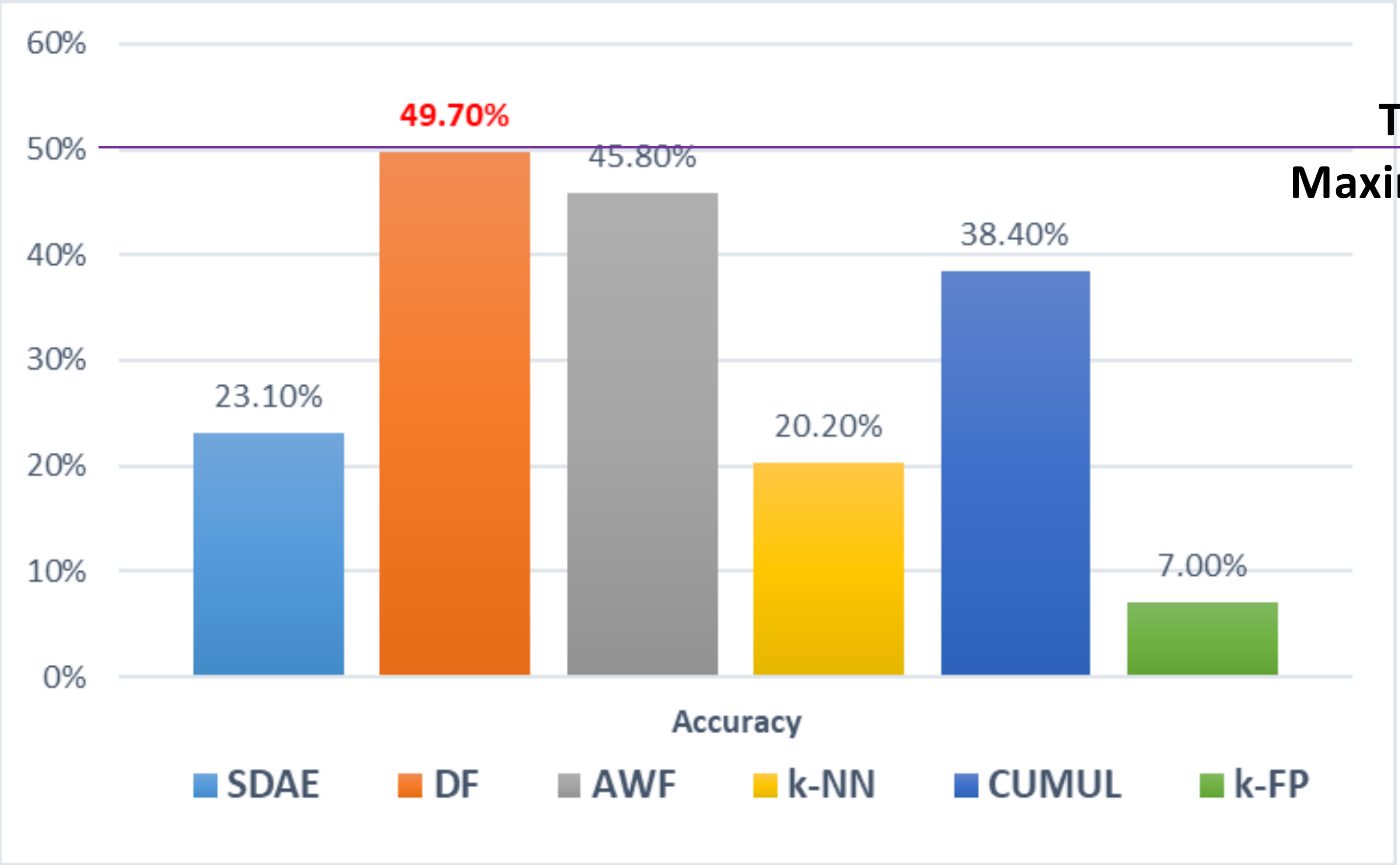
- Non-defended Dataset



Deep Fingerprinting

Experimental Evaluation

- Walkie-Talkie
  - 31% Bandwidth, 34% Latency

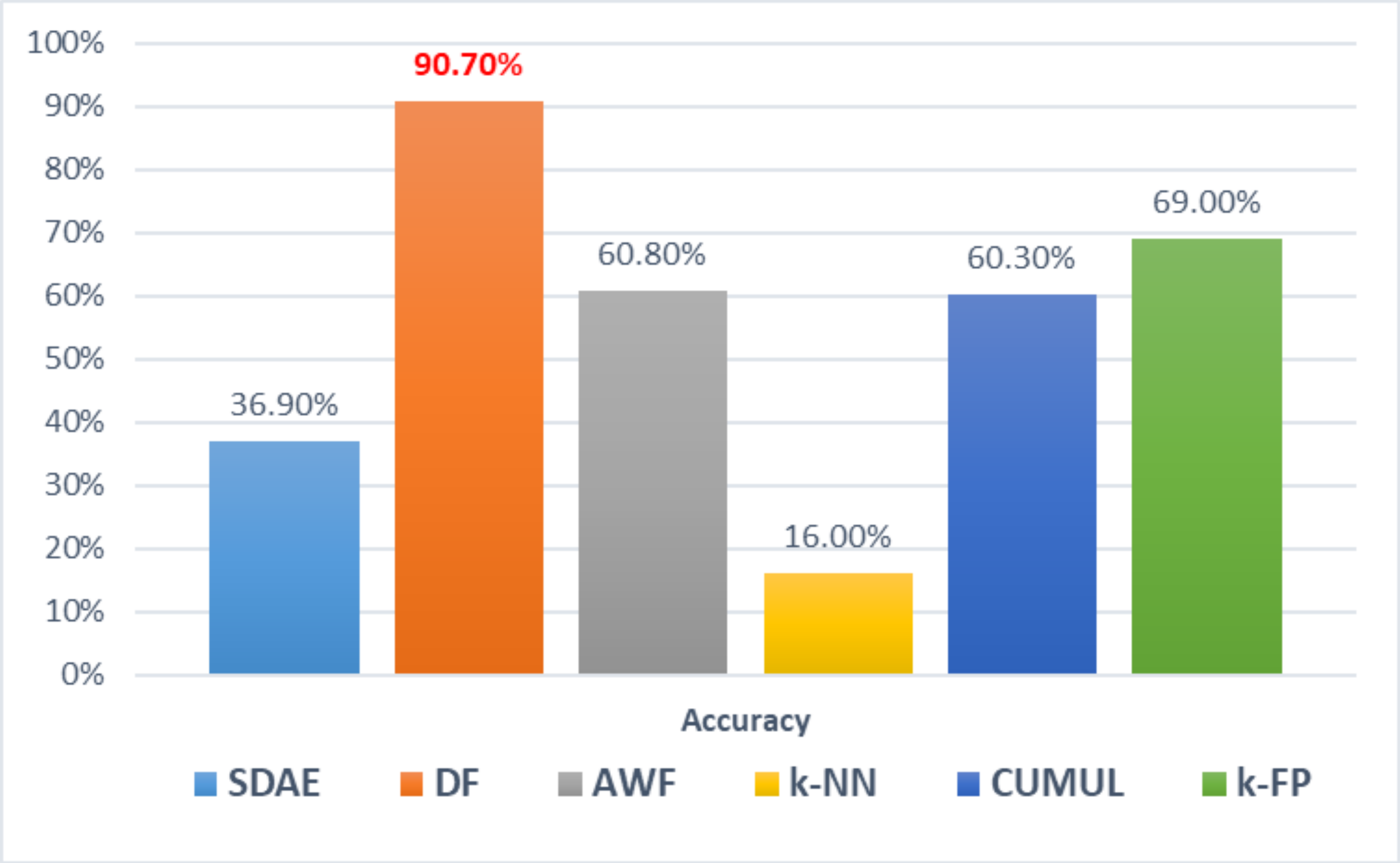


Theoretical Maximum Accuracy

Deep Fingerprinting

Experimental Evaluation

- WTF-PAD
  - 64% Bandwidth, 0% Latency

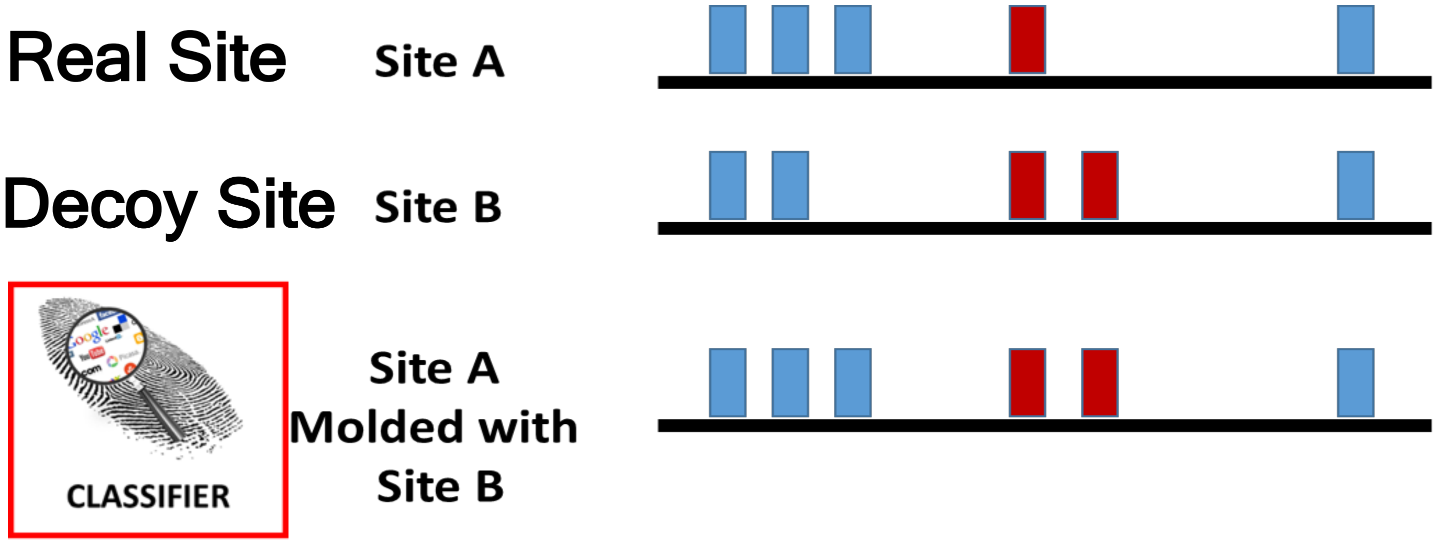


Deep Fingerprinting

Walkie-Talkie: Discussion

- At most 50% accuracy in closed world

- Top-N prediction



Real Site

Site A

Decoy Site

Site B

Site A Molded with Site B

DF: Top-2 prediction → 98.44 Accuracy

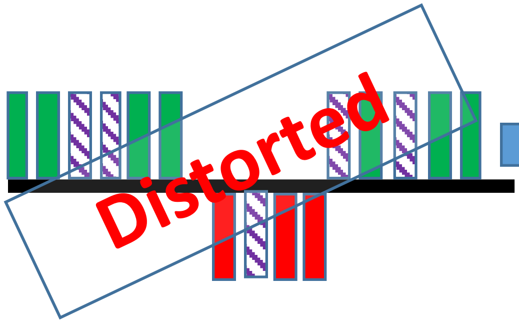
Conclusion

Conclusion

Distorted

CNN Model

Effective?



Network Traffic
with Defenses

DF Model

>90% Accuracy
(WTF-PAD)

This material is based upon work supported by the National Science Foundation under Grant No. CNS-1423163, CNS-1722473, and CNS-1816851. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

I’m back, baby!

CENTER FOR CYBERSECURITY

iSec
THE INFORMATION SECURITY LAB

COSIC

https://github.com/deep-fingerprinting/df

Deep Fingerprinting

Undermining Website Fingerprinting Defenses with Deep Learning