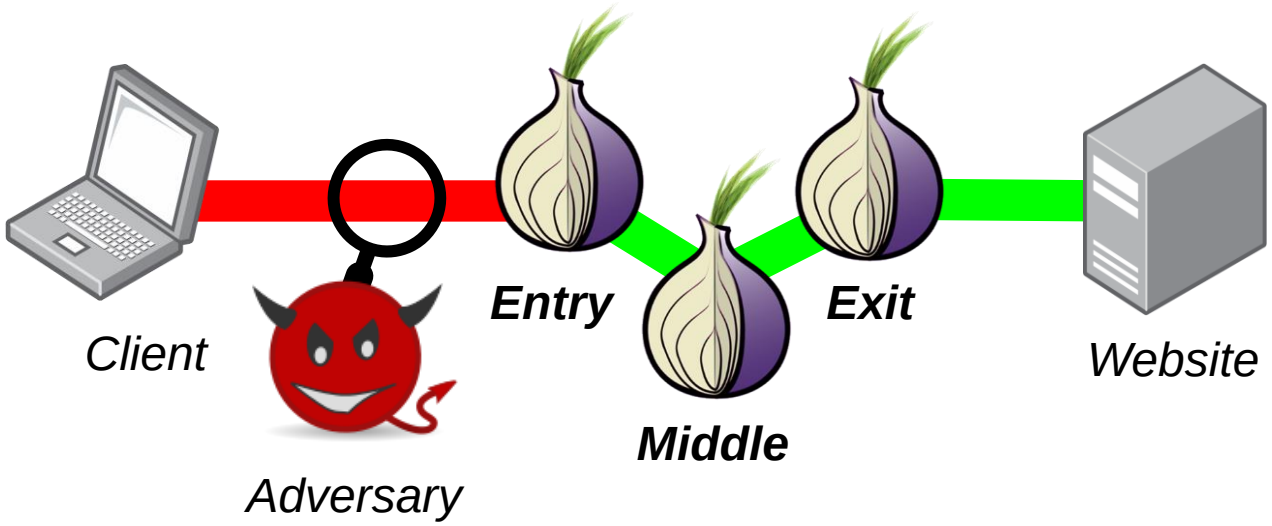
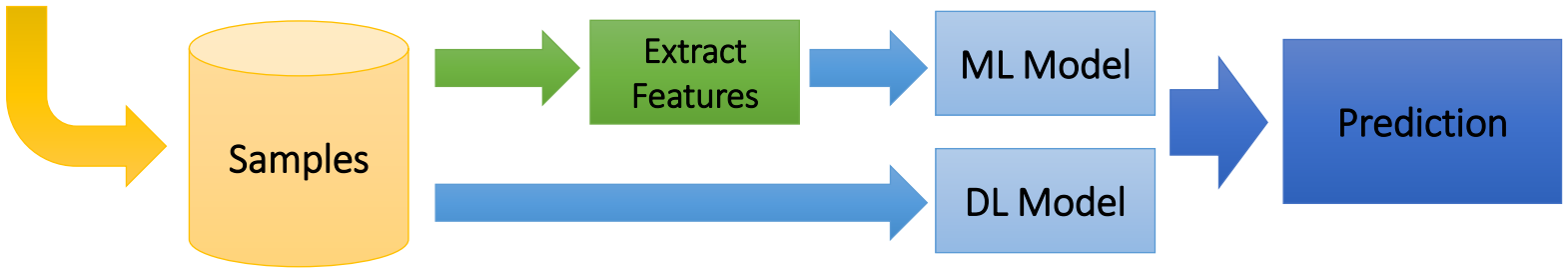
Nate Mathews, nate.mathews@mail.rit.edu

Website Fingerprinting (Single-tab)

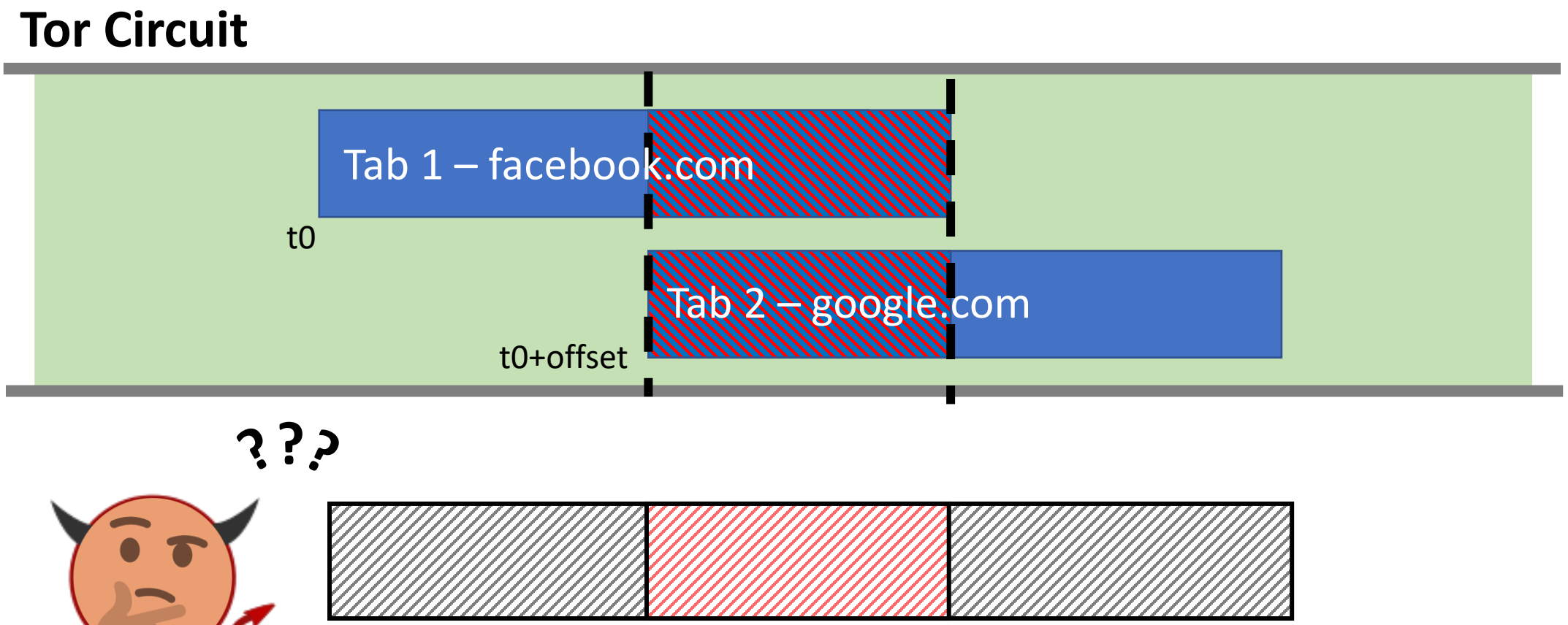


– Use traffic patterns to link client to website



Nate Mathews, nate.mathews@mail.rit.edu

Website Fingerprinting (Multi-tab)



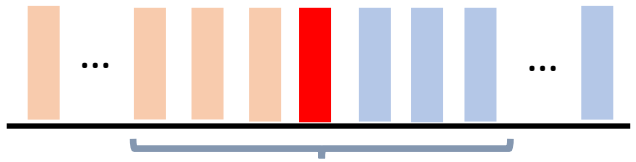
- WF accuracy reduced significantly *[PETS'16]*

Nate Mathews, nate.mathews@mail.rit.edu

WF Multi-tab (background)

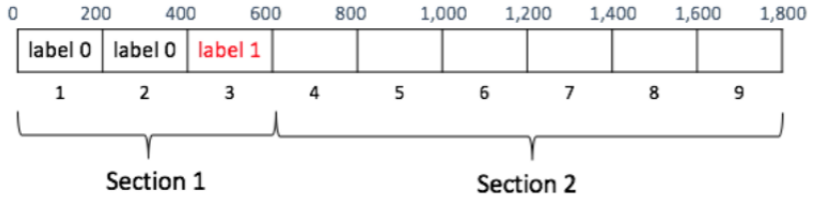
- **2016 [PETS] - On Realistically Attacking Tor with Website Fingerprinting**
  - Generate 23 features for every packet in a sample
  - Use k-Nearest Neighbors to score most probable packet for start of 2nd page
- **2018 [ACSAC] - A Multi-tab Website Fingerprinting Attack**
  - Re-uses [PETS'16]'s 23 features
  - Uses XGBoost with undersampling
- **2019 [AsiaCCS] - Revisiting Assumptions for Website Fingerprinting Attacks**
  - Split sample into blocks
  - Use a Hidden Markov Model to classify each block as *Tab 1* or *Tab 2*

PETS'16 & ACSAC'18



$[f_1, f_2, \ldots, f_{23}]$

AsiaCCS'19



Nate Mathews, nate.mathews@mail.rit.edu

Deep Learning for Multi-tab

- ***Why?***
  - Automatic extraction of features from 'raw' inputs
  - More 'powerful' features learned
    - *DL in Single-page increased acc. and defeated several defenses* [CCS'18]

- ***How?***
  - Treat it like an audio segmentation problem:
    - *E.g. Ingest the sample as time-series data and divide into overlap & non-overlap regions*
    - Bi-directional LSTM w/ convolutional layers for feature extraction [EURASIP'20]

[CCS'18] Sirinam et al. "*Deep Fingerprinting: Undermining Website Fingerprinting Defenses with Deep Learning*"
[EURASIP'20] Gimeno et al. "*Multiclass audio segmentation based on recurrent neural networks for broadcast domain data*"