D1, D3: Leveraging Asymmetric Data for Detecting Cyber Attacks

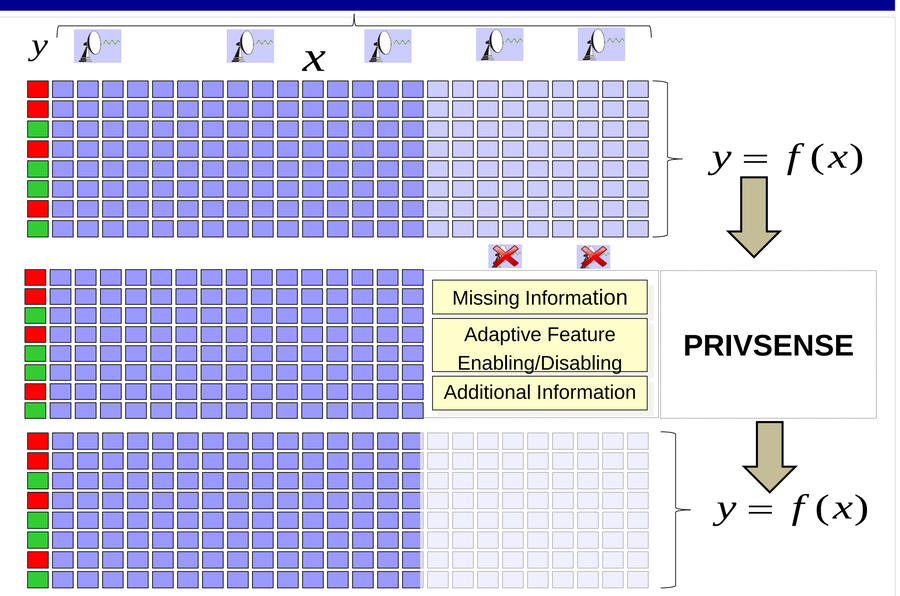
Research Lead: Dr. Ritu Chadha, Applied Communication Sciences, rchadha@appcomsci.com

Introduction and Research Goals

- **Problem:** Data for detecting cyber attacks may not be available at both the training and the deployment phases:
  - Data from a source may be "missing" and therefore unavailable for processing.
  - Data from a source may be ignored if it is suspected to be compromised
  - Data from a source may be restricted due to privacy issues
  - Data from a source may be too expensive or impractical to obtain at deployment time, such as expert information produced by human analysts.

Terminology: We call features that are available for training but are not available at deployment "*privileged features*", and we call feature that are available both at training and deployment time "standard features"

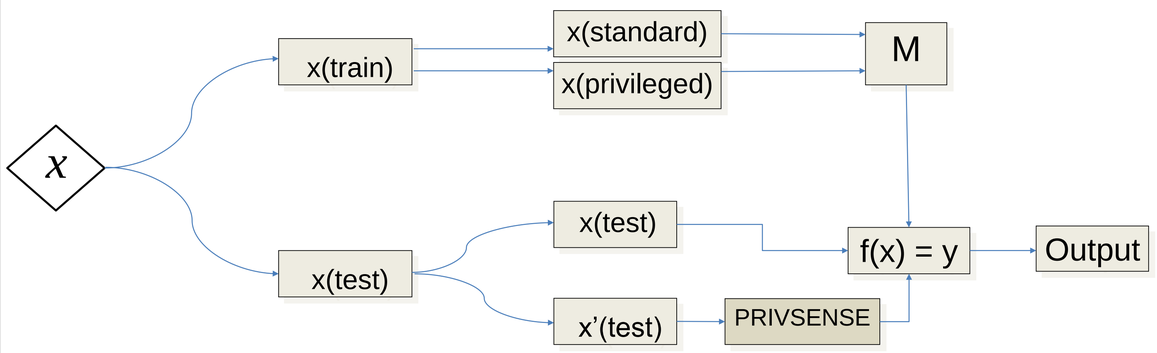
PrivSense: A novel system that aims to make use of privileged features only for training in order to improve the accuracy of cyber attack detection



Technical Approach

State of the Art:

- Current machine learning techniques cannot make use of observables for training if the same observables are not available at test time
- Drawback: Degraded accuracy of detection due to loss of information



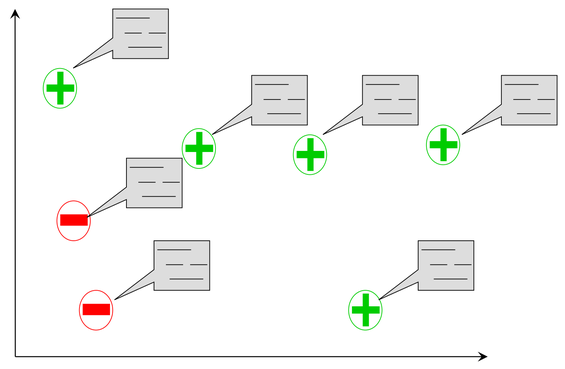
PrivSense Design:

- x(train) is the dataset available during training, x(test) is the test dataset which is in same space as x(train), and x'(test) is the test dataset without privileged features.
- M is the model used as a learning algorithm, and f is the hypothesis it learns from the x(train).

PrivSense makes continuous evaluation of f(x) on x'(eval) by predicting the privileged features from standard ones.

Underlying Theory: Learning Using Privileged Information (LUPI) [1]

- Advanced ML paradigm enables development of models for cyber attack detection



- Given training data $(x_1,y_1),\ldots,(x_L,y_L)$ and privileged data $x_1^*,\ldots, x_L^*$
- Generalize data to a rule (function) $y=f(x)$, where $x \in X$, $x^* \in X^*$, and $y \in \{-1, +1\}$

1 Where does "Privileged" Information come from? Some possible sources include:

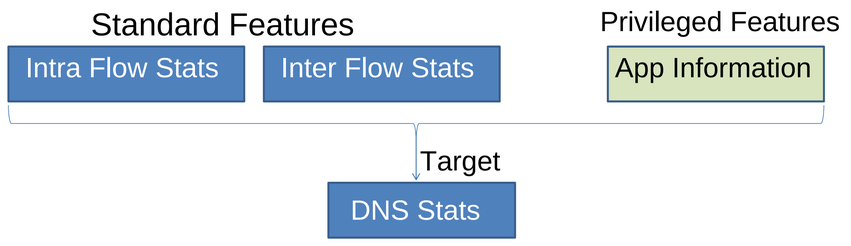
- Features that are too expensive to collect on a deployed system, such as detailed system performance data
- Features that are not available at deployment time, such as expert analysis
- Features that are missing, e.g., from malfunctioning sensors
- Features that may be compromised due to infection of the source system

2 How to use privileged features:

- Transferring knowledge from standard features to derive privileged features (Knowledge Transfer)
- Use privileged features during training to learn a more accurate unified model (SVM+)

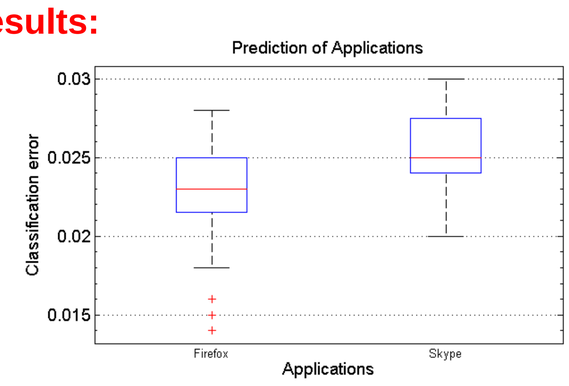
Results

Case Study:



- Investigating applicability of PrivSense to detection of botnet clients that use DGA (Domain Generation Algorithm) for botnet C&C
- Differentiate between "legitimate" DNS queries resulting in NXDOMAIN responses (e.g., generated through genuine user/app typos/errors) and those generated by a DGA

Preliminary Results:



Knowledge Transfer accuracy for privileged features

Research Plan/Next Steps

- Complete PrivSense evaluation on DGA case study using CyberVAN testbed
- Use new evaluation metrics to measure PrivSense performance gain

New Evaluation Metrics:

- Expected case: Use **both available** spaces, error rate = Y
- PrivSense: Use **unavailable** space, error rate = Z
- Current ML paradigm: Use **one available** space, error rate = X

Performance Gain = (X-Y) / (Z-Y)

Primary Researchers

Z. Berkay Celik	PSU	zbc102@cse.psu.edu
Patrick McDaniel	PSU	mcdaniel@cse.psu.edu
Rauf Izmailov	ACS	rizmailov@appcomsci.com
Ritu Chadha	ACS	rchadha@appcomsci.com
Constantin Serban	ACS	cserban@appcomsci.com
Roberto Pagliari	ACS	rpagliari@appcomsci.com

Task Rotations (listed by PI)

Dr. Ritu Chadha, PI (ACS), 10 days, ARL
Z. Berkay Celik, PhD. Student (PSU), 3 months, ACS

Collaborations

PSU	Penn State University
ACS	Applied Communication Sciences

[1] D. Pechyony, R. Izmailov, A. Vashist, V. Vapnik, "SMO-style Algorithms for Learning Using Privileged Information," in *Proceedings of the 2010 International Conference on Data Mining* (DMIN), 2010. (Best Academic Research Paper Award).