CENTER FOR CYBERSECURITY

http://www.rit.edu/cybersecurity

Matthew Wright, PhD

Director of the Center for Cybersecurity

Professor of Computing Security

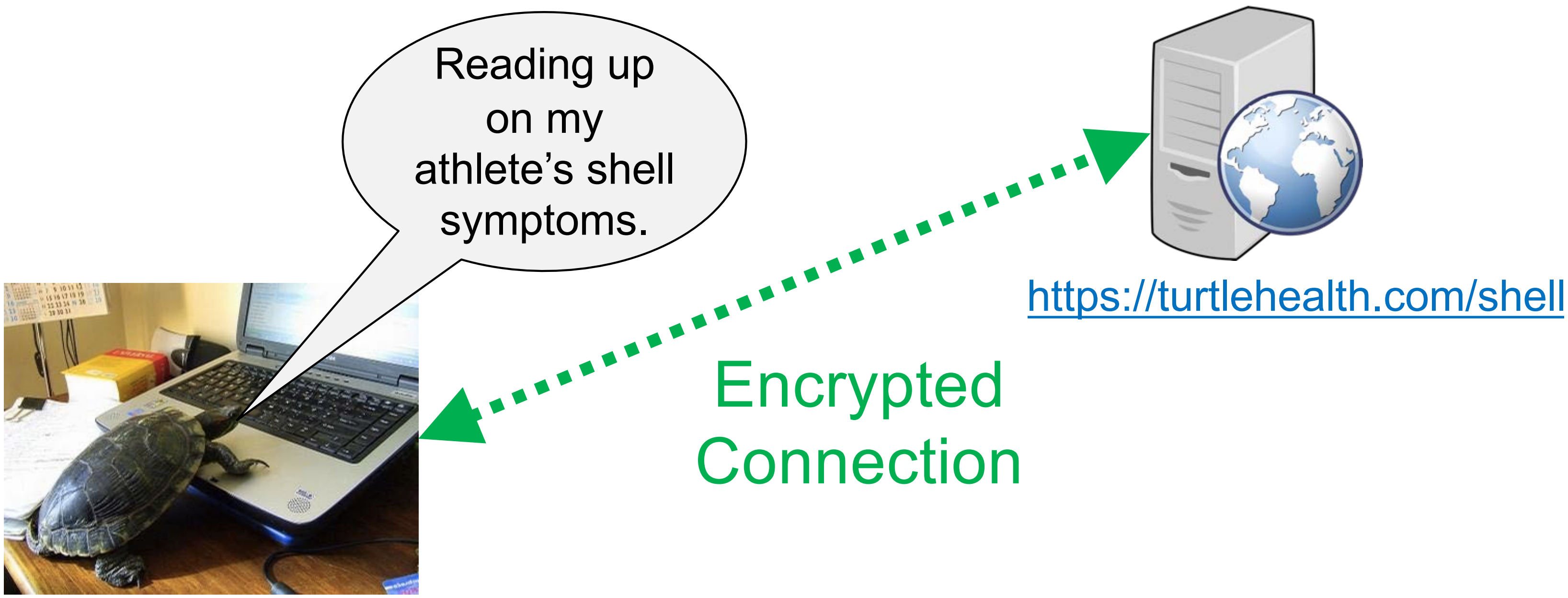
RIT | Rochester Institute of Technology

How Attackers Can Read Your Encrypted Traffic …

and Can We Stop It?

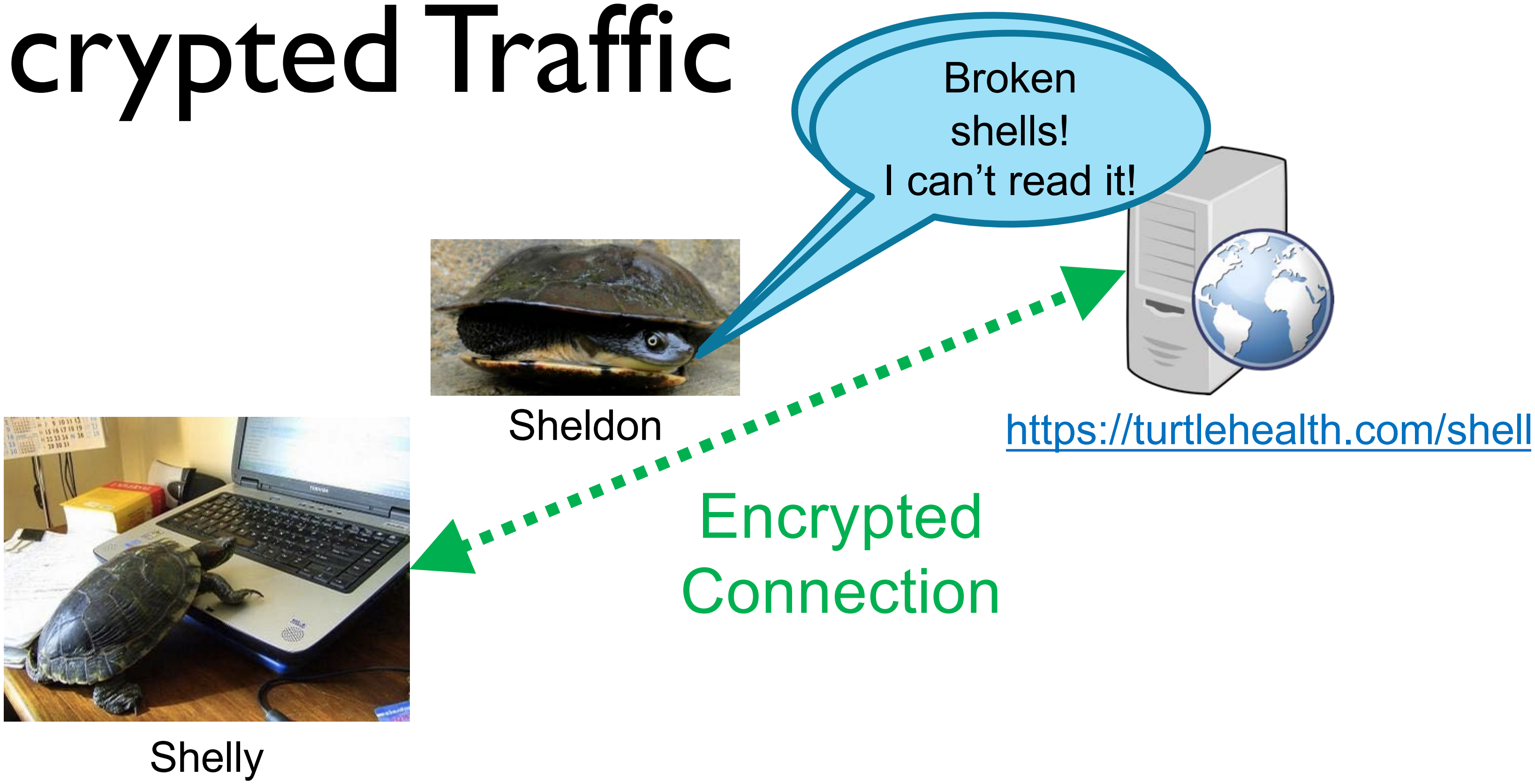
RIT

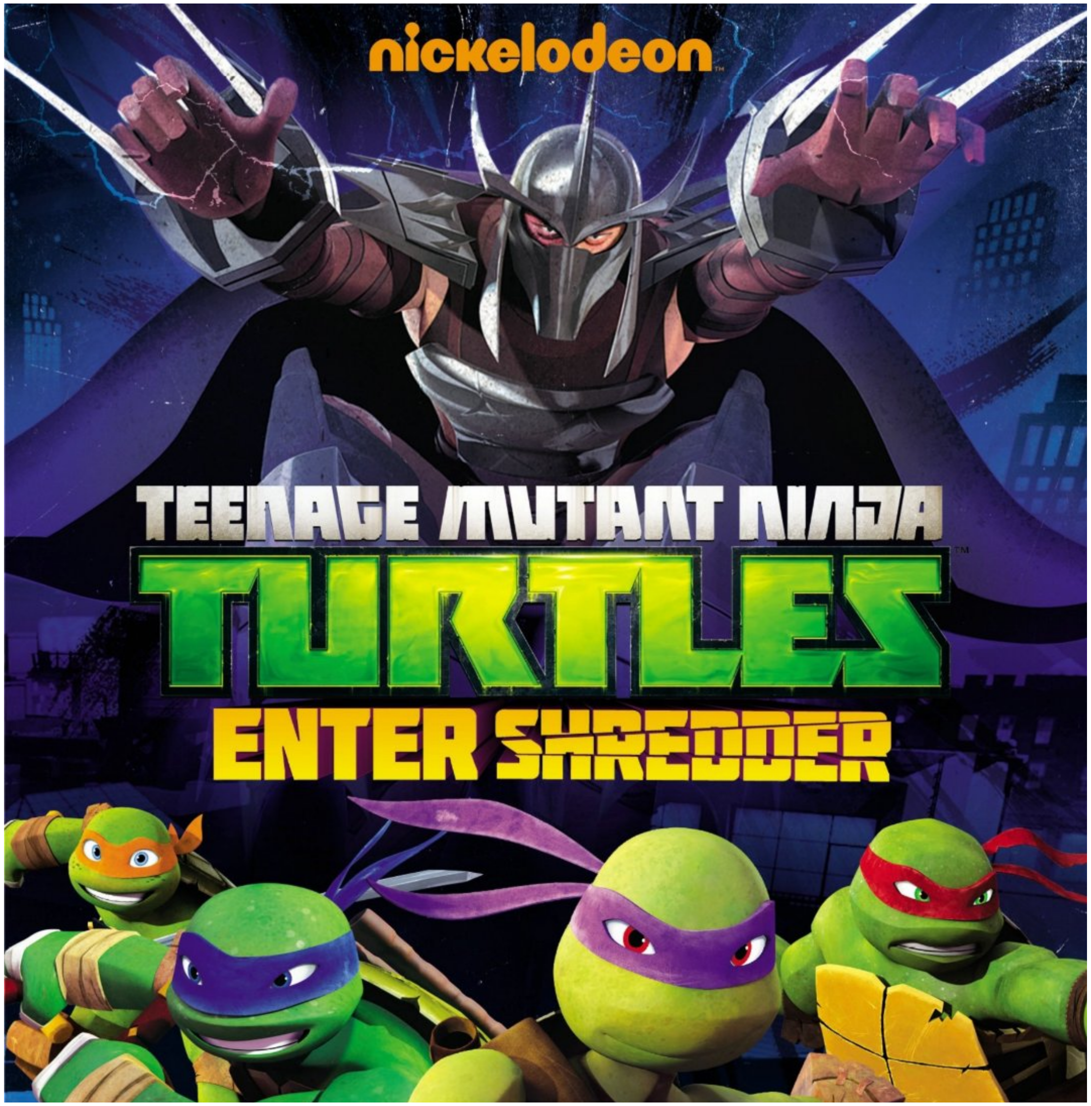
Encrypted Traffic



Shelly

Encrypted Traffic

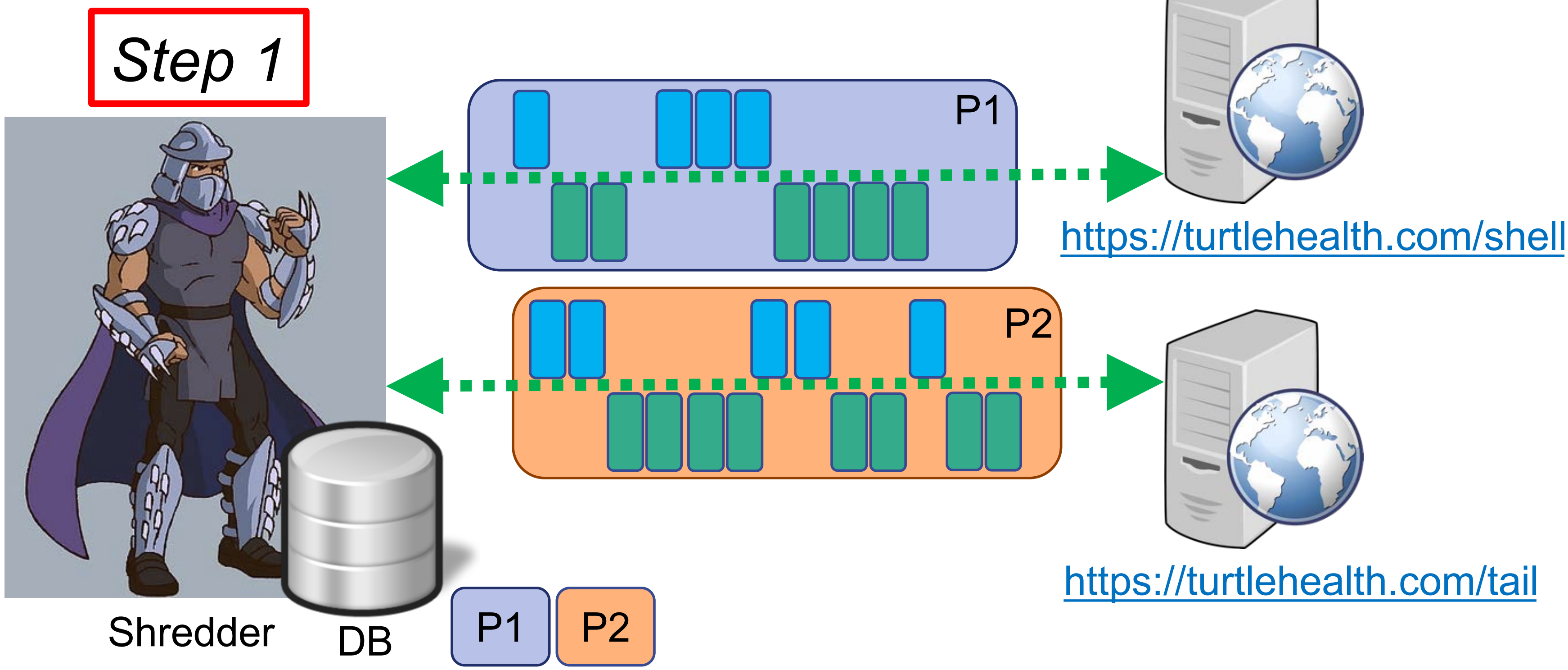




http://www.nickandmore.com/wordpress/wp-content/uploads/2013/08/cover.jpg

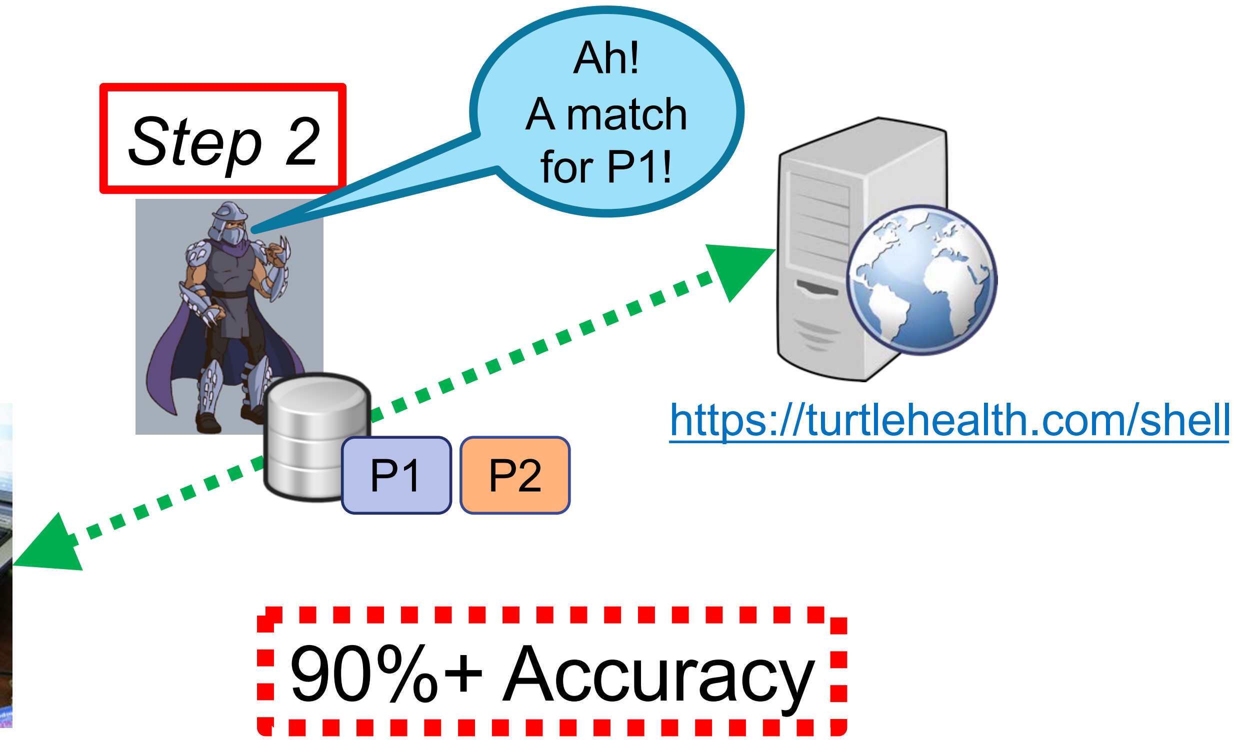
RIT

Website Fingerprinting

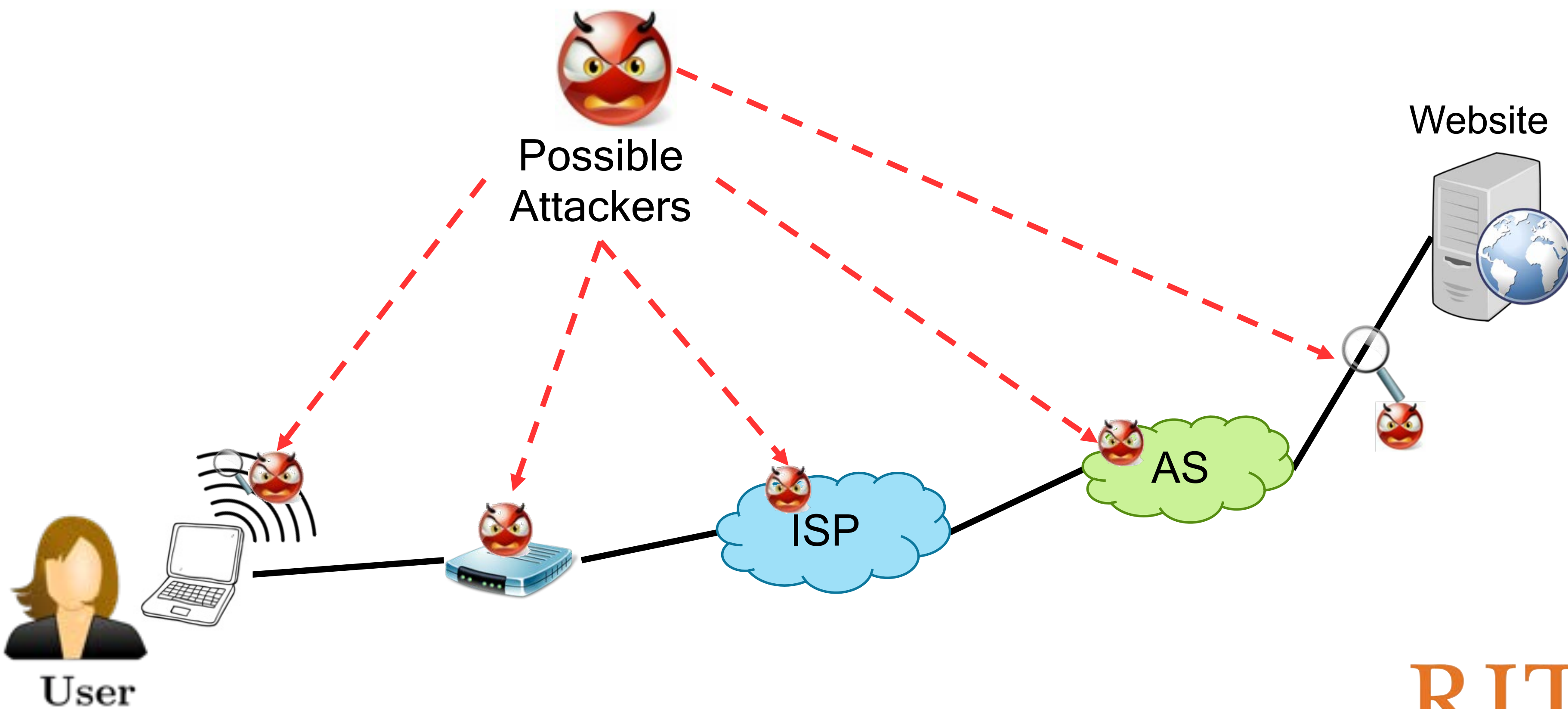


RIT

Website Fingerprinting



Shelly


Possible
Attackers
Website
ISP
AS
User

Meet Jerome

RIT

Jerome* Goes Online

ESPN

FOX SPORTS

BLACK LIVES MATTER

NRA

NARAL PRO-CHOICE AMERICA

THE CENTER
THE LESBIAN, GAY, BISEXUAL & TRANSGENDER COMMUNITY CENTER

glaad

PRIDE AT WORK

betterhelp

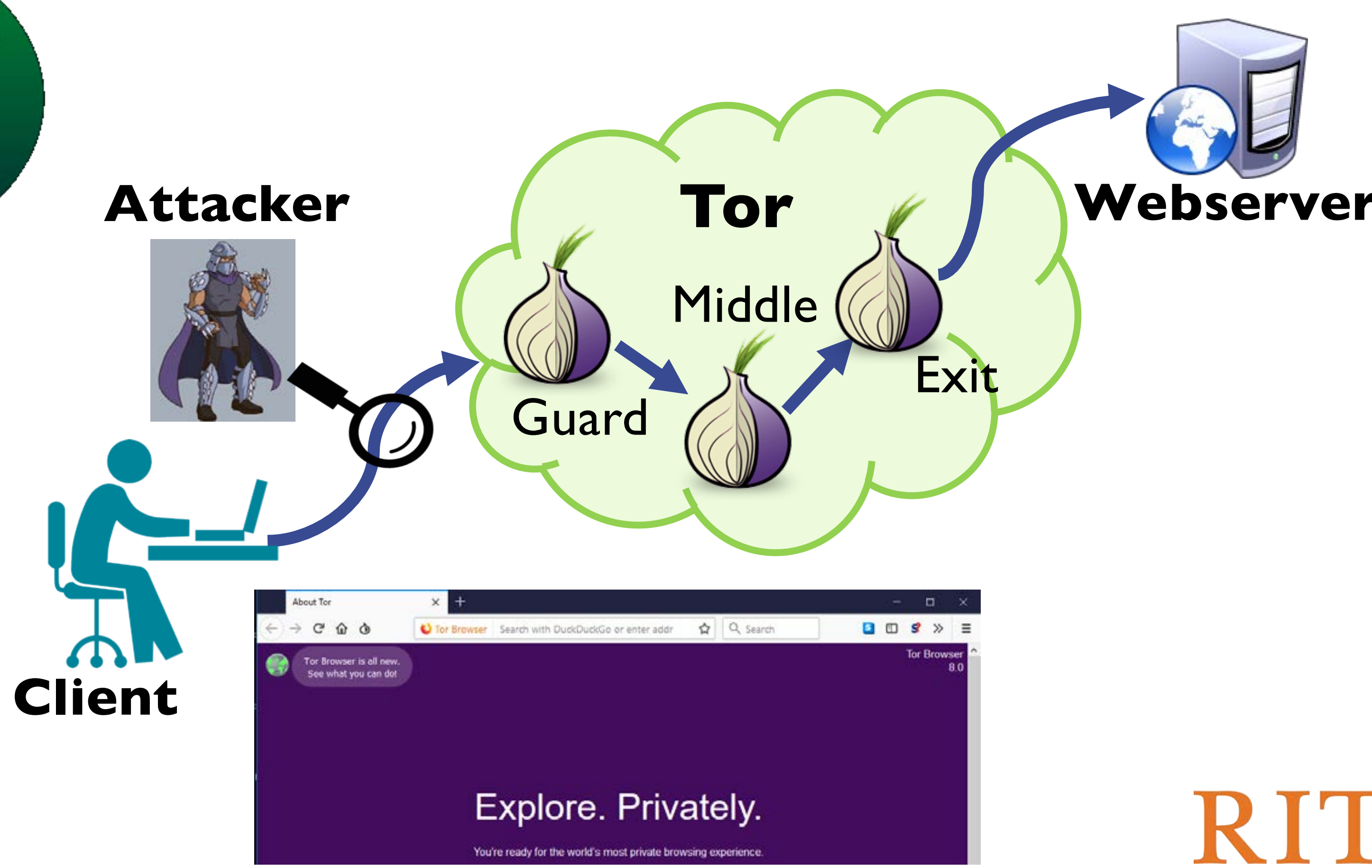
CORE MENTAL HEALTH COUNSELING

Transitions COUNSELING SERVICES
Counseling & Life Coaching

WHITE SPRUCE COUNSELING
individual & flexible counseling

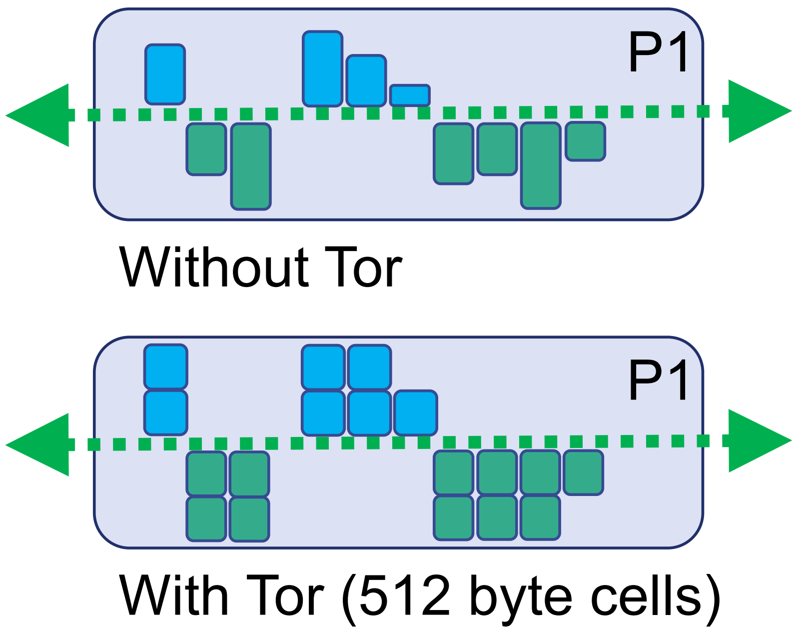
* Not related to actual interests of any Jeromes Bettises

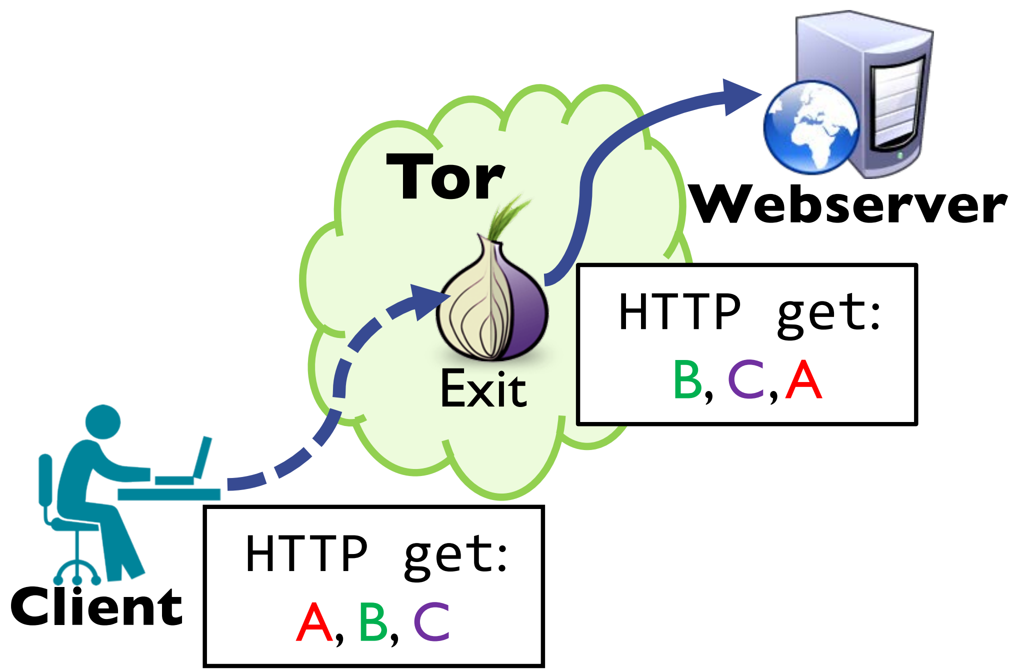
RIT


Attacker
Tor
Middle
Exit
Guard
Webserver
Client
About Tor
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.

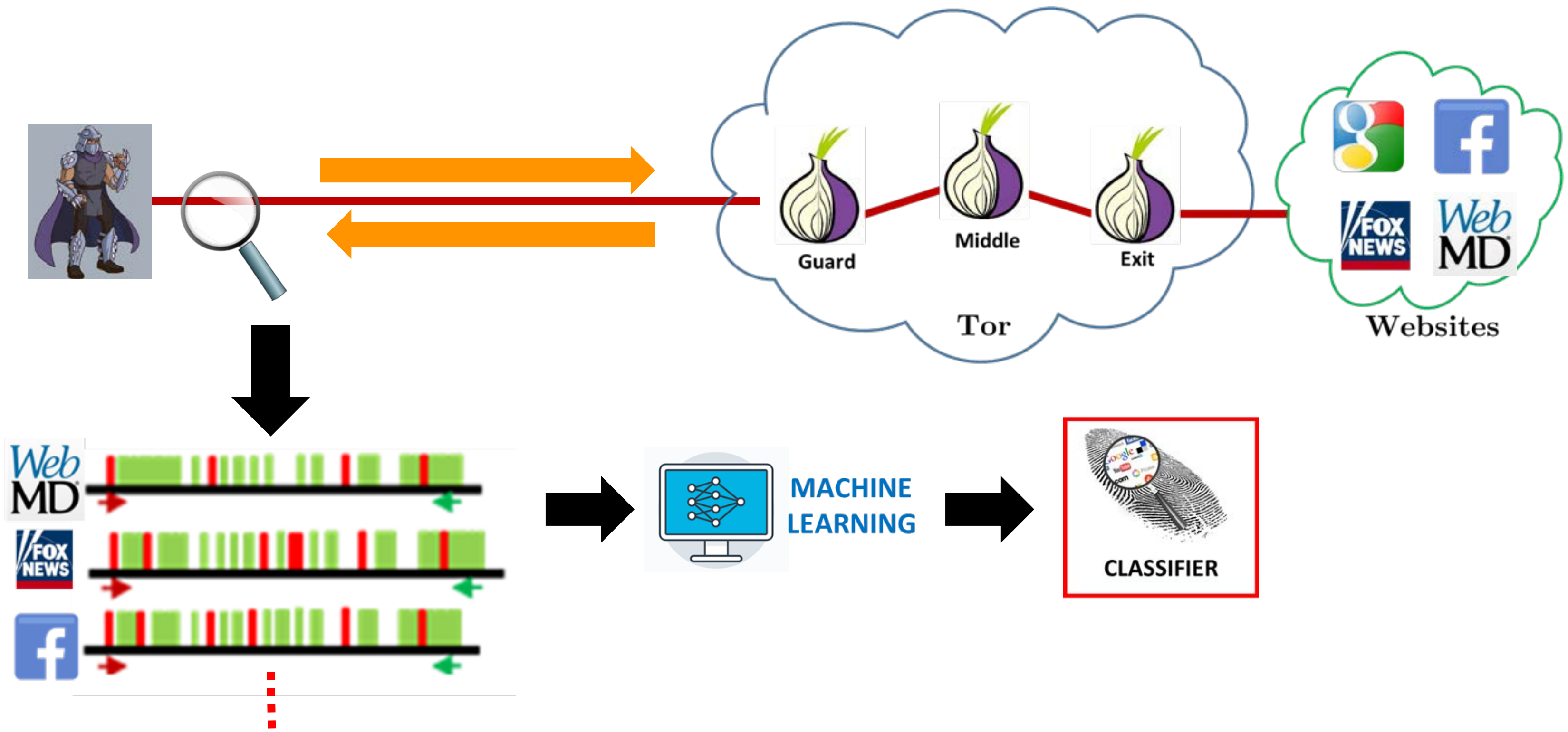
RIT

Tor's WF Defenses



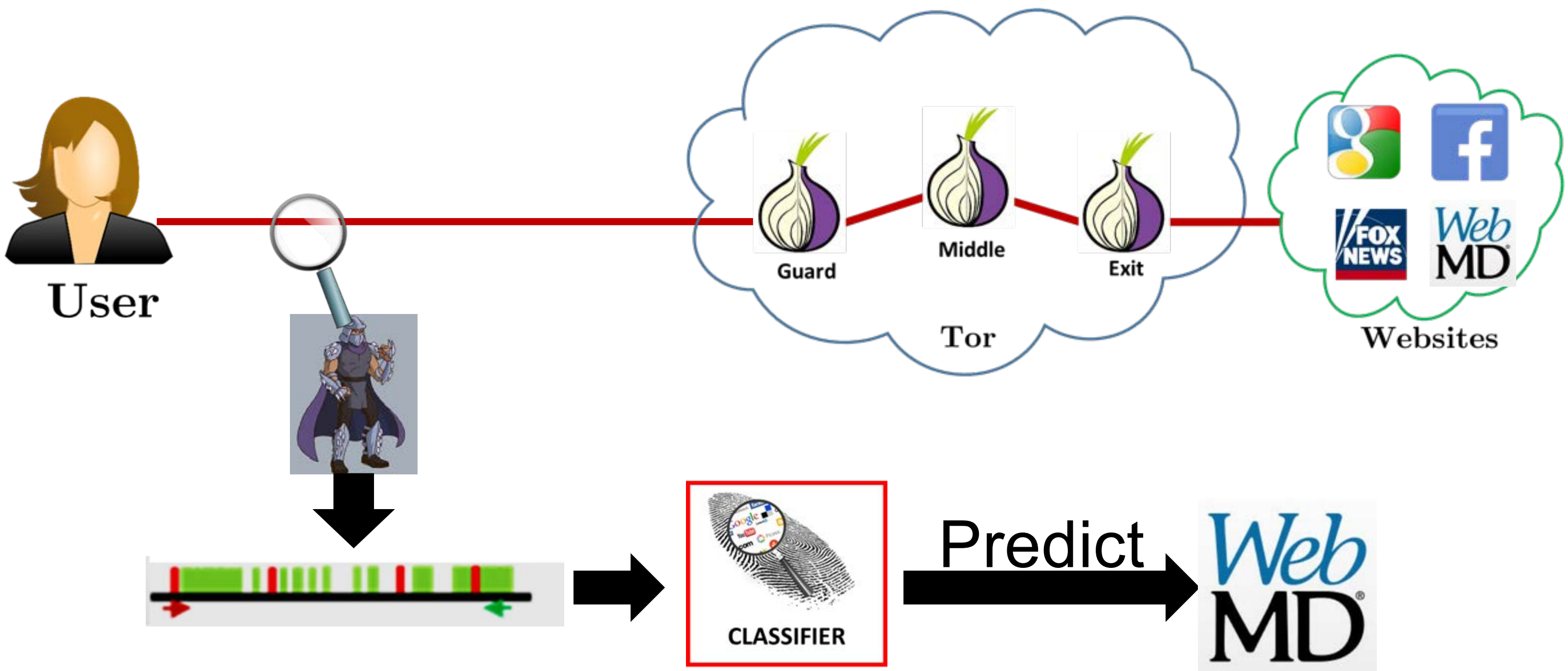


WF in Tor



1. Train the classifier

WF in Tor



2. Perform the attack

90%+ Accuracy*

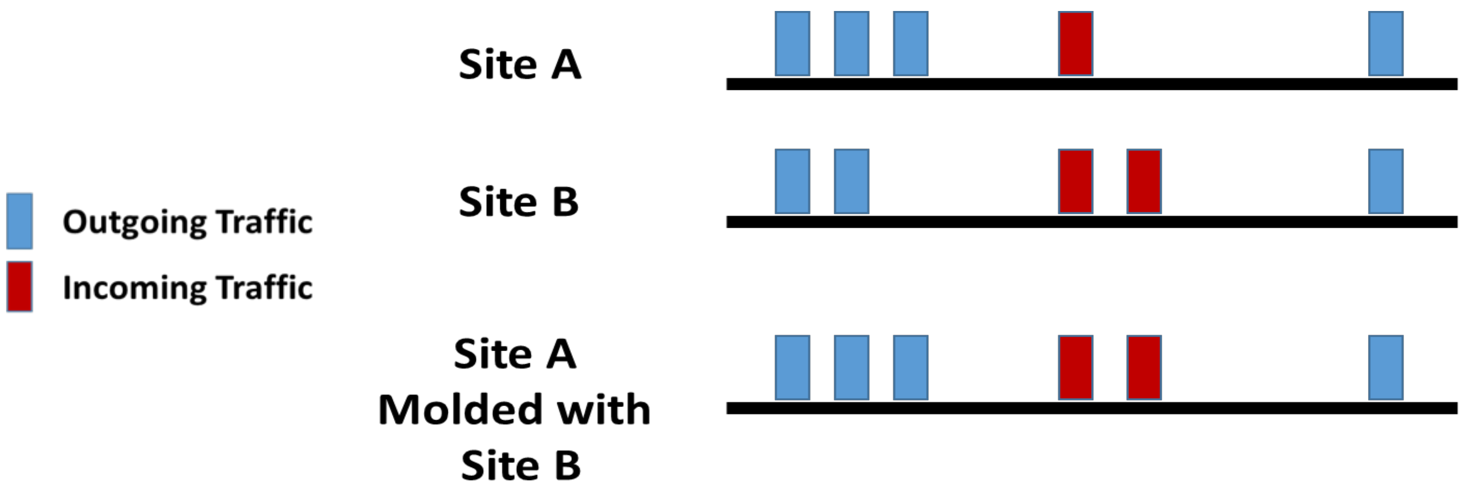
* For ~100 sites, not pages

++?

RIT

Walkie-Talkie (W-T) [WG17]

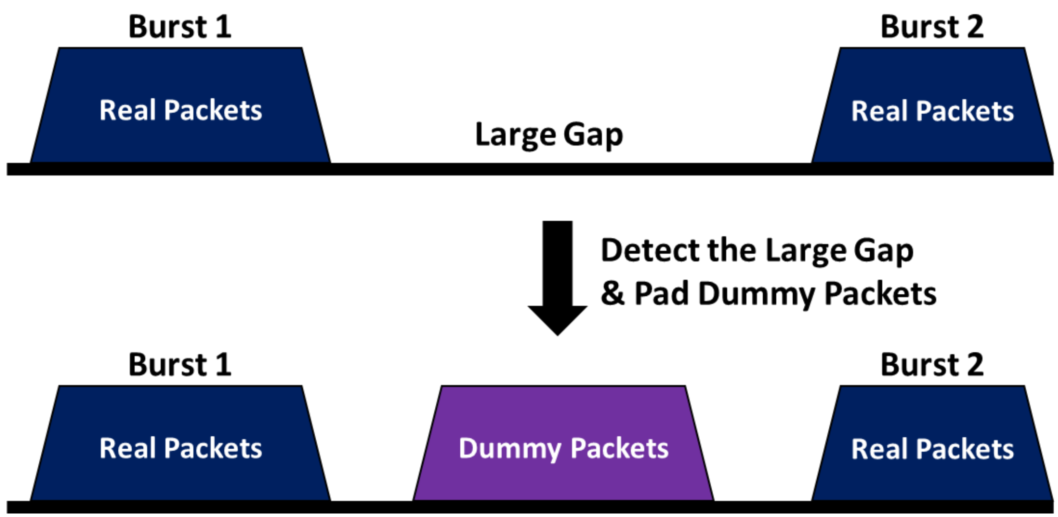
- 31% bandwidth overhead; 34% added delay
- Reduce accuracy < 30%



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

WTF-PAD [JIP16]

- 54% bandwidth overhead; No added delay*
- 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.

RIT


WTF!?!

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

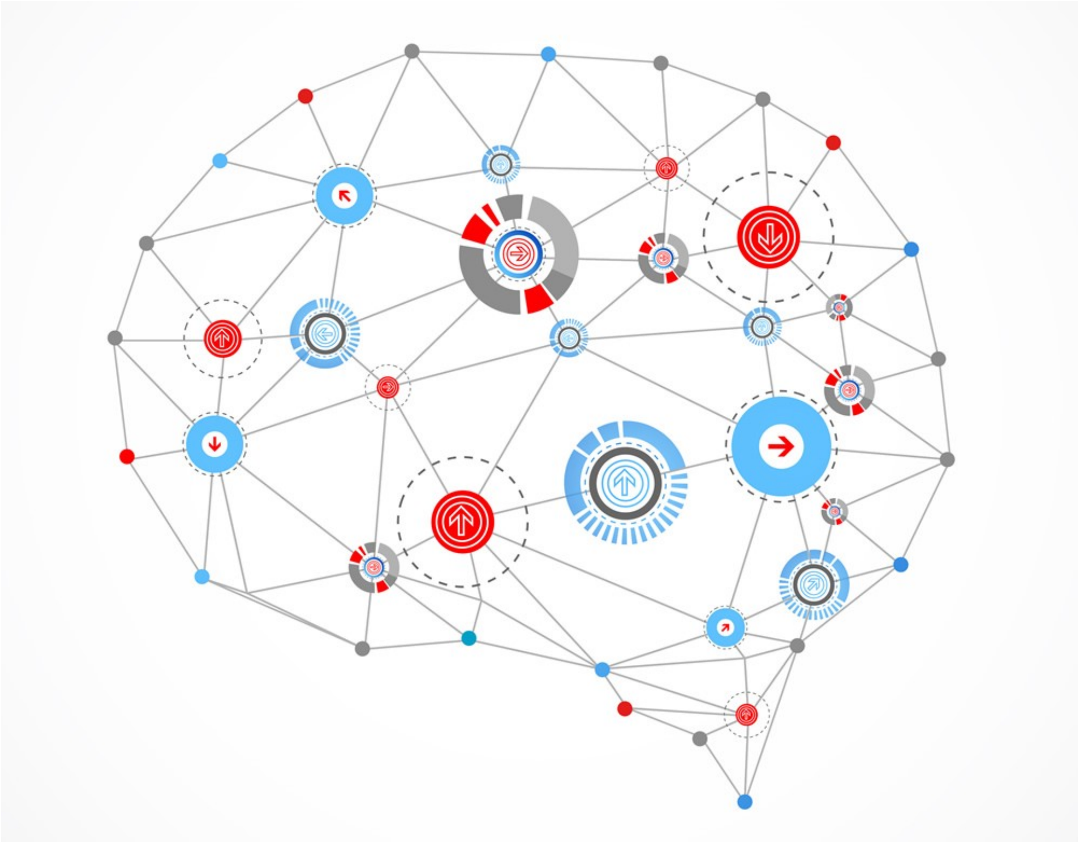
Payap

Mohsen

Marc

RIT

Deep Learning

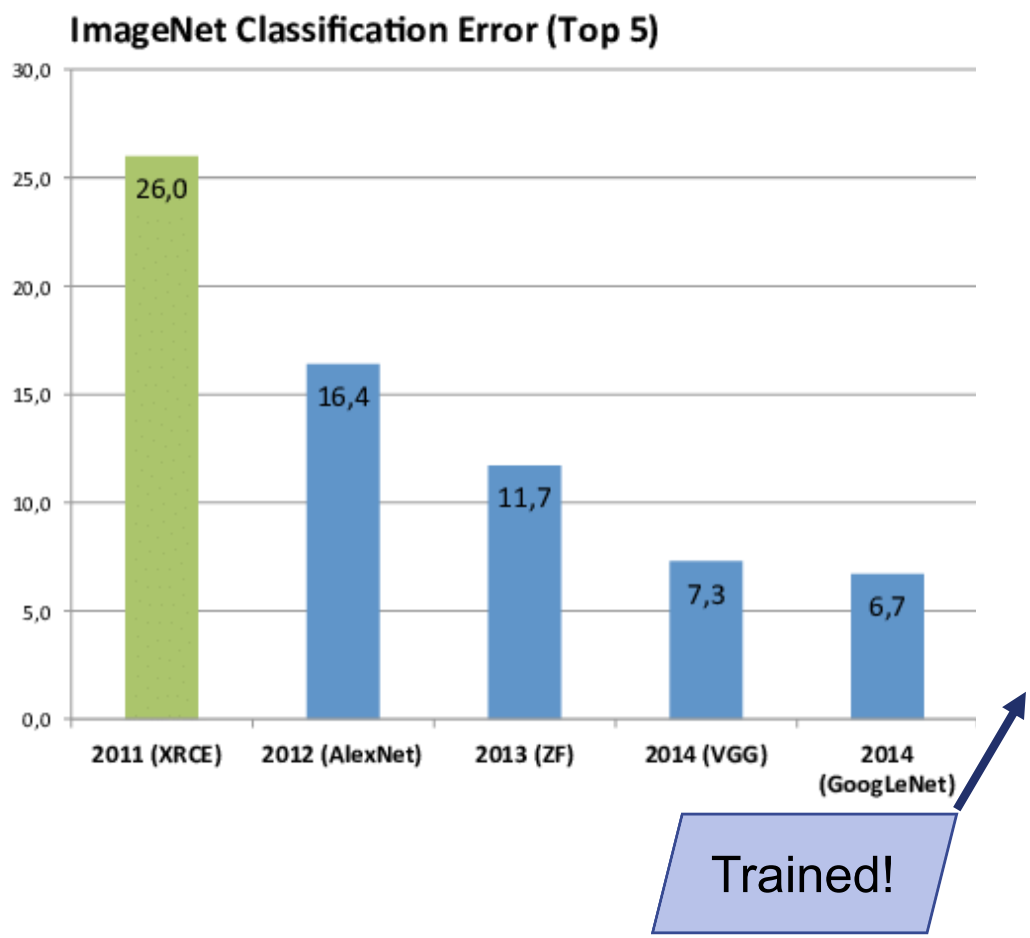


https://codeburst.io/deep-learning-what-why-dd77d432f182

ILSVRC: 1.2M images, 1.2K categories

RIT

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!

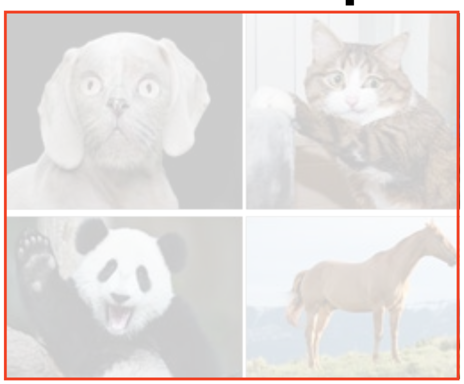
RIT

Research Goals (1)

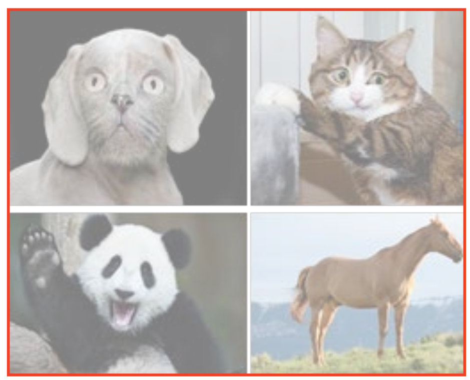
- Prior work: early CNN

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

- Improvements of CNN in the literature



~55% Accuracy
AlexNet (2012)



~71% Accuracy
VGG19 (2014)

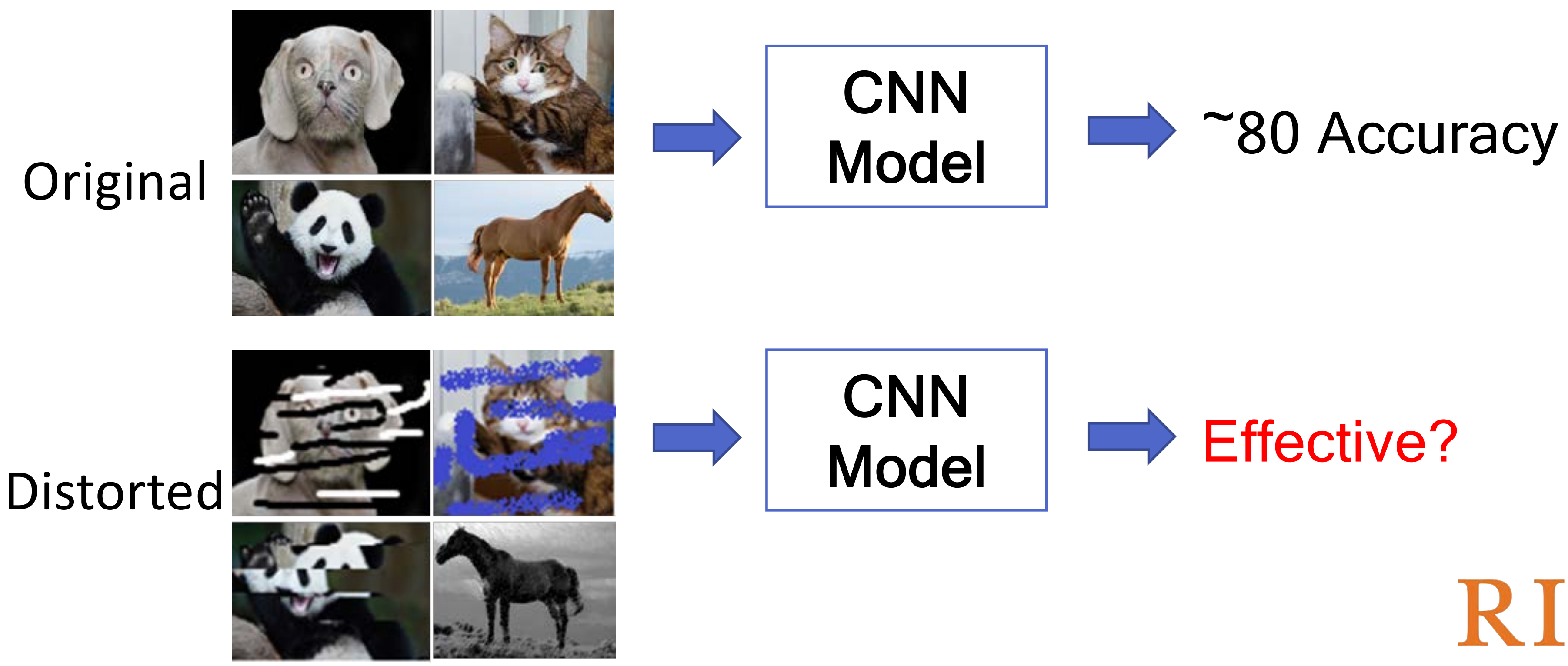
~80% Accuracy
Inception V4 (2016)

RIT

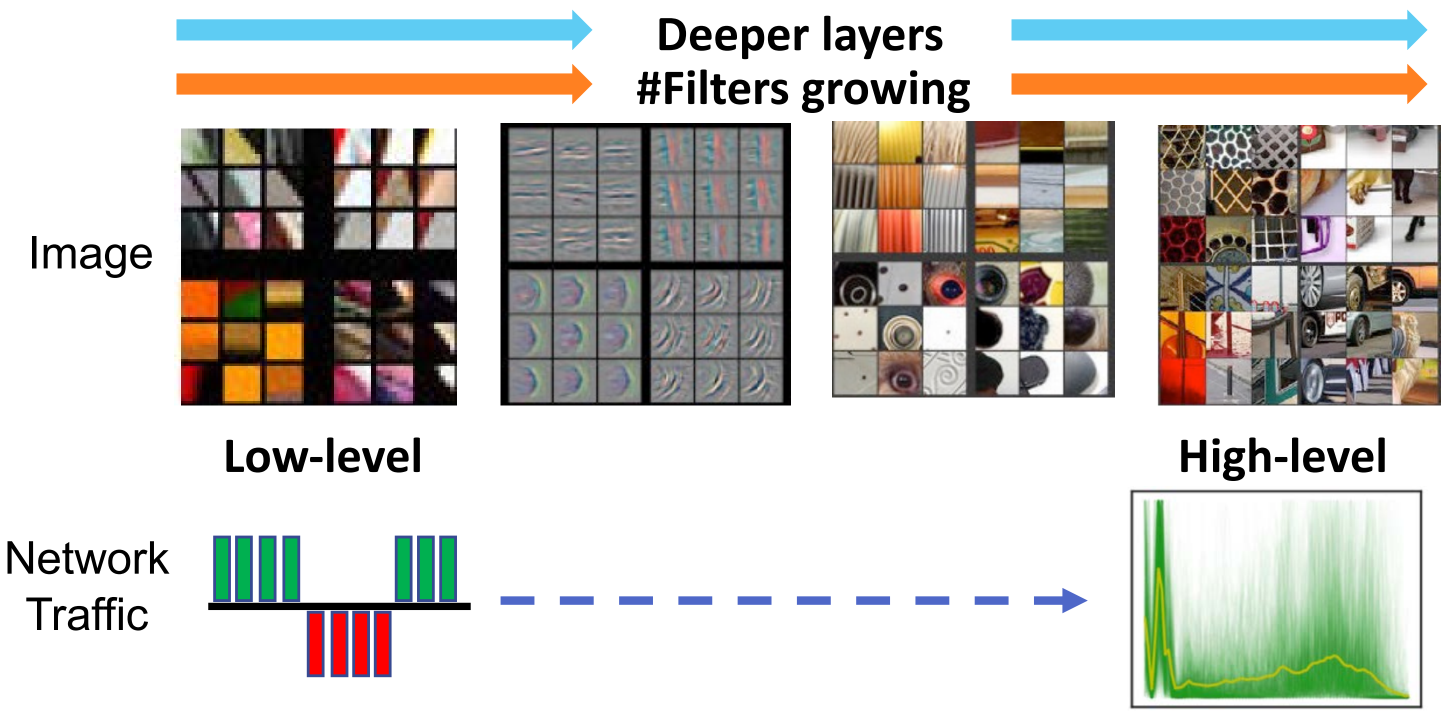
Canziani et al. An Analysis of Deep Neural Network Models for Practical Applications., arXiv:1605.07678

Research Goals (2)

- Evaluation against WF defenses

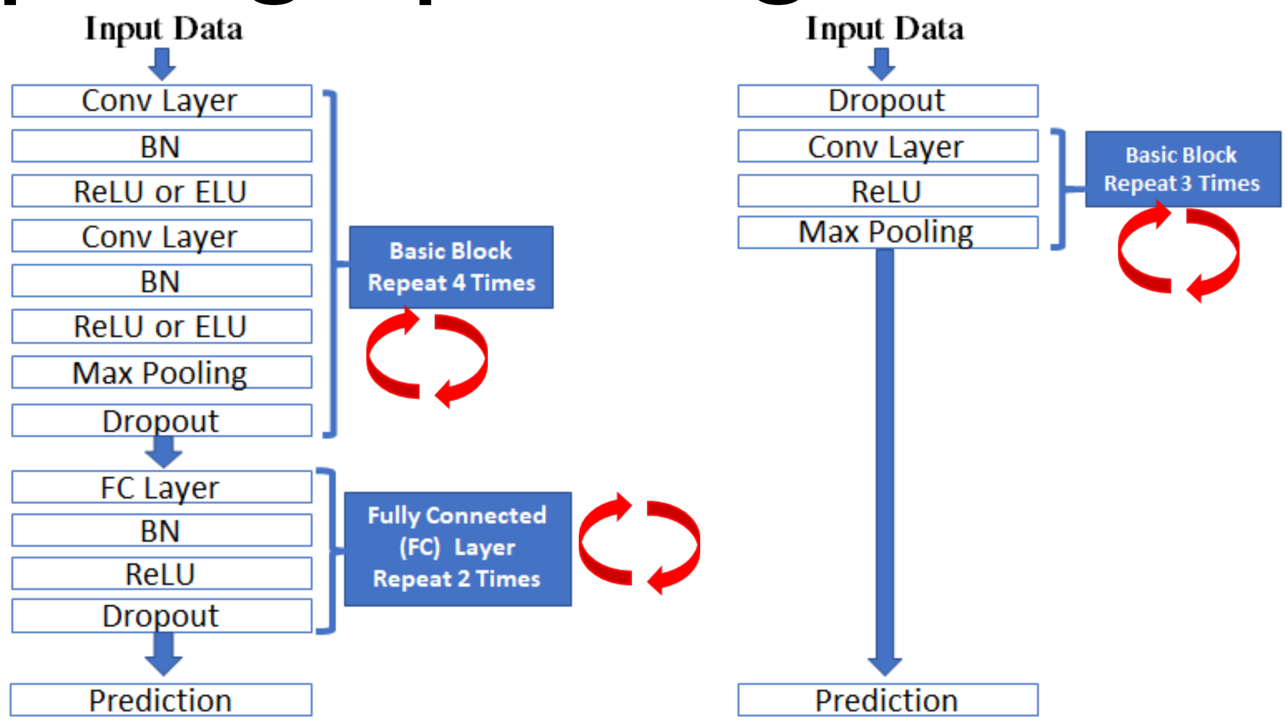


Deep Fingerprinting



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

Deep Fingerprinting



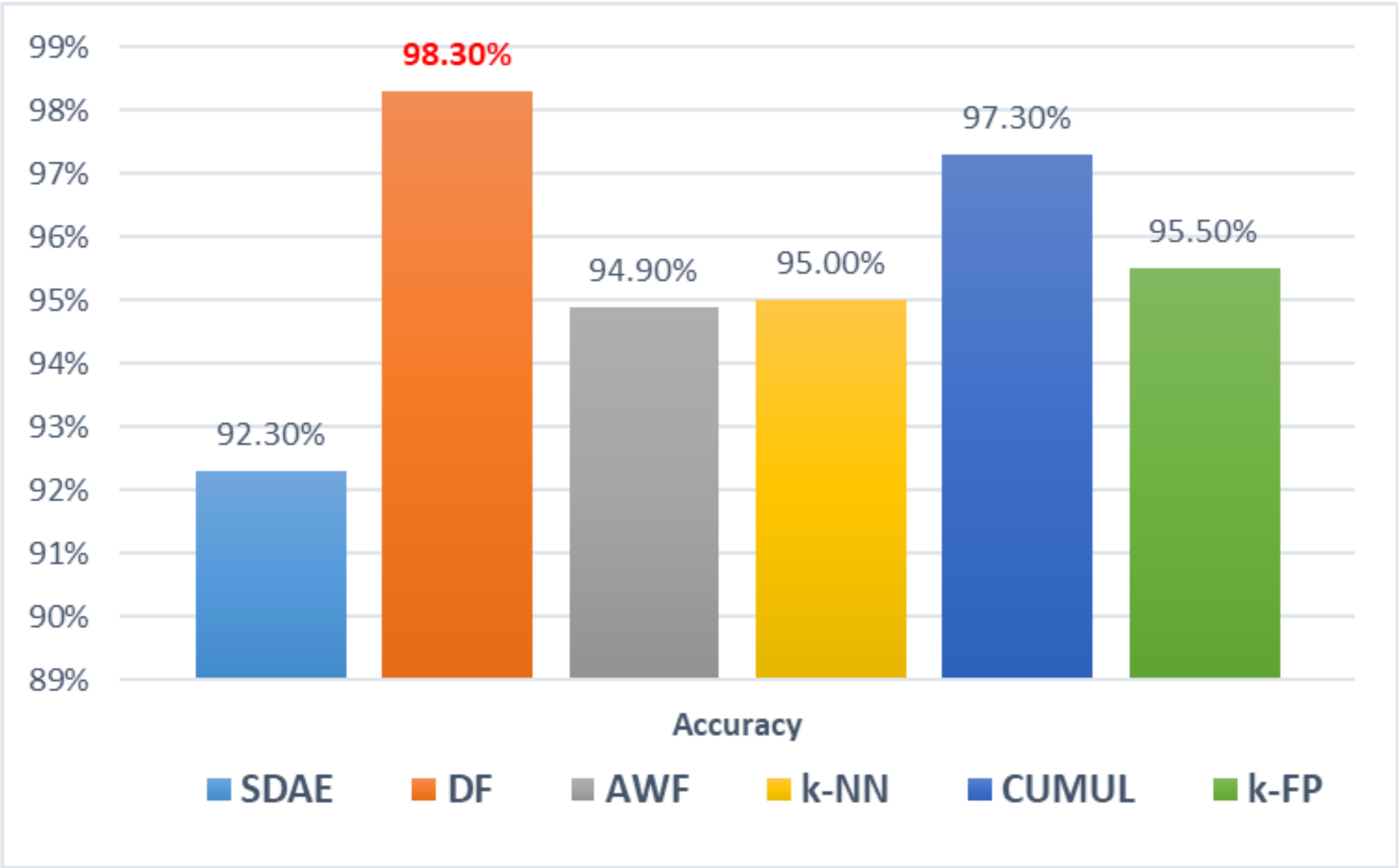
DF Model
(Our)

~3X deeper

AWF Model
(Rimmer et al.)

RIT

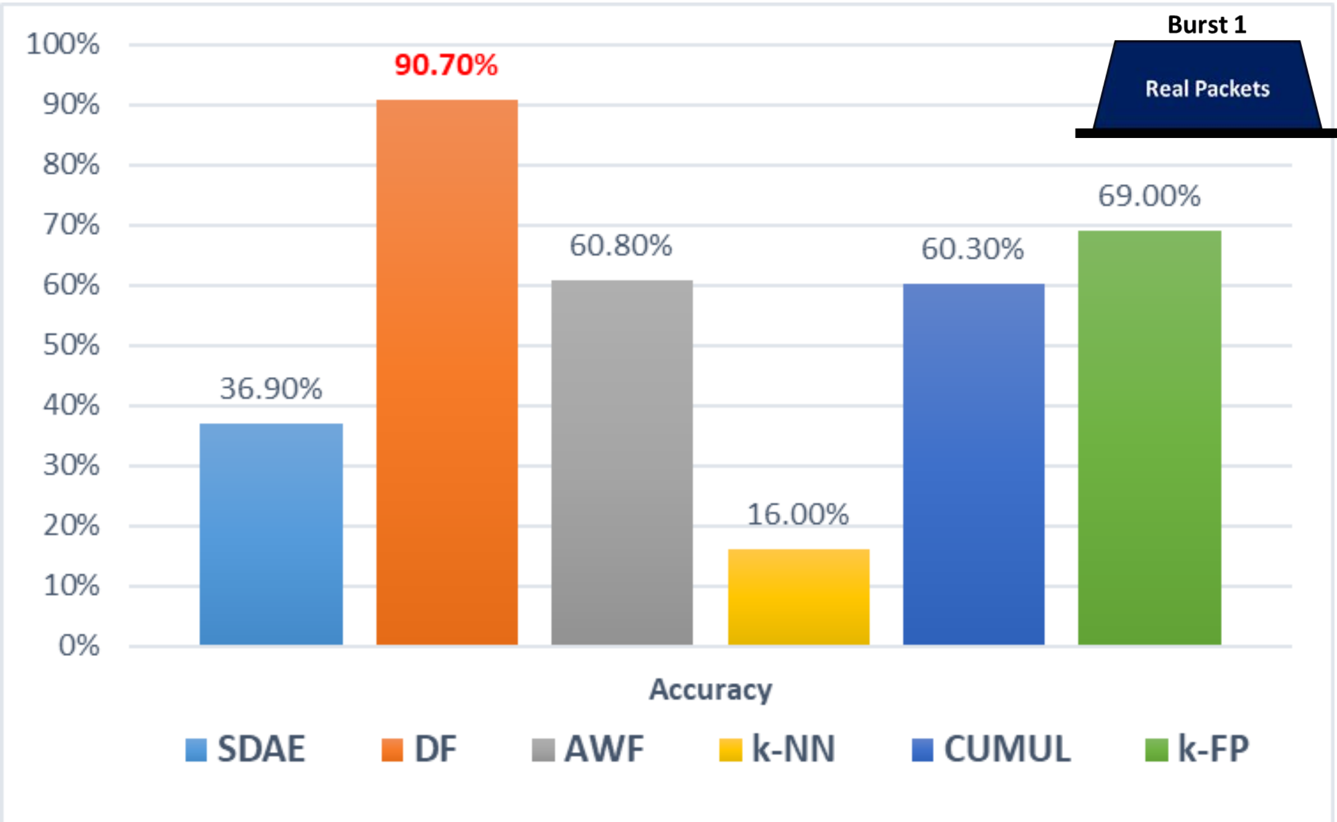
Evaluation: No Defense

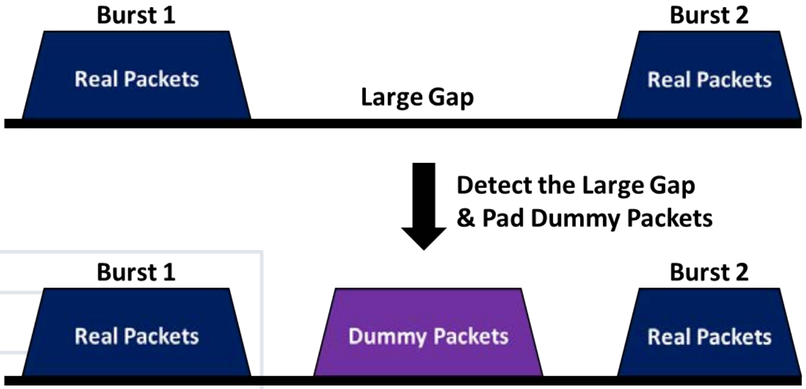


RIT

WTF-PAD

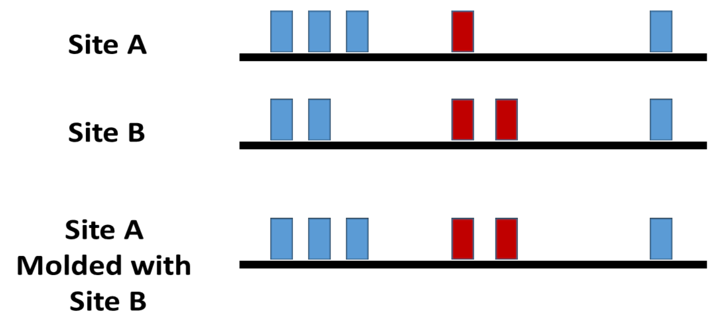
- 64% Bandwidth

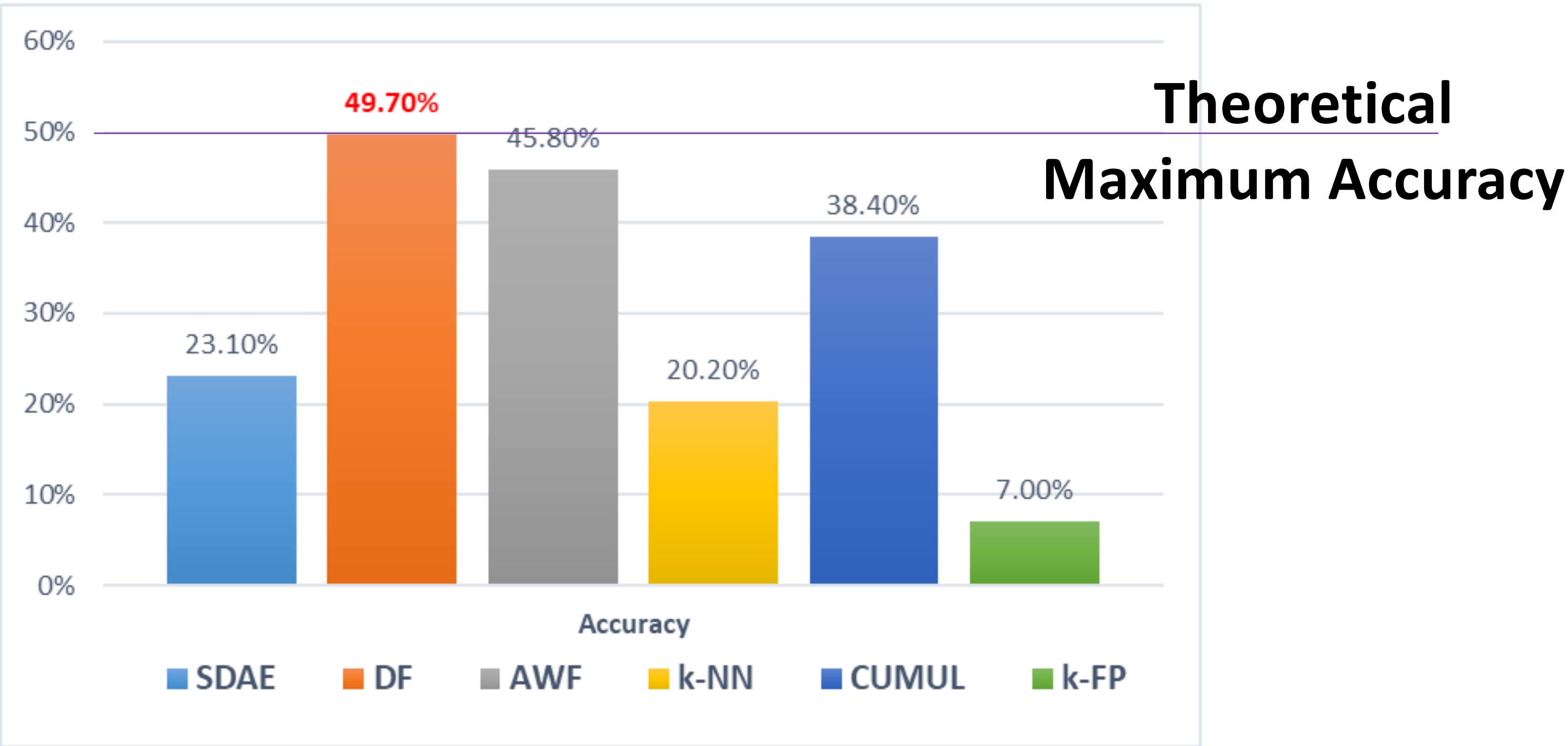




Walkie-Talkie

- 31% Bandwidth, 34% Latency



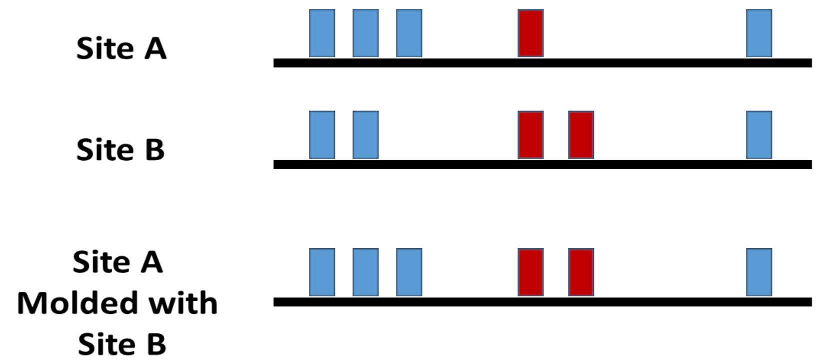


RIT

Walkie-Talkie: Discussion

- Top-N prediction

Top-2 prediction: 98.44 Accuracy



- Implementation Challenges

Conclusion

Original

DF Model

Very Effective

Distorted

DF Model

90% accuracy
or
98% Top-2


I’m back, baby!

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.

Questions?

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

Deep Fingerprinting

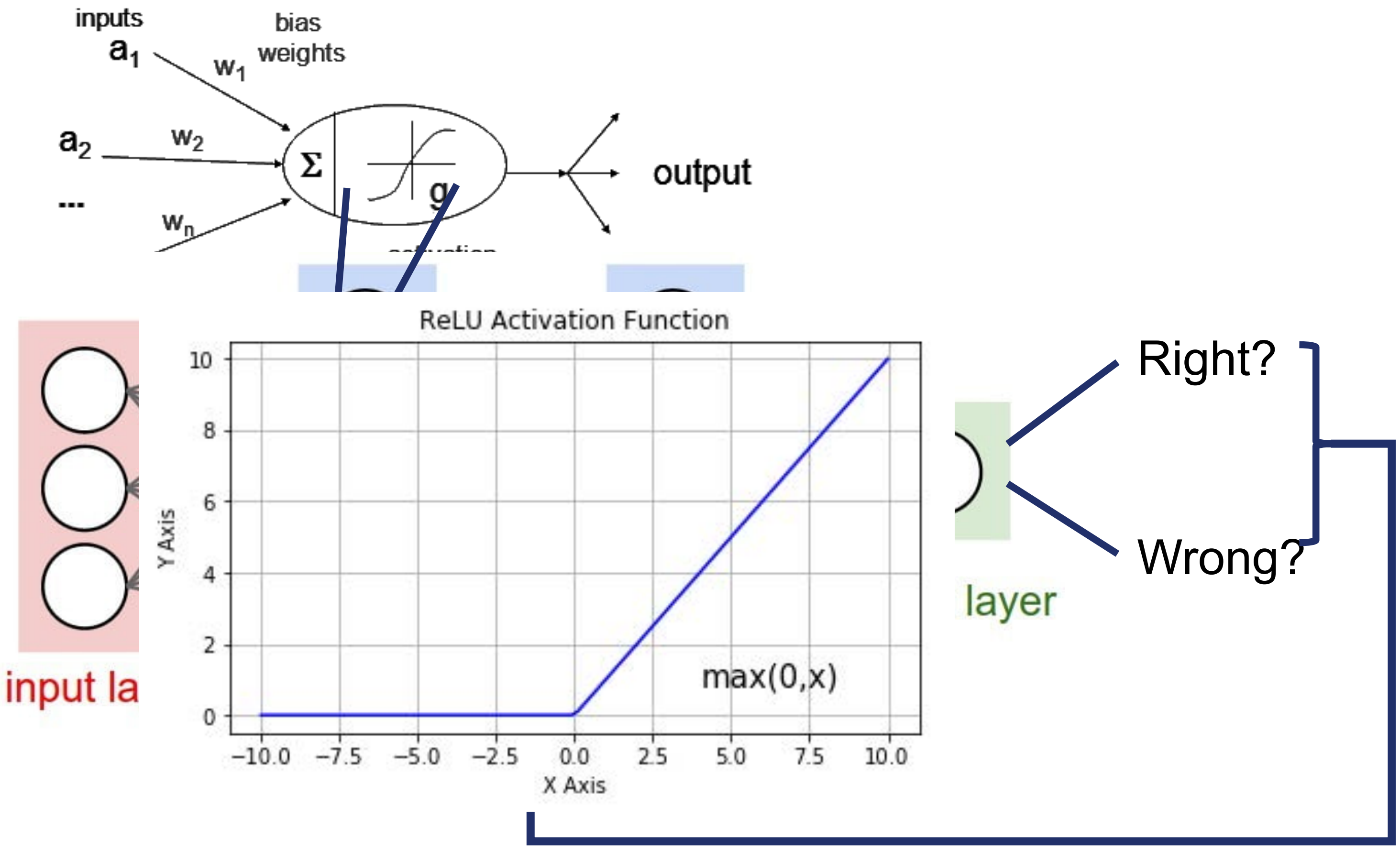
Undermining Website Fingerprinting Defenses with Deep Learning https://github.com/deep-fingerprinting/df

Deep Fingerprinting

Undermining Website Fingerprinting Defenses with Deep Learning

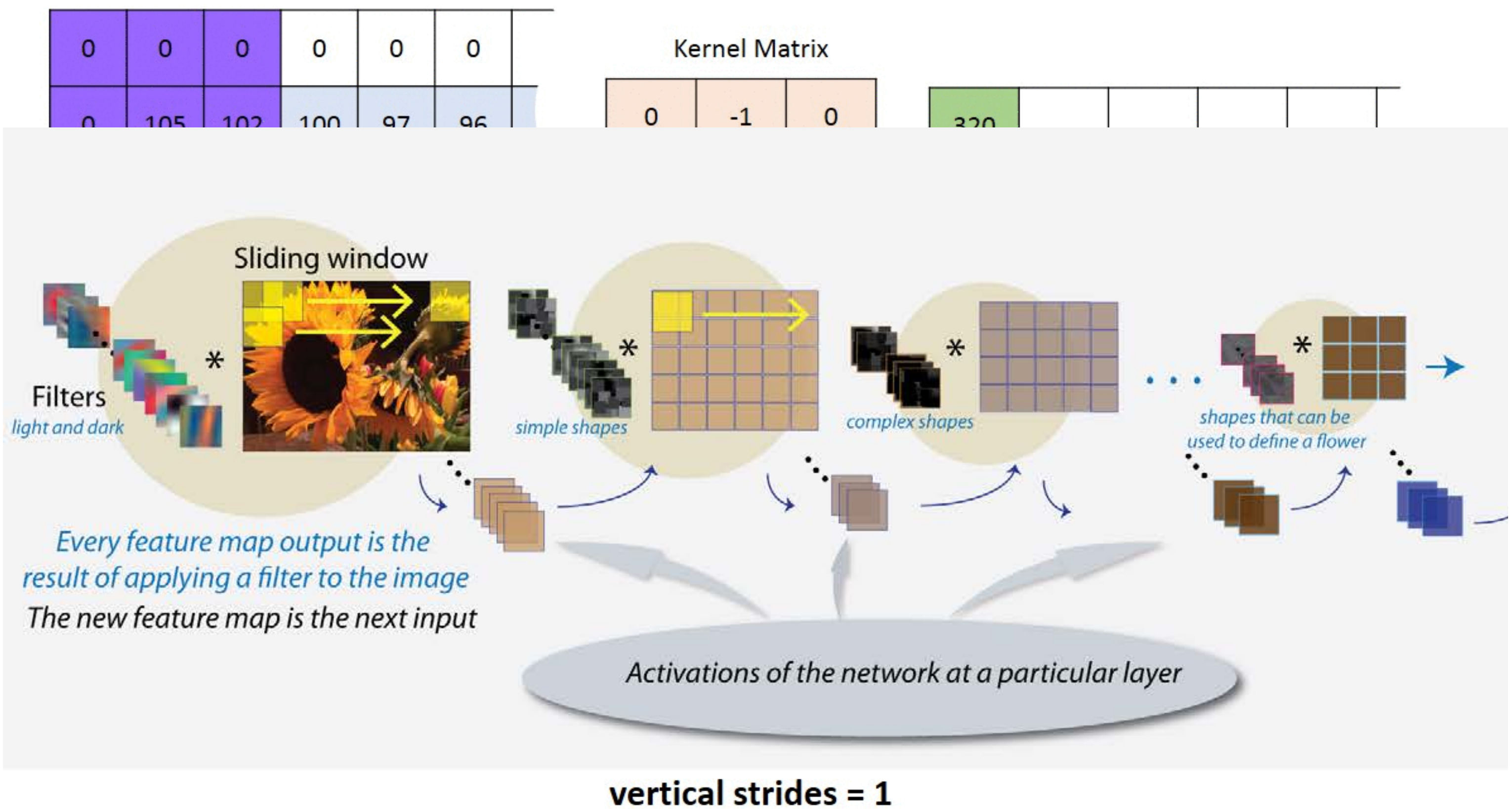
Backup Slides

Neural Networks (in 1 slide)



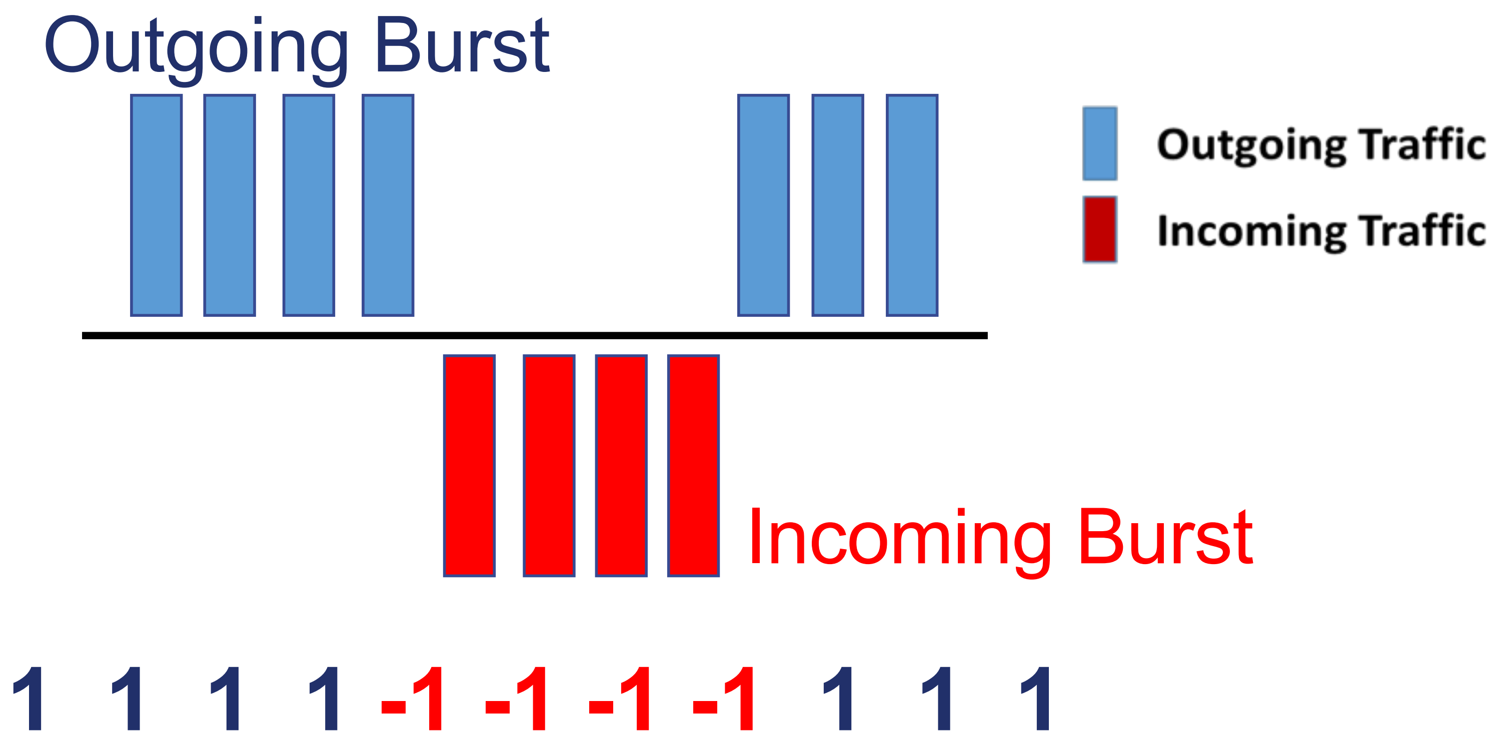
https://stats.stackexchange.com/questions/188277/activation-function-for-first-layer-nodes-in-an-ann
https://www.digitaltrends.com/cool-tech/what-is-an-artificial-neural-network/

CNNs (in 1 slide)

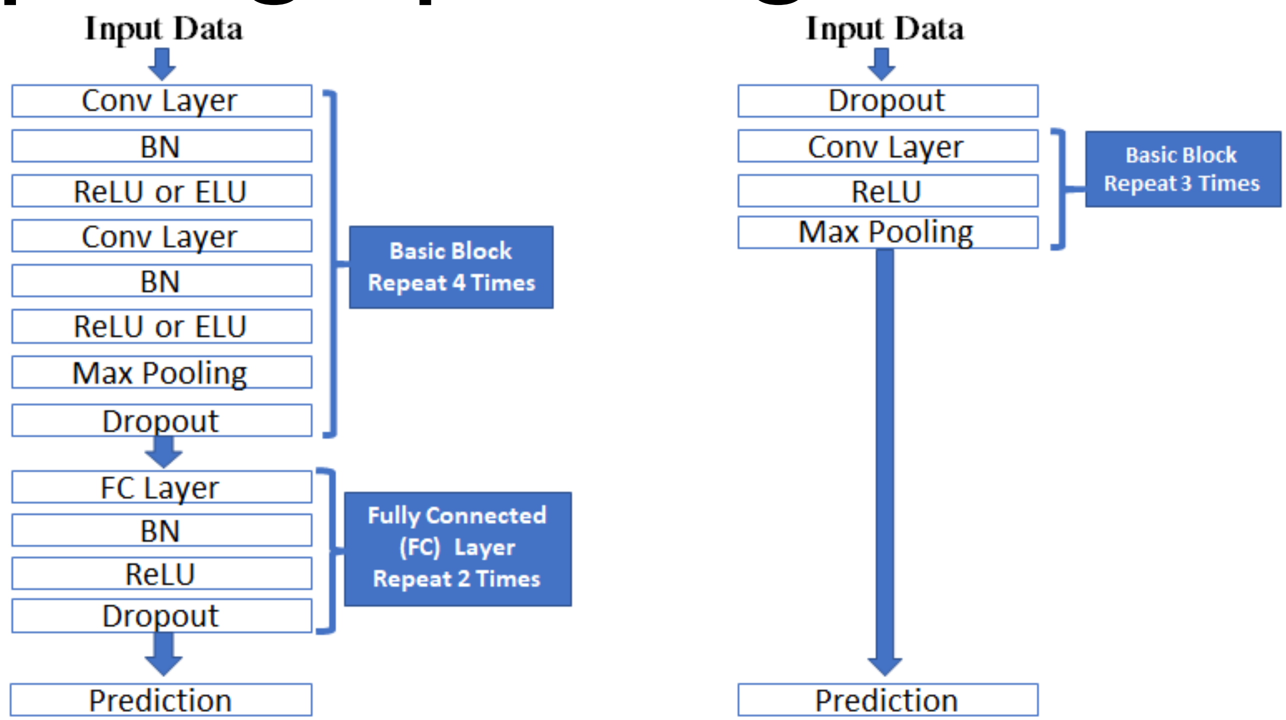


https://stats.stackexchange.com/questions/188277/activation-function-for-first-layer-nodes-in-an-ann
https://www.digitaltrends.com/cool-tech/what-is-an-artificial-neural-network/

Data Representation



Deep Fingerprinting

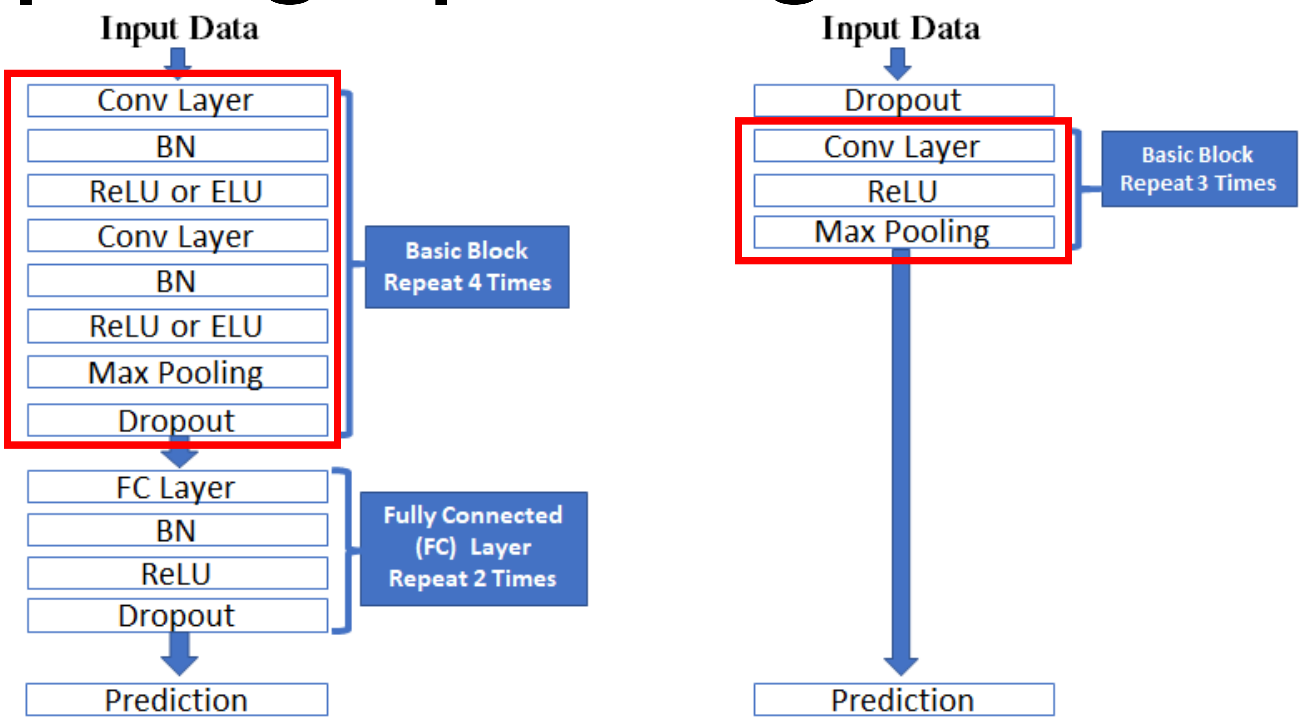


DF Model
(Our)

AWF Model
(Rimmer et al.)

RIT

Deep Fingerprinting



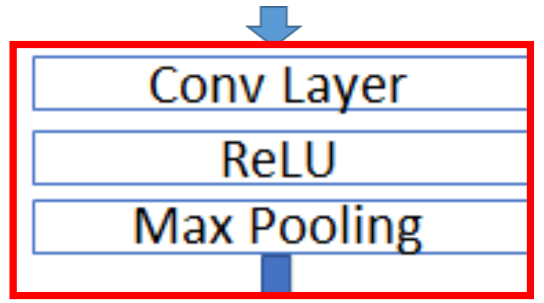
DF Model
(Our)

AWF Model
(Rimmer et al.)

RIT

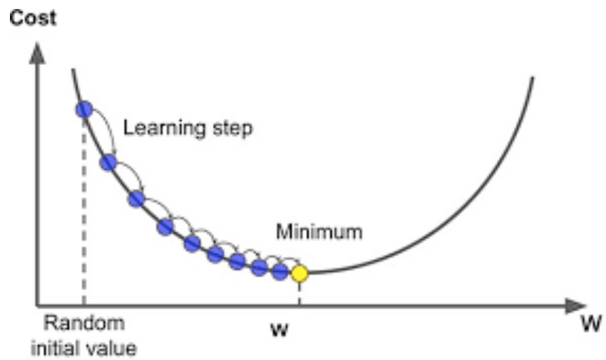
Deep Fingerprinting

DF Model
(Our)

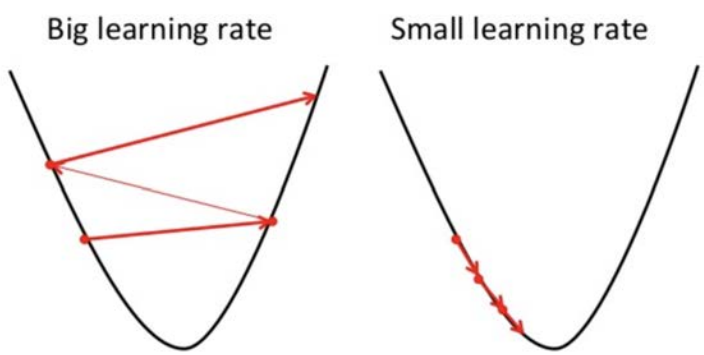


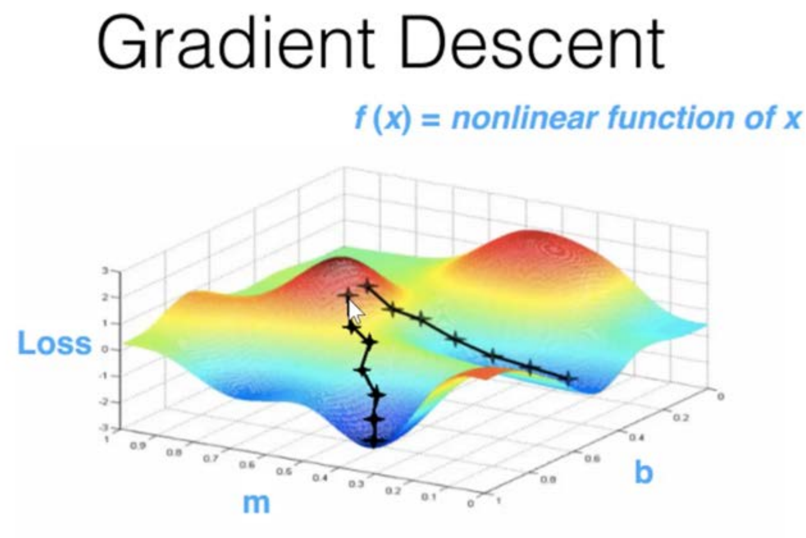
AWF Model
(Rimmer et al.)

Batch Norm



Gradient Descent



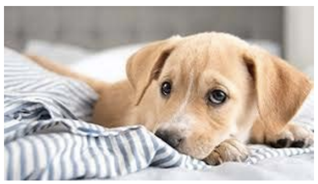


BN: **30 cm max**

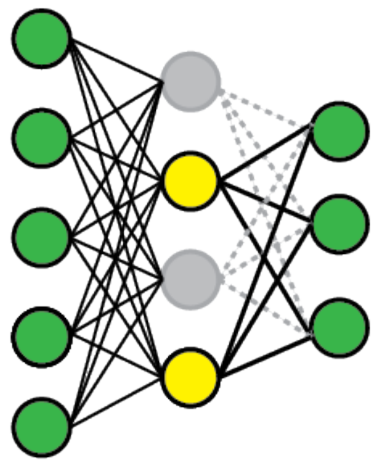
https://saugatbhattarai.com.np/what-is-gradient-descent-in-machine-learning/
https://towardsdatascience.com/gradient-descent-in-a-nutshell-eaf8c18212f0
https://medium.com/@julian.harris/stochastic-gradient-descent-in-plain-english-9e6c10cdba97

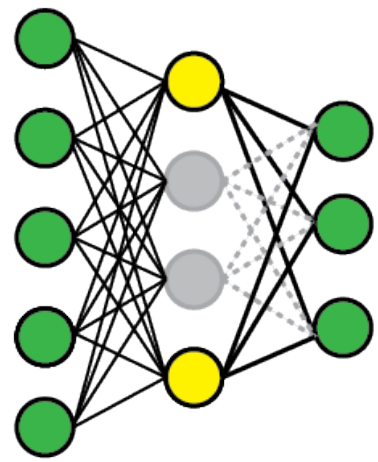
Dropout

Train



Test





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

Closed vs. Open World

Monitored

facebook.com

humanrights.com

.....

Unmonitored

cartoon.com

alibaba.com

.....

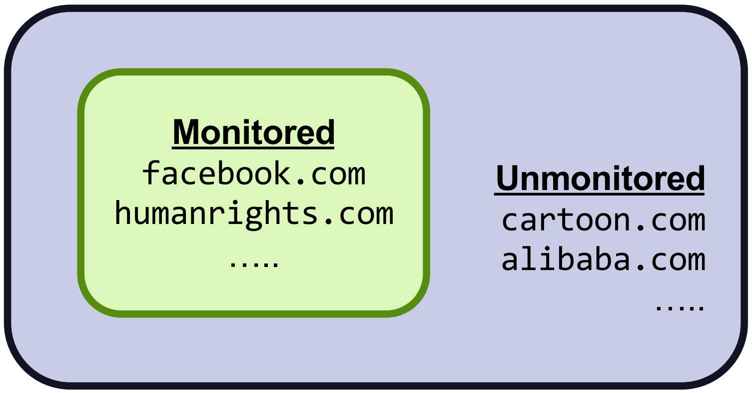
Closed vs. Open World

Monitored
`facebook.com`
`humanrights.com`
.....

Closed-World Scenario

- Users only visit monitored sites
- **Accuracy** of the attack
- Unrealistic

Closed vs. Open World

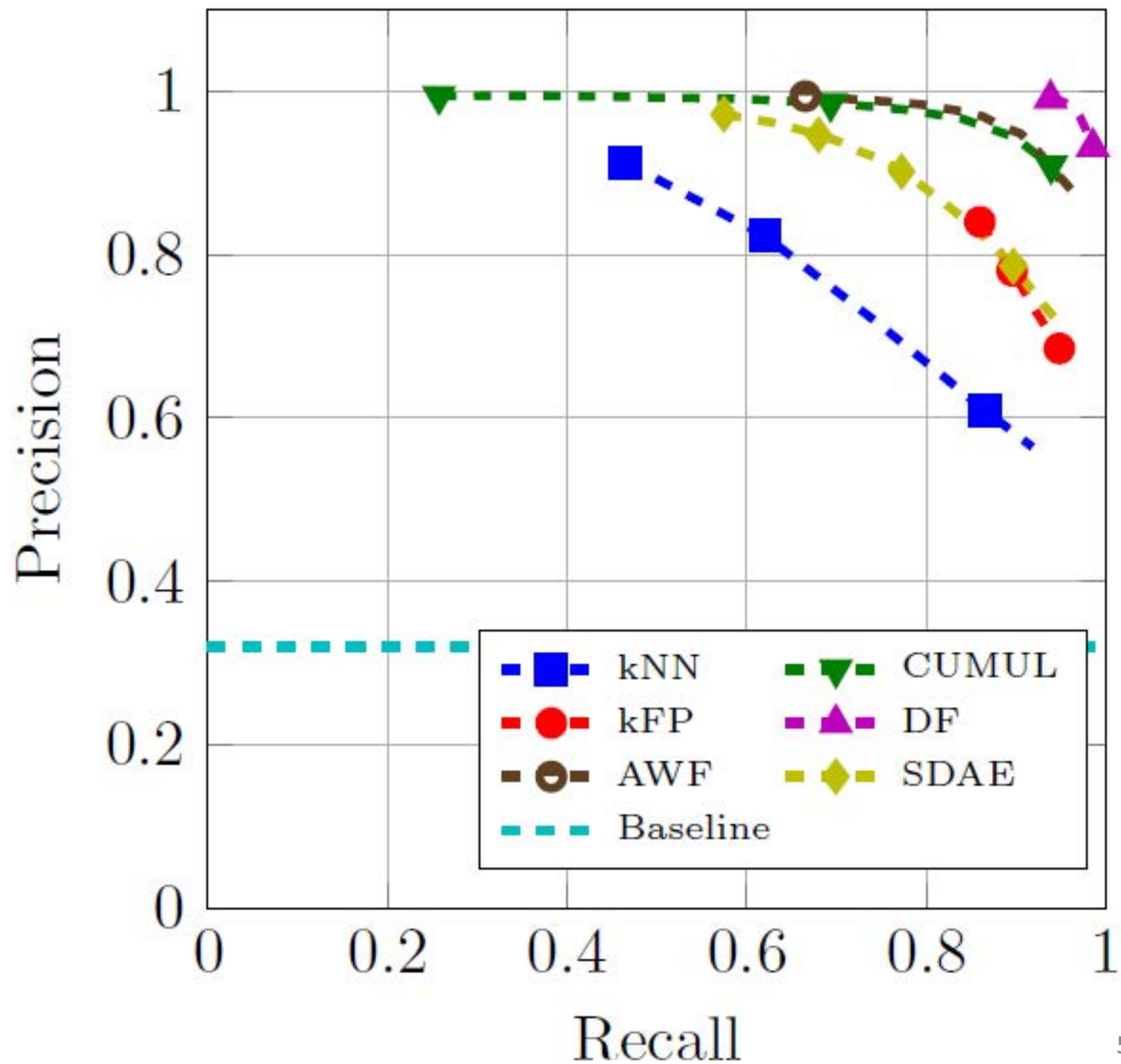


Open-World Scenario

- Users can visit any site
- Attacker goal: ID monitored sites
- **Precision** & **Recall**

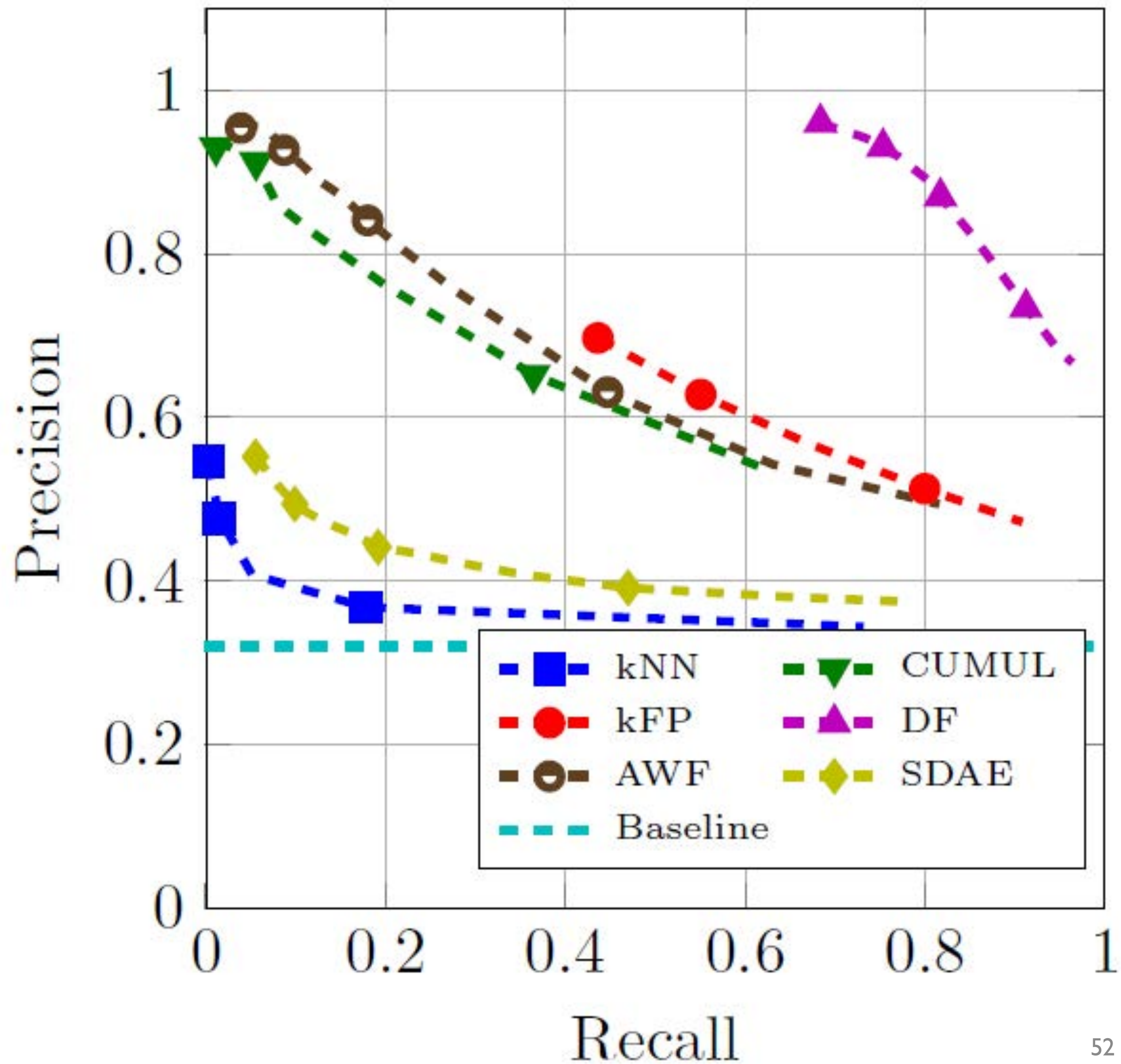
Open World

- 99% precision
- 94% recall



WTF-PAD: Open World

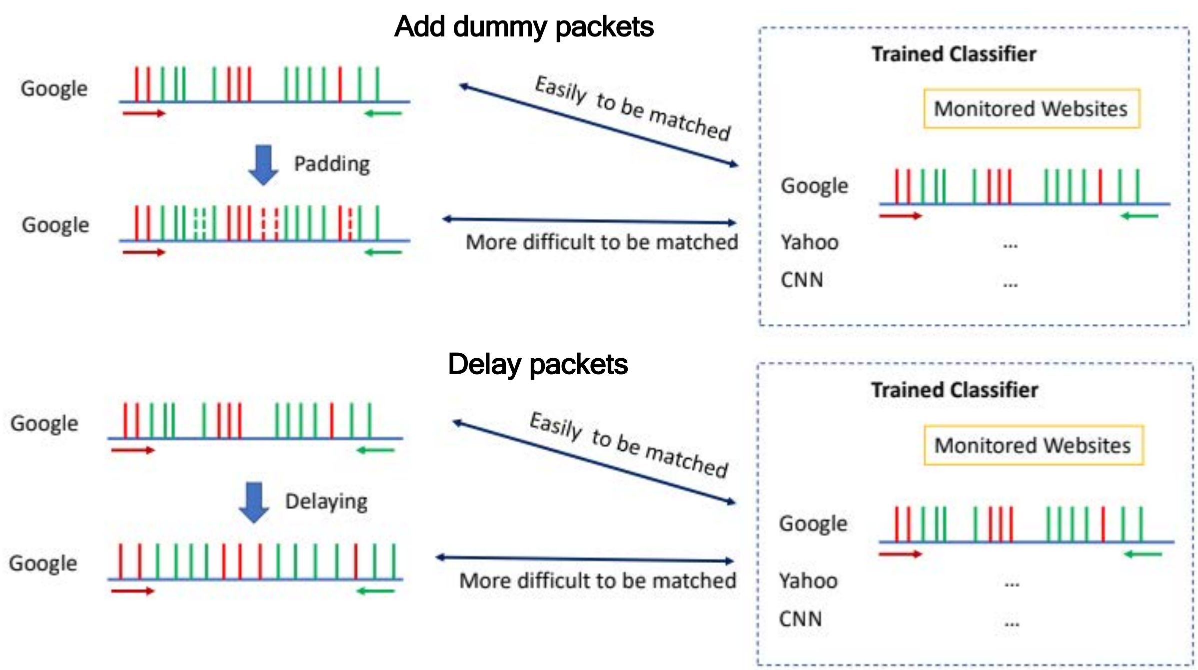
- 96% precision
- 68% recall



Website Fingerprinting Attacks & Defenses

WF Defenses

- Basic mechanisms



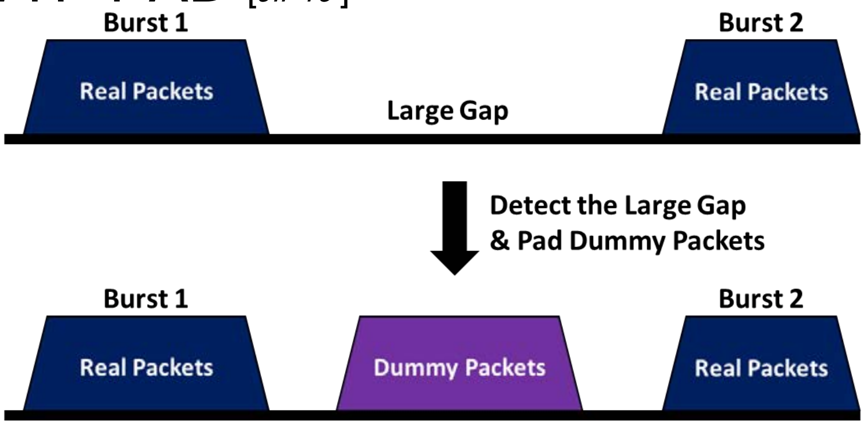
Transition to Practice

- Working with Tor to deploy this

Website Fingerprinting Attacks & Defenses

Lightweight WF Defenses

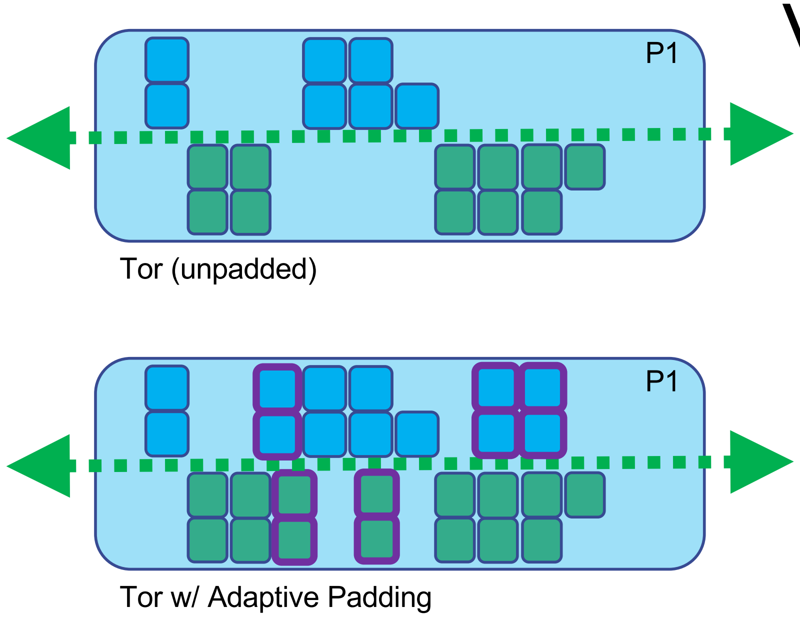
- WTF-PAD [*JIP16*]



- 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.

Adaptive Padding

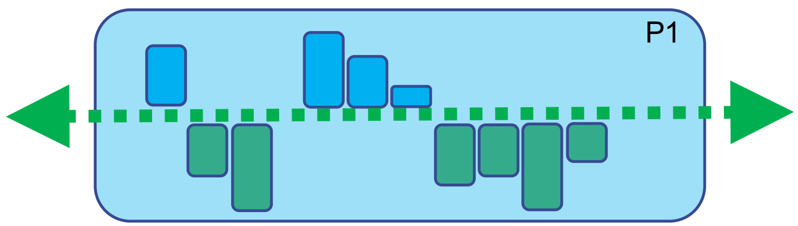


WTF-PAD

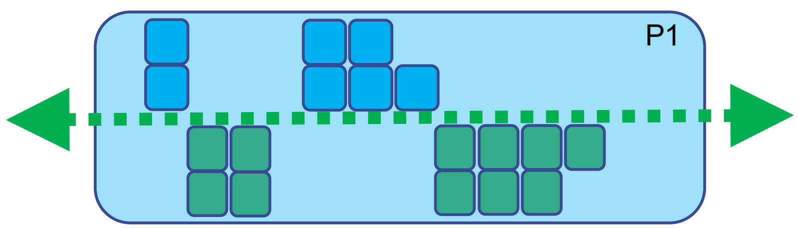
- AP for Tor
- 90% accuracy → **17%**
- 54-64% bandwidth overhead
- Minimal added delay

RIT

Tor Cells



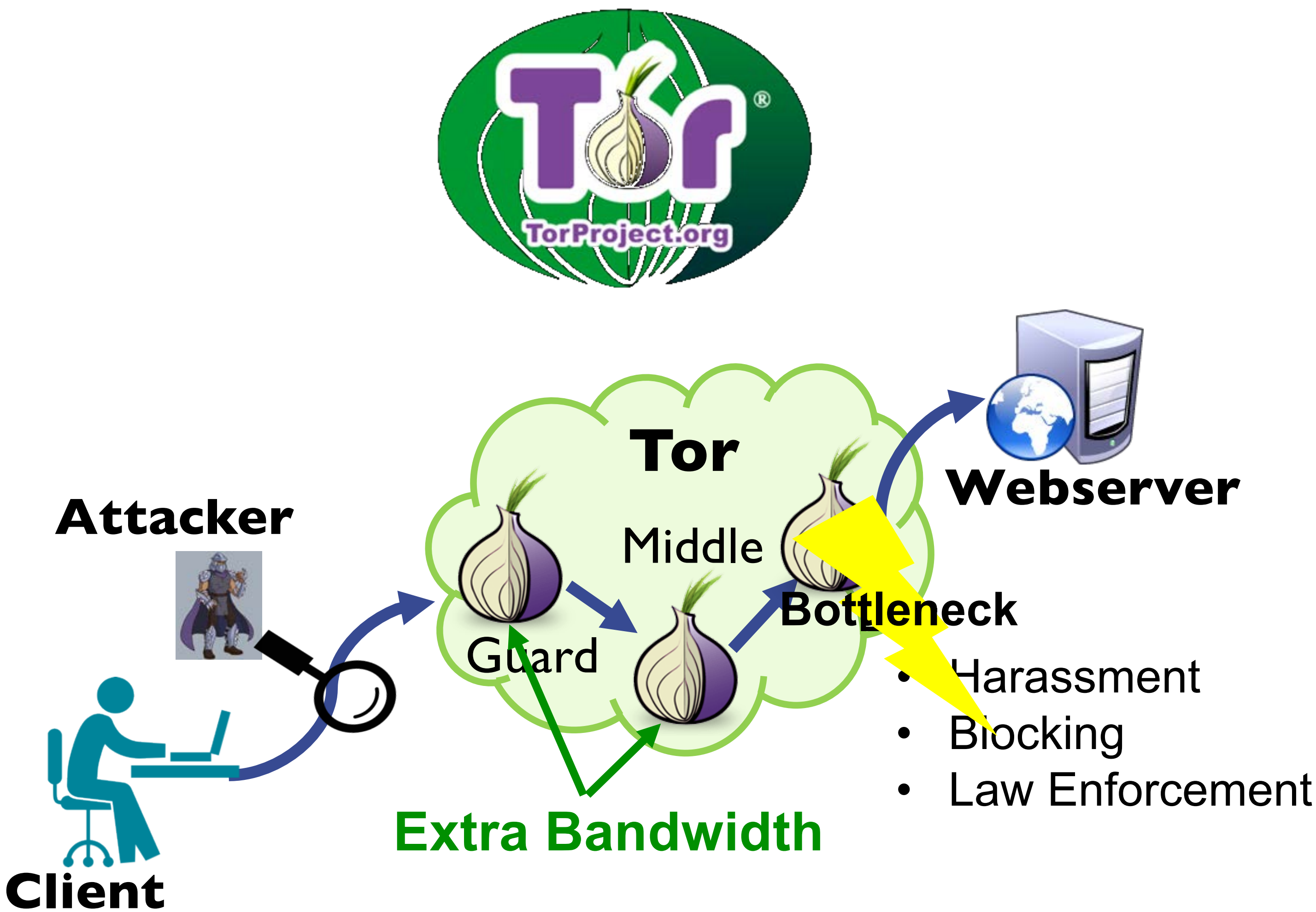
Without Tor



With Tor (512 byte cells)

90%+ Accuracy*

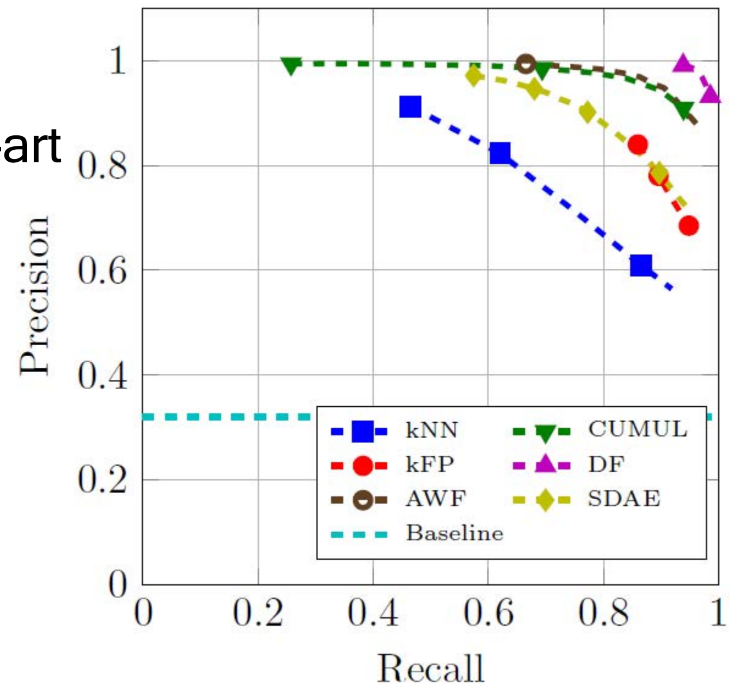
* For ~100 sites, not pages


TorProject.org
Tor
Attacker
Middle
Webserver
Bottleneck
Guard
Harassment
Blocking
Law Enforcement
Extra Bandwidth
Client

Deep Fingerprinting

Experimental Evaluation (Open World)

- Non-Defended
  - DF outperforms other state-of-the-art

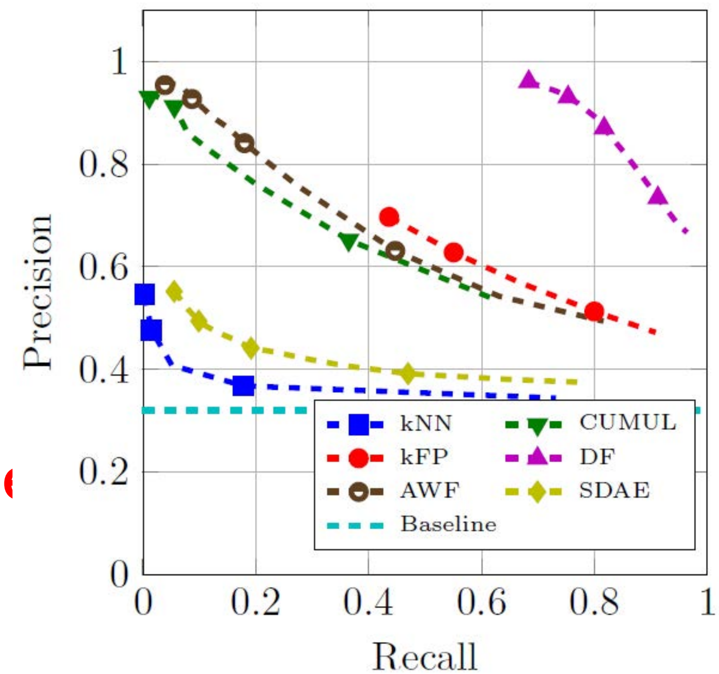


RIT

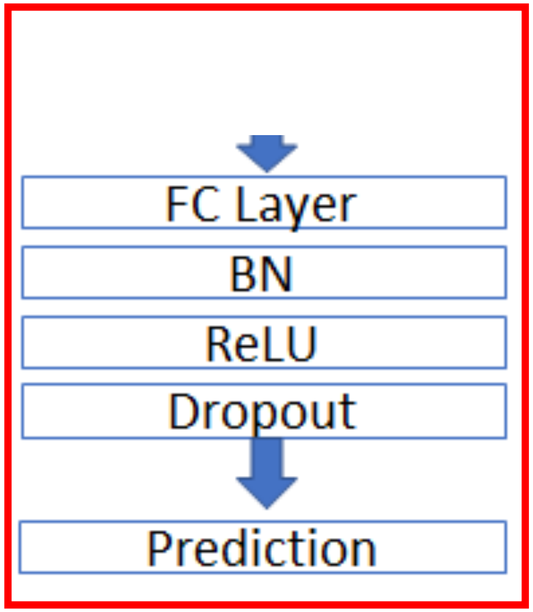
Deep Fingerprinting

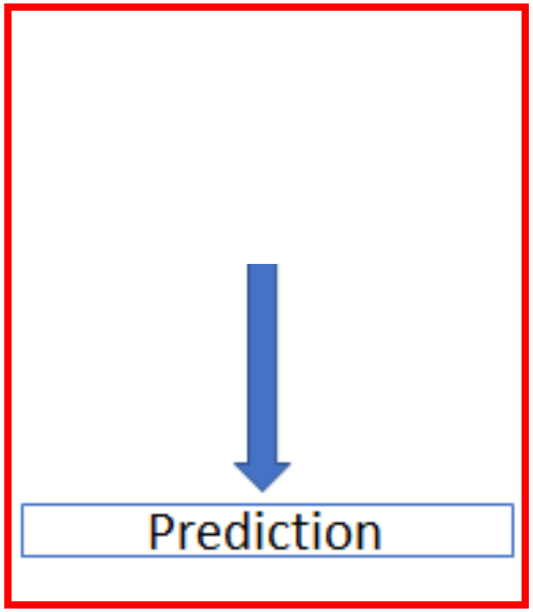
Experimental Evaluation (Open World)

- WTF-PAD
  - DF perform the best
  - DF significantly outperforms other state-of-the-art
- The DF can undermine WTF-PAD



Deep Fingerprinting



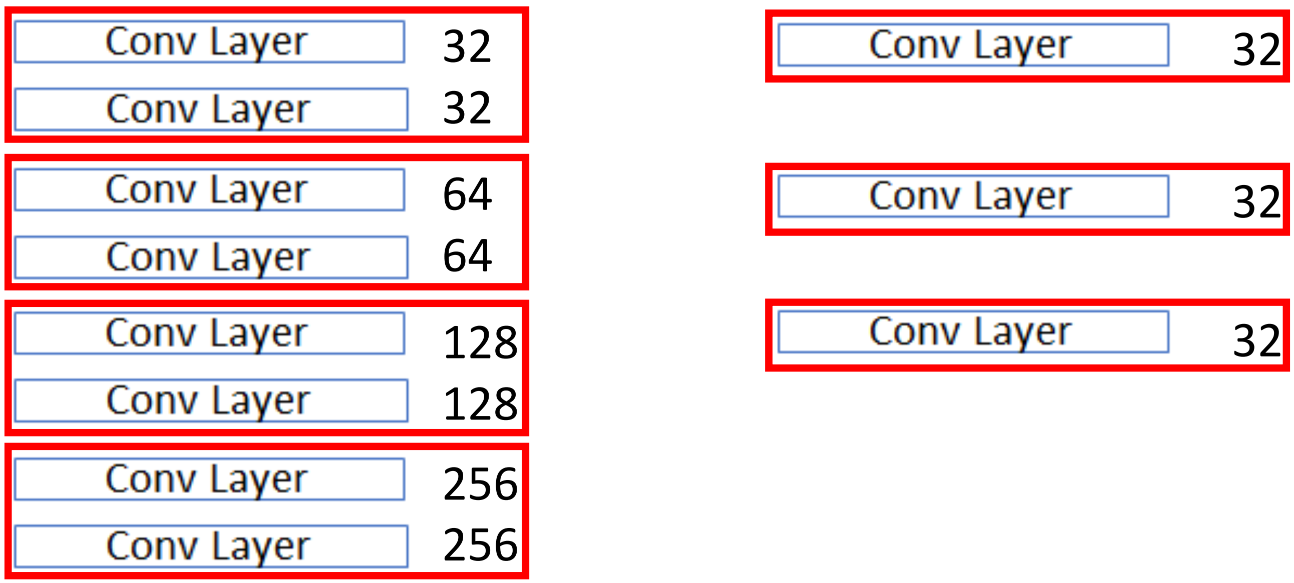


DF Model
(Our)

AWF Model
(Rimmer et al.)

RIT

Deep Fingerprinting



DF Model
(Our)

AWF Model
(Rimmer et al.)

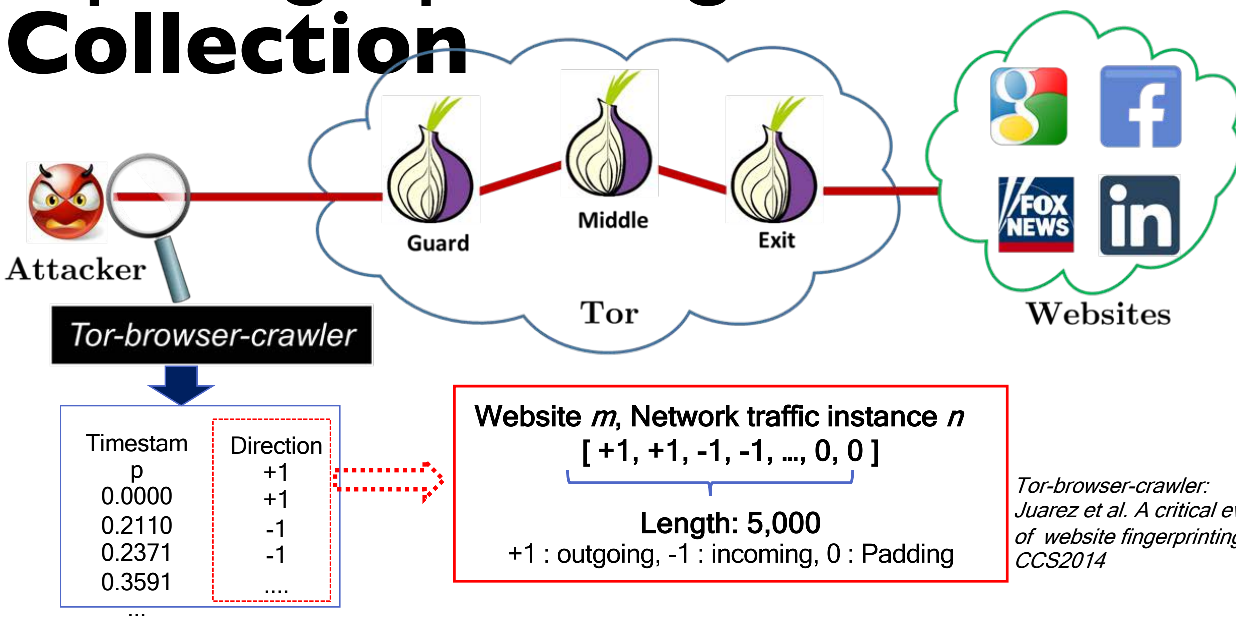
Deep Fingerprinting

Walkie-Talkie: Discussion

- **Deployablity**
  - Requires database
    - Distribute to the clients and Tor’s nodes
  - Only apply to static website
  - Half-duplex communication
    - 31 % additional latency
    - Direct cost to end-user performance
    - Tor is now slower than regular browsing

Deep Fingerprinting

- **Data Collection**

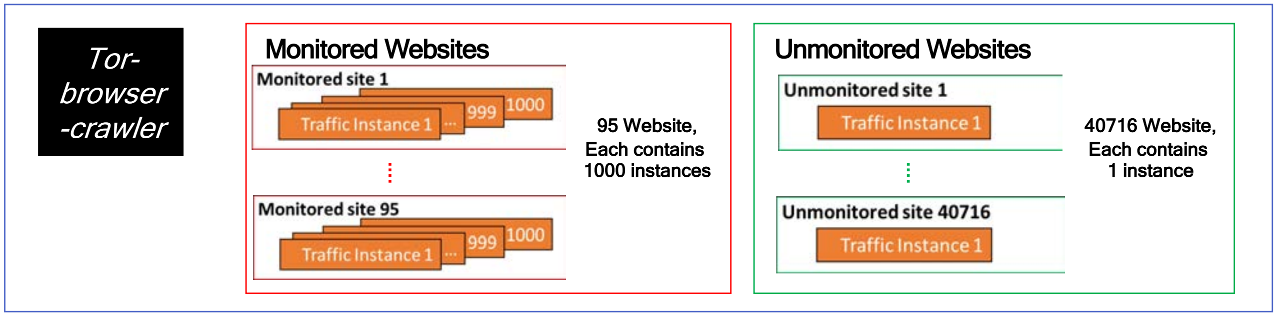


Tor-browser-crawler:
Juarez et al. A critical evaluation of website fingerprinting attacks., CCS2014

RIT

Deep Fingerprinting

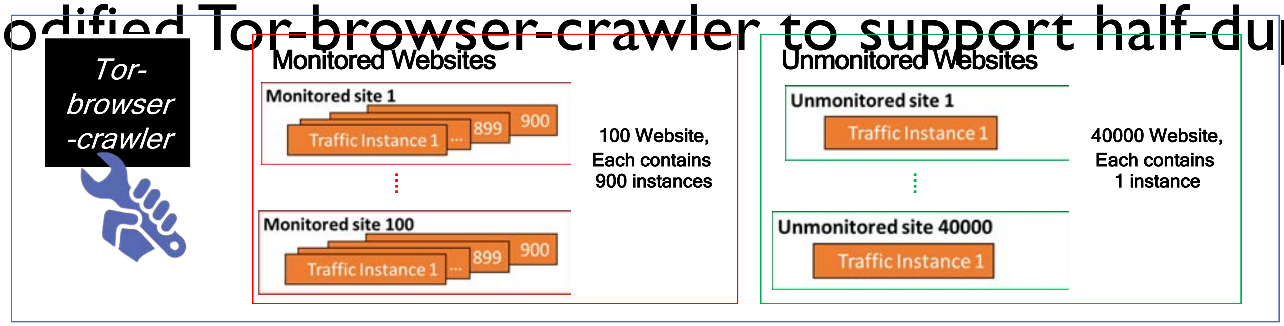
- **Data Collection**
  - Non-Defended Dataset



  - WTF-PAD Dataset
    - Simulated from non-defended dataset (same size)

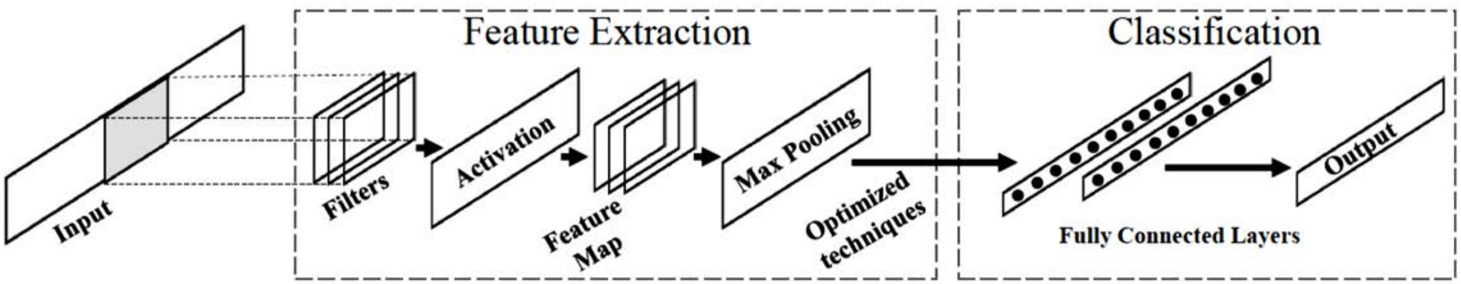
Deep Fingerprinting

- **Data Collection (Cont.)**
  - **Walkie-Talkie**
    - Modified Tor-browser-crawler to support half-duplex



Deep Fingerprinting

- **DF Model**
  - Based on the CNN architecture



Deep Fingerprinting

Failure Causes of WTF-PAD

- **Ability to detect the hidden features**
  - WTF-PAD handle WF attacks using hand-crafted features
  - Defense hides the deterministic features
- **Robustness against small change**
  - WTF-PAD aim to fill the gap with the faked burst
  - Insufficient distortion and still leave fingerprint

Background & Related Work

- **WF Attacks using Hand-crafted Features (Cont.)**
  - k-NN [Wang et al.]
    - Packets ordering, #incoming & outgoing, #bursts etc.
    - k-Nearest Neighbors
    - 91% Accuracy (closed world)
    - 86% TPR and 0.6% FPR (open world)

- *Wang et al. Effective attacks and provable defenses for website fingerprinting. , USENIX 2014*

Background & Related Work

- **WF Attacks using Hand-crafted Features (Cont.)**
  - CUMUL [Panchenko et al.]
    - Cumulative sum of packet lengths.
    - SVM
    - 91% Accuracy (closed world)
    - 96% TPR and 1.9% FPR (open world)

- Panchenko et al. "Website fingerprinting at internet scale.", NDSS 2016

RIT

Background & Related Work

- **WF Attacks using Hand-crafted Features (Cont.)**
  - k-FP [Hayes and Danezis]
    - Traditional features such as #packets
    - Random Forest to extract the features
    - Analyze the importance of the features
    - 91% Accuracy (closed world)
    - 88% TPR and 0.5% FPR (open world)

Hayes and Danezis. k-Fingerprinting: A robust scalable website fingerprinting technique. , USENIX 2016.

RIT

- **ML Techniques Used in the DF**

Deeper Networks

- Krizhevsky et al. Imagenet classification with deep convolutional neural networks., NIPS 2012.
- Szegedy et al. Going deeper with convolutions. CVPR 2015.
- Karen and Andrew. Very deep convolutional networks for large-scale image recognition. ArXiv2015.

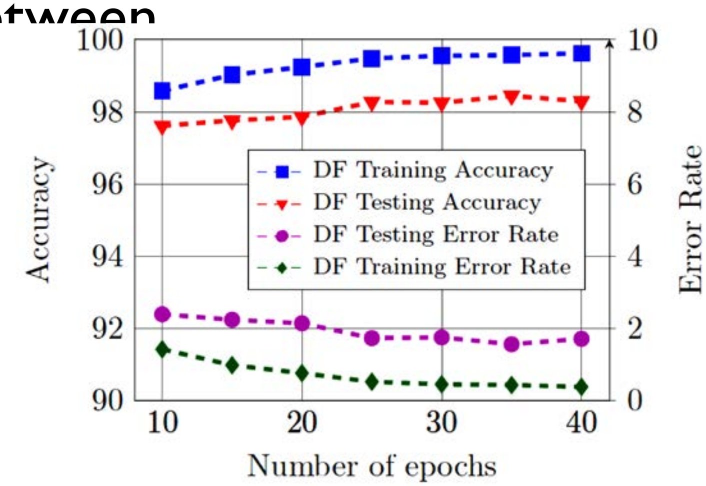
Appropriate Activation Functions

- Clevert et al. Fast and accurate deep networks learning by exponential linear units (elus). ICCV2015.
- Mishkin et al. Systematic evaluation of CNN advances on the imagenet. CoRR, abs/1606.02228, 2016.

Prevent Overfitting

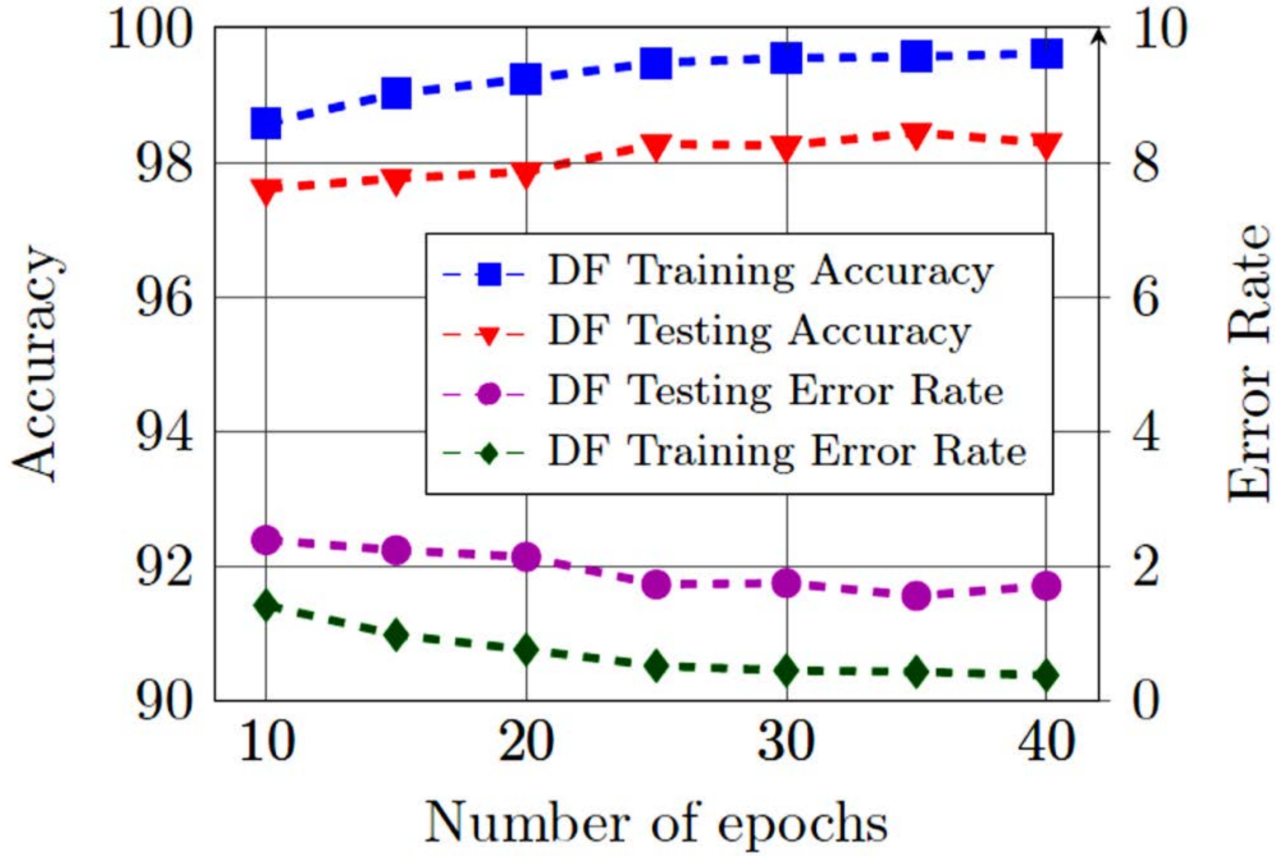
- Srivastava et al. Dropout: A simple way to prevent neural networks from overfitting. Journal of Machine Learning Research 2014
- Ioffe and Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift., International Conference on Machine Learning, 2015

- **Experimental Evaluation**
  - **Convergence of the DF model**
    - 97 % Accuracy (10 epoch)
    - Level off after 30 epochs
  - **Overfitting measurements**
    - Small difference between training and testing rates (< 2%)
    - Overfitting is unlik

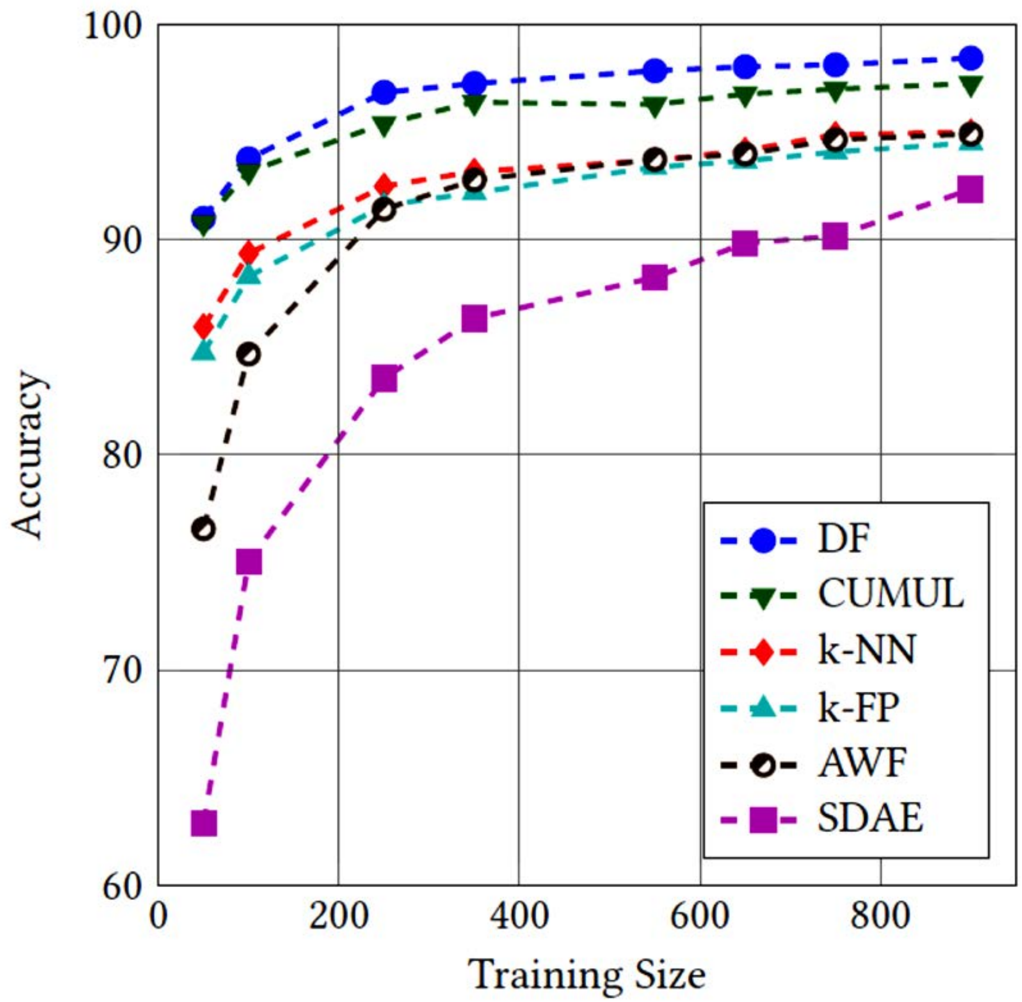


Closed World: Impact of the number of training epochs on the DF model's accuracy and error rate

RIT



100
98
96
94
92
90
10
8
6
4
2
0
10
20
30
40
Accuracy
Error Rate
Number of epochs
DF Training Accuracy
DF Testing Accuracy
DF Testing Error Rate
DF Training Error Rate



100
90
80
70
60
0
200
400
600
800
Accuracy
Training Size
DF
CUMUL
k-NN
k-FP
AWF
SDAE

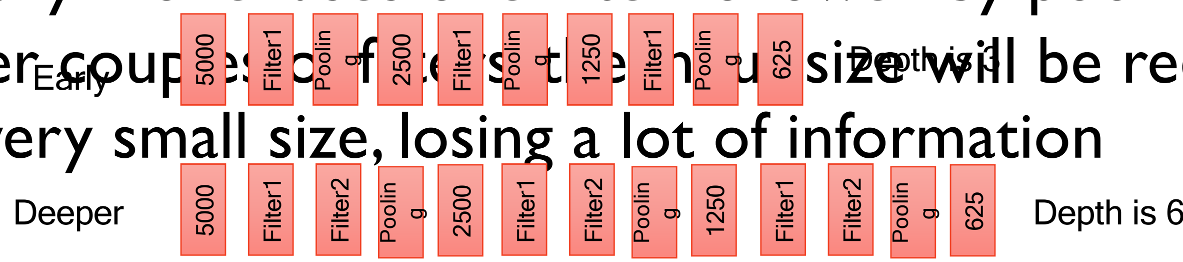
Deeper Model

- How to go deeper
  - Note that, we don't need the extremely deep network like Inception
    - We tested with Inception, Xception, GoogleNet, there is no noticeable improvement for the accuracy of the attack
  - The model just needs to be deep enough to provide the effective performance
  - Deeper network does not always provide the better result

RIT

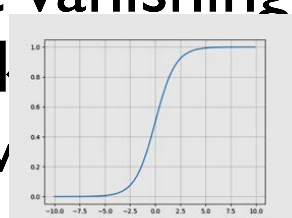
Deeper Model

- How to go deeper
  - Multiple filters before pooling
    - Pooling always reduce the size of the input
    - The early model uses one filter followed by pooling
      - After couples of filters the input size will be reduced to very small size, losing a lot of information



Deeper Model

- Batch Normalization
  - Normalize the inputs to layers with in the network
    - Mean activation close to 0, activation S.D. close to 1
  - Batch normalization helps reduce the sensitivity to the initial starting weights
  - Prevent vanishing gradient problem when the network eper



    - Even v he model sometimes stops learning

Performance Metric

- **Accuracy**

$$\text{Accuracy} = \frac{P_{correct}}{N}$$

$P_{correct}$ is the total number of correct predictions. A correct prediction is defined as the output of the classier matching the label of the website to which the test trace belongs. N is the total number of instances in the test set.

Performance Metric

- **Precision & Recall**

$$Precision = \frac{TP}{TP + FP} \qquad Recall = \frac{TP}{TP + FN}$$

TP is the total number of test samples of monitored websites that are correctly classified as monitored websites.
TN is the total number of test samples of unmonitored websites that are correctly classified as unmonitored websites.
FP is the total number of test samples of unmonitored websites that are misclassified as monitored websites.
FN is the total number of monitored websites that are misclassified as unmonitored websites.

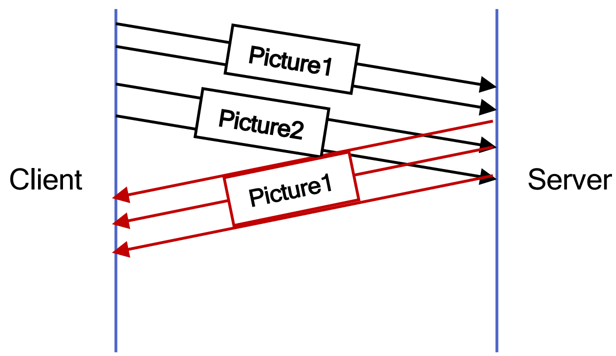
Website Fingerprinting Attacks & Defenses

WF Defenses

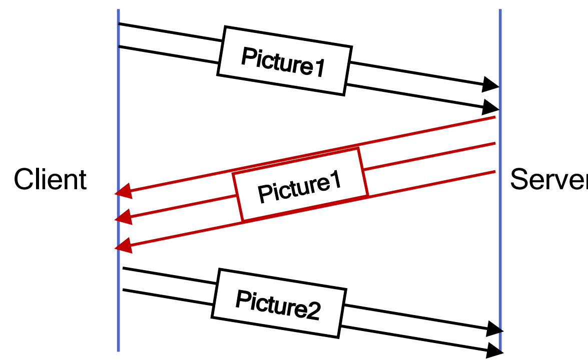
- Basic mechanisms
  - Add and/or delay packets
  - Reduce the distinctive features
- Early WF Defenses
  - BuFLO [*DCR12*] and Tamaraw [*CNJ14*]
  - Make traffic look constant rate
  - 200 – 400% extra latency → 2-4X as long to get the website
  - Over 130% extra bandwidth

[DCR12] Dyer et al. Peek-a-Boo, I still see you: Why efficient traffic analysis countermeasures fail., IEEE S&P 2012
[CNJ14] Cai et al. A systematic approach to developing and evaluating website fingerprinting defenses., CCS 2014

Half-Duplex Communication

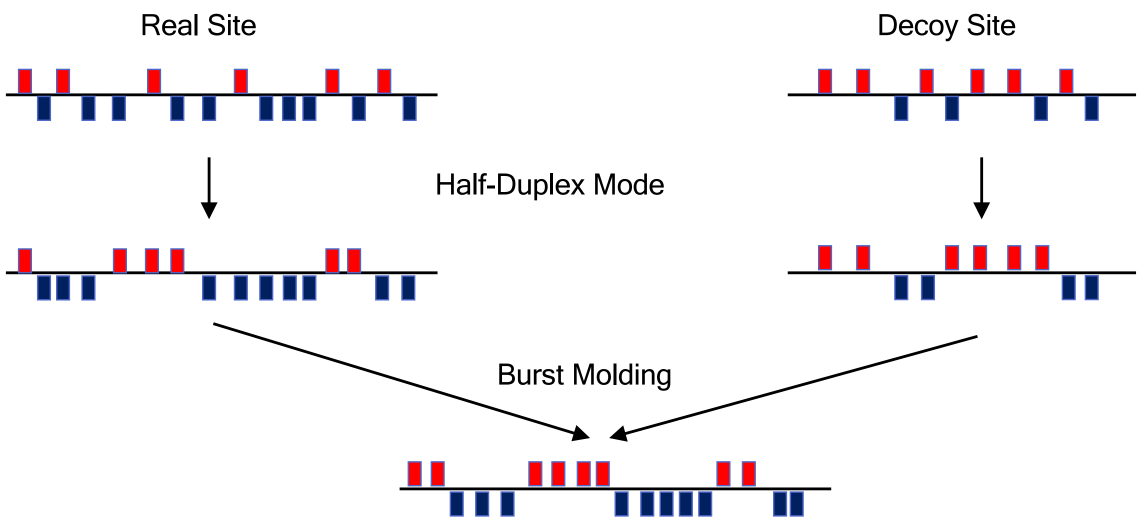


Full-Duplex



Half-Duplex

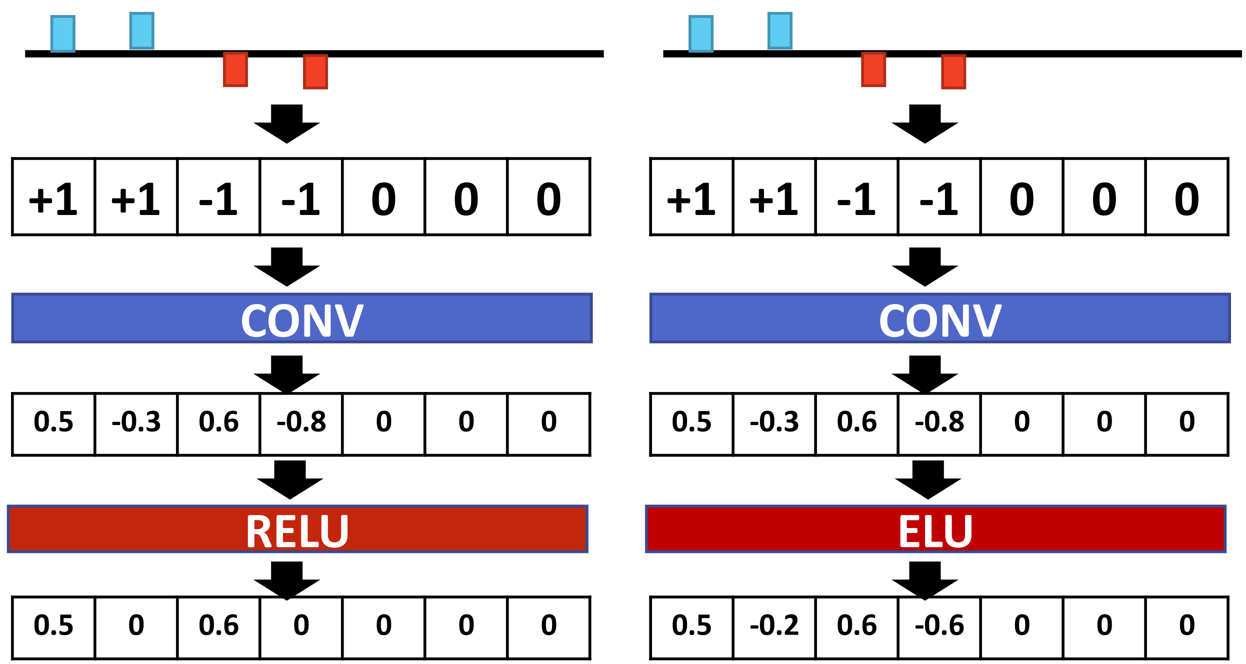
Mold Padding



Deep Fingerprinting

DF Model: Improved Design of CNN

- ELU vs ReLU

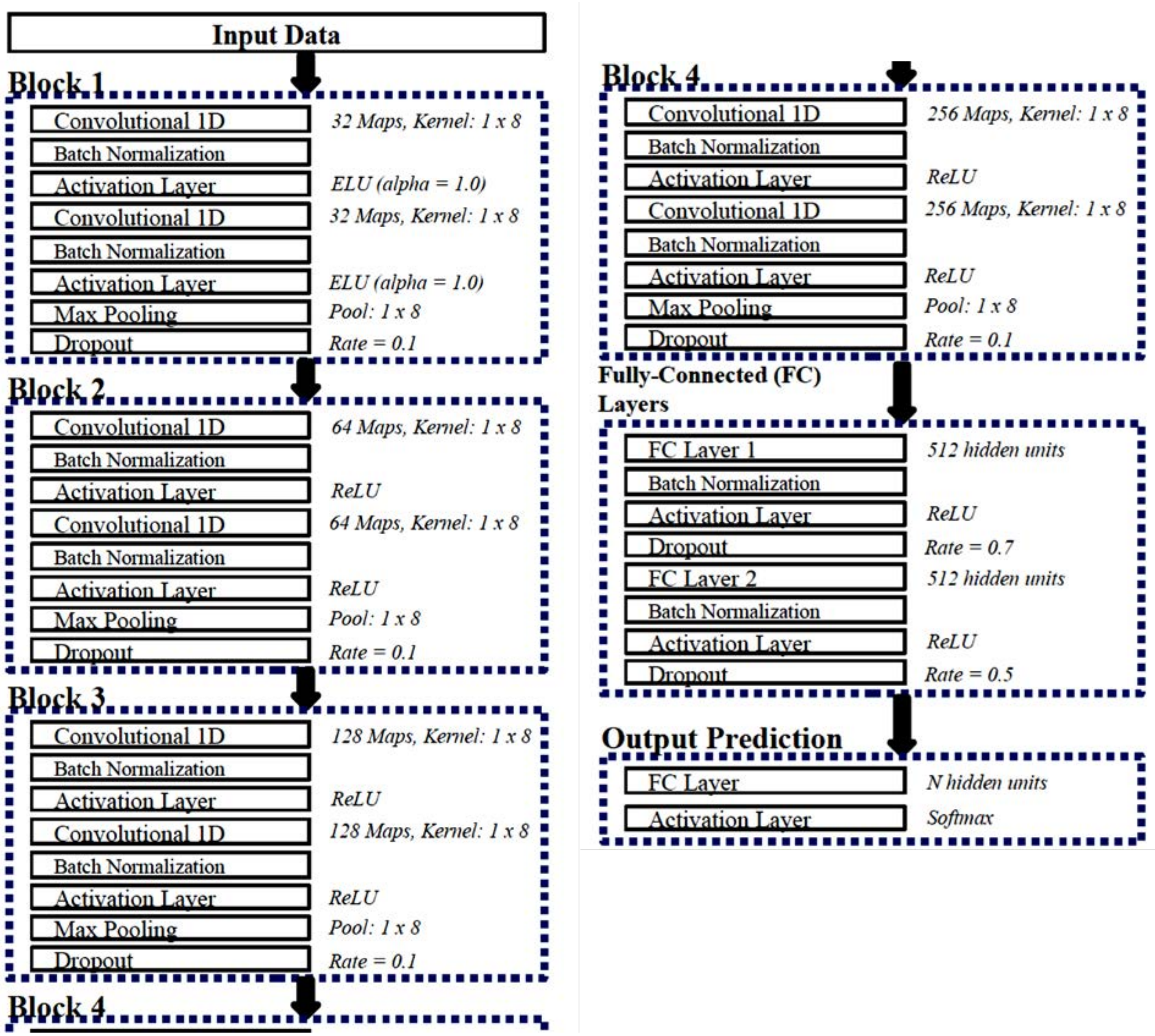


RIT

Table 1: Hyperparameters selection for DF model from Extensive Candidates Search method

Hyperparameters	Search Range	Final
Input Dimension	[500 ... 7000]	5000
Optimizer	[Adam, Adamax, RMSProp, SGD]	Adamax
Learning Rate	[0.001 ... 0.01]	0.002
Training Epochs	[10 ... 50]	30
Mini-batch Size	[16 ... 256]	128
[Filter, Pool, Stride] Sizes	[2 ... 16]	[8, 8, 4]
Activation Functions	[Tanh, ReLU, ELU]	ELU, ReLU
Number of Filters		
Block 1 [Conv1, Conv2]	[8 ... 64]	[32, 32]
Block 2 [Conv3, Conv4]	[32 ... 128]	[64, 64]
Block 3 [Conv5, Conv6]	[64 ... 256]	[128, 128]
Block 4 [Conv7, Conv8]	[128 ... 512]	[256, 256]
Pooling Layers	[Average, Max]	Max
Number of FC Layers	[1 ... 4]	2
Hidden units (each FCs)	[256 ... 2048]	[512, 512]
Dropout [Pooling, FC1, FC2]	[0.1 .. 0.8]	[0.1, 0.7, 0.5]

RIT


Input Data
Block 1
Convolutional 1D
32 Maps, Kernel: 1 x 8
Batch Normalization
Activation Layer
ELU (alpha = 1.0)
Convolutional 1D
32 Maps, Kernel: 1 x 8
Batch Normalization
Activation Layer
ELU (alpha = 1.0)
Max Pooling
Pool: 1 x 8
Dropout
Rate = 0.1
Block 2
Convolutional 1D
64 Maps, Kernel: 1 x 8
Batch Normalization
Activation Layer
ReLU
Convolutional 1D
64 Maps, Kernel: 1 x 8
Batch Normalization
Activation Layer
ReLU
Max Pooling
Pool: 1 x 8
Dropout
Rate = 0.1
Block 3
Convolutional 1D
128 Maps, Kernel: 1 x 8
Batch Normalization
Activation Layer
ReLU
Convolutional 1D
128 Maps, Kernel: 1 x 8
Batch Normalization
Activation Layer
ReLU
Max Pooling
Pool: 1 x 8
Dropout
Rate = 0.1
Block 4
Block 4
Convolutional 1D
256 Maps, Kernel: 1 x 8
Batch Normalization
Activation Layer
ReLU
Convolutional 1D
256 Maps, Kernel: 1 x 8
Batch Normalization
Activation Layer
ReLU
Max Pooling
Pool: 1 x 8
Dropout
Rate = 0.1
Fully-Connected (FC) Layers
FC Layer 1
512 hidden units
Batch Normalization
Activation Layer
ReLU
Dropout
Rate = 0.7
FC Layer 2
512 hidden units
Batch Normalization
Activation Layer
ReLU
Dropout
Rate = 0.5
Output Prediction
FC Layer
N hidden units
Activation Layer
Softmax

RIT

User Interfaces

Impact of Two Factor Authentication

Josephine Wolff
Public Policy

Why Use 2FA?

- Mitigate phishing
- Password breaches

Research Questions:

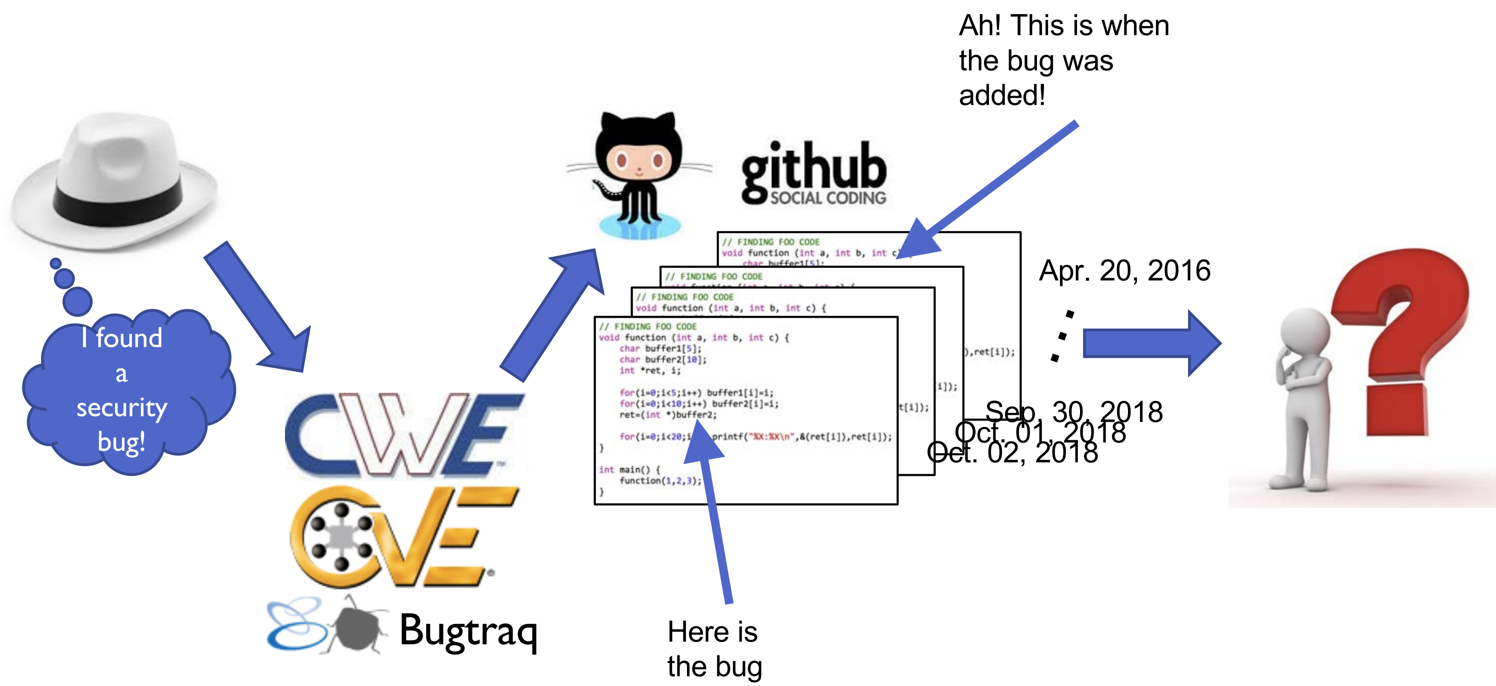
- Impact of 2FA on account compromises
- Which technologies do users adopt?
  - Key fob, smartphone app, SMS (text) code, phone call
- Barriers to usability and adoption

RIT

Tools for Professionals

Mining to understand security bugs

Andy Meneely
Software Engineering

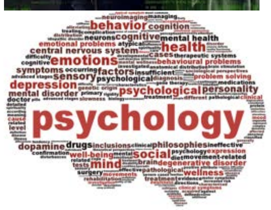


RIT

Modeling

Modeling Attackers

Jay Yang
Computer Engineering

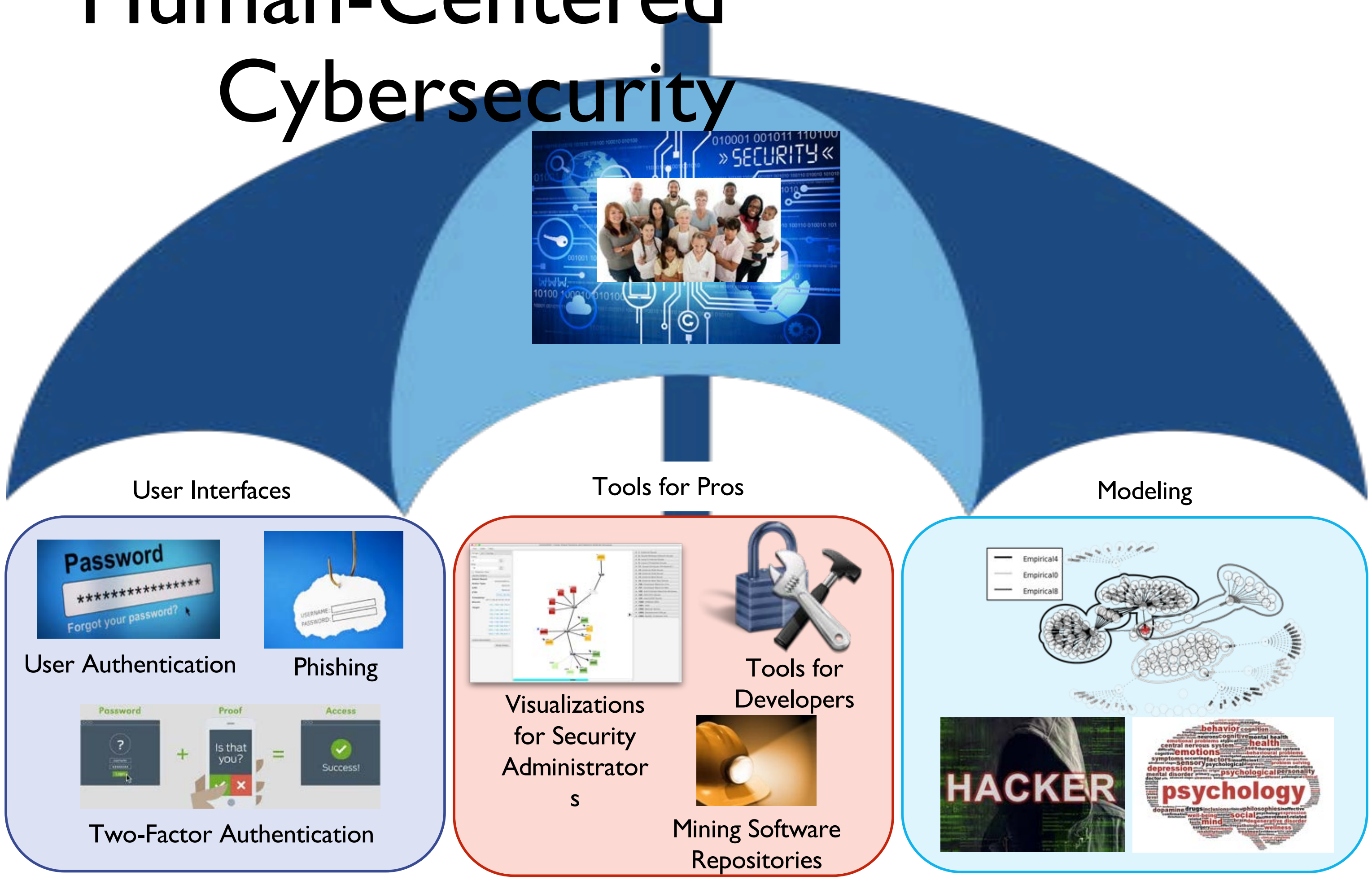


Advanced Cyber Attack Simulations

Better Understanding

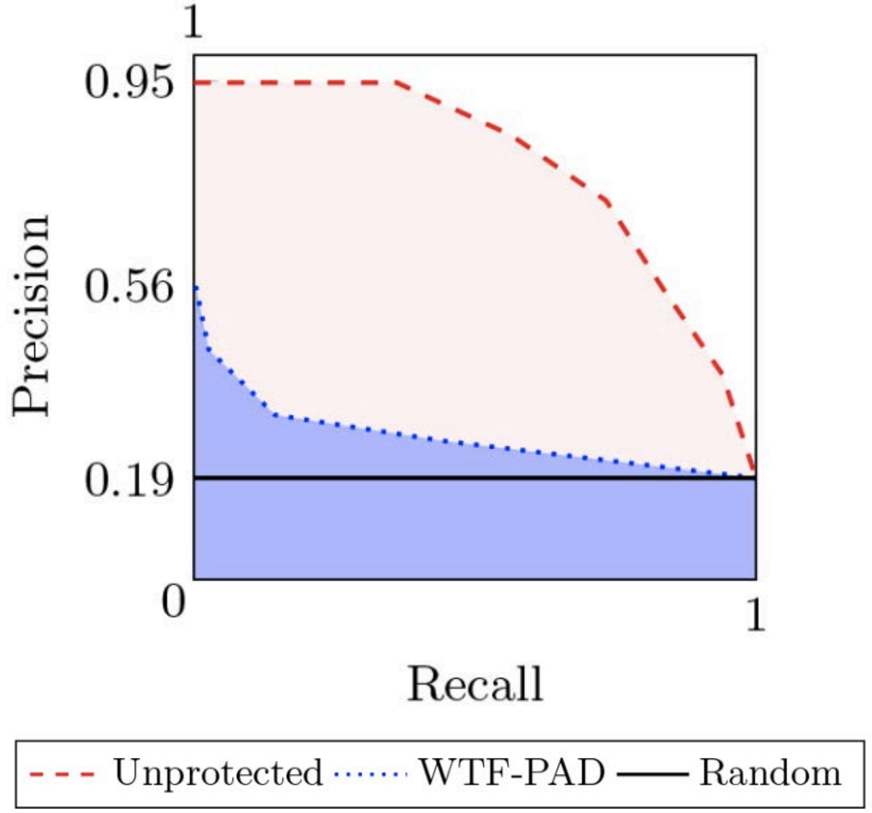
Better Defenses

Human-Centered Cybersecurity



Results

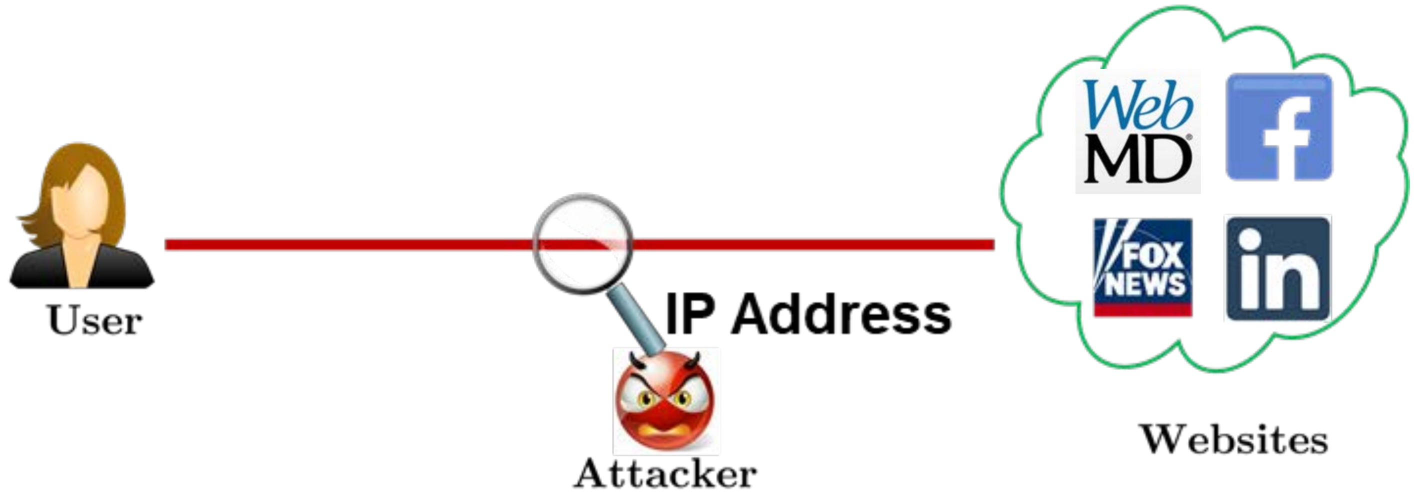
- No added delays
- 54% bandwidth overhead
- Much worse for the attacker



RIT

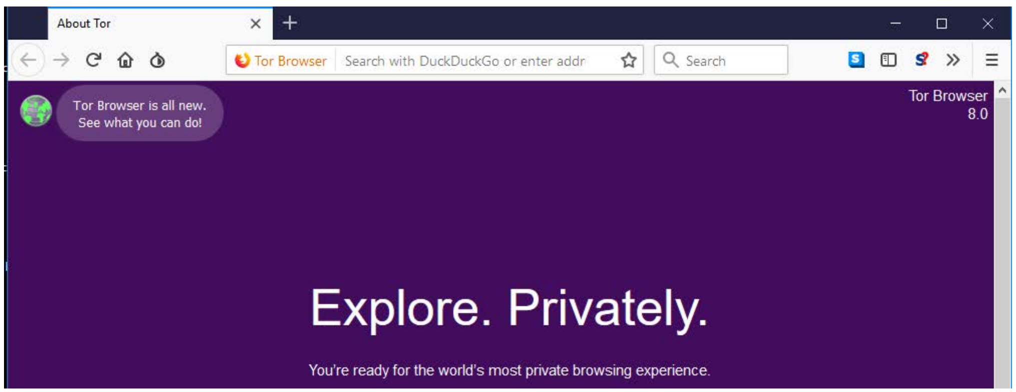
Website Fingerprinting in Tor

Website Fingerprinting in Tor



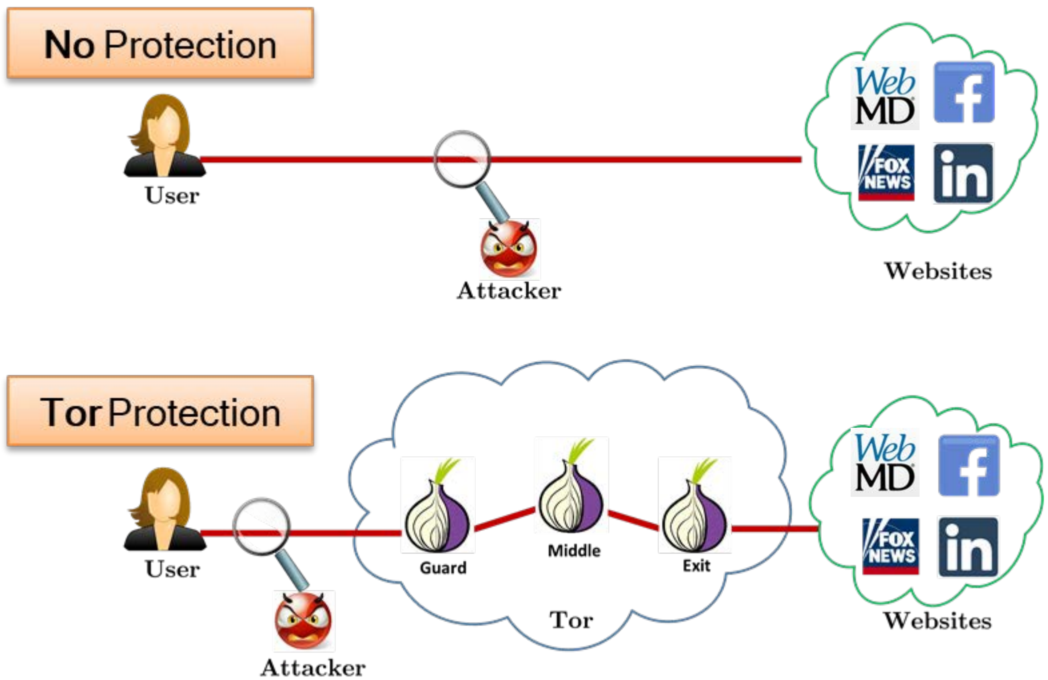
The attacker can easily learn user's Internet behavior

Website Fingerprinting in Tor



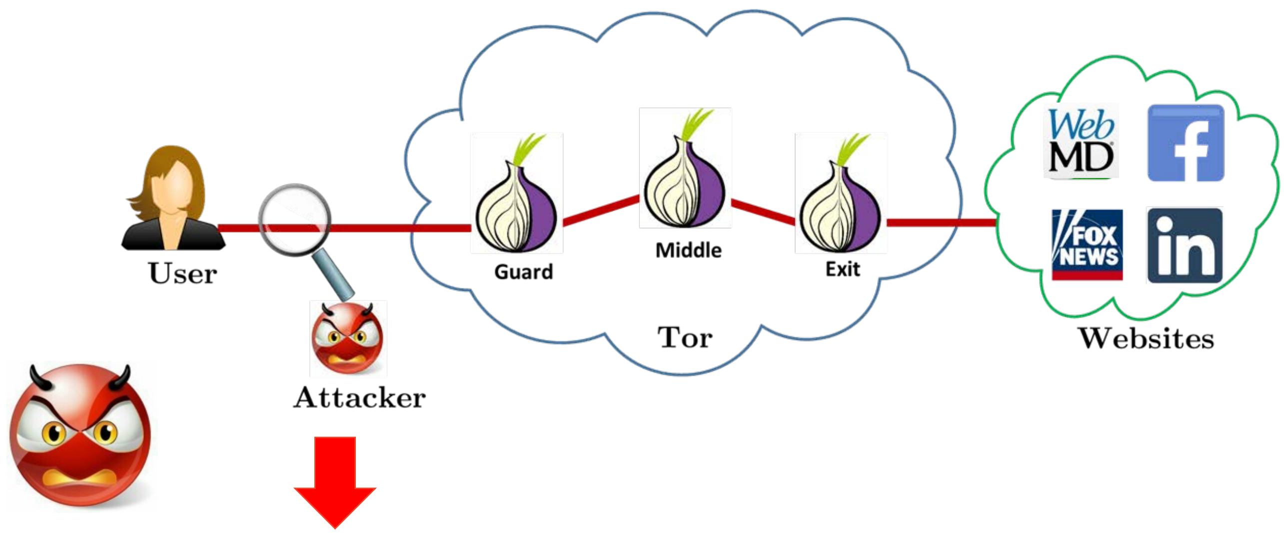
Tor: Privacy Enhancing Technology

Website Fingerprinting in Tor



No individual node has the complete path information

Website Fingerprinting in Tor

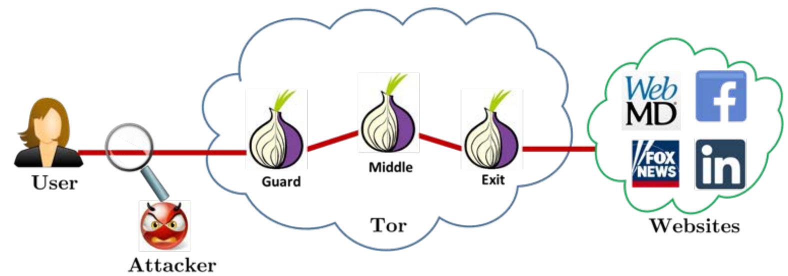


User ← → Guard Node

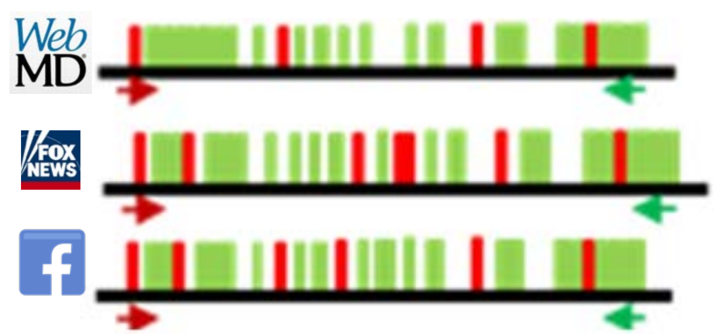
The attacker <u>fails to link</u> user to the actual website

Website Fingerprinting in Tor

- **WF Attacks**
  - Try to link **the user** to **the website**



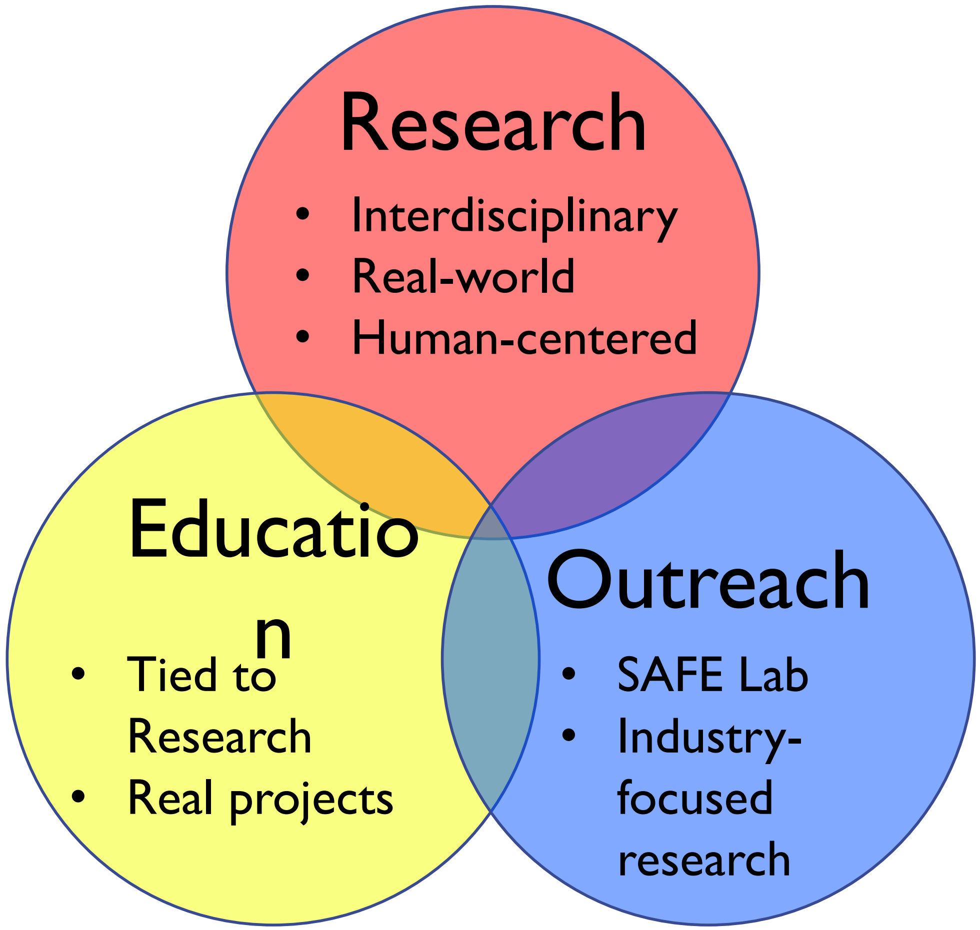
Side-channel information
e.g. Packets statistics



Information Leak
- Packets Statistic
- Burst of packets

* Unique for each website

Center Mission



90%+ Accuracy*

* For ~100 sites, not pages

Closed vs. Open World

Monitored- vs Unmonitored Websites

Closed vs. Open World

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

Closed-World Scenario

- Users only visit monitored sites
- **Accuracy** of the attack
- Unrealistic

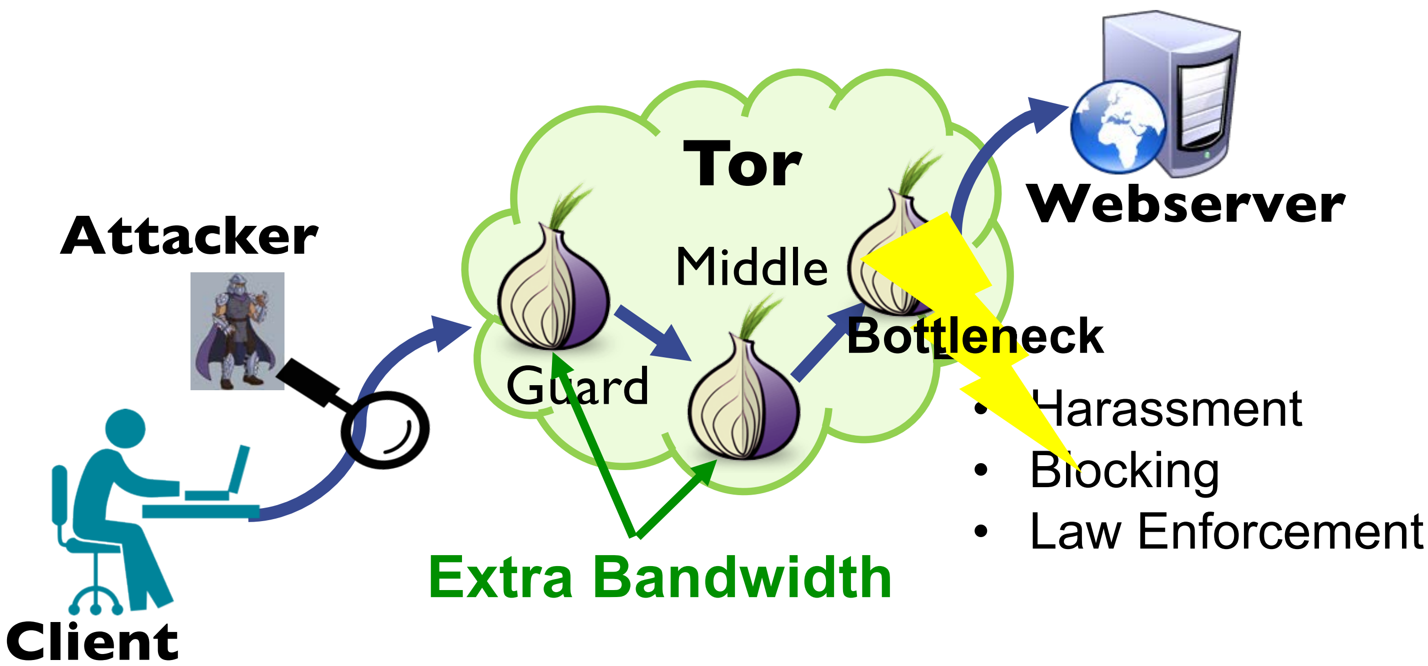
Closed vs. Open World

Open-World Scenario

- Users can visit any website (> Billions)
- Recognizing monitored vs. unmonitored
- Matt's Rule of Thumb
  - 90+% CW Accuracy → **High Danger**

++?

RIT


Tor
Attacker
Middle
Webserver
Bottleneck
Guard
• Harassment
• Blocking
• Law Enforcement
Extra Bandwidth
Client