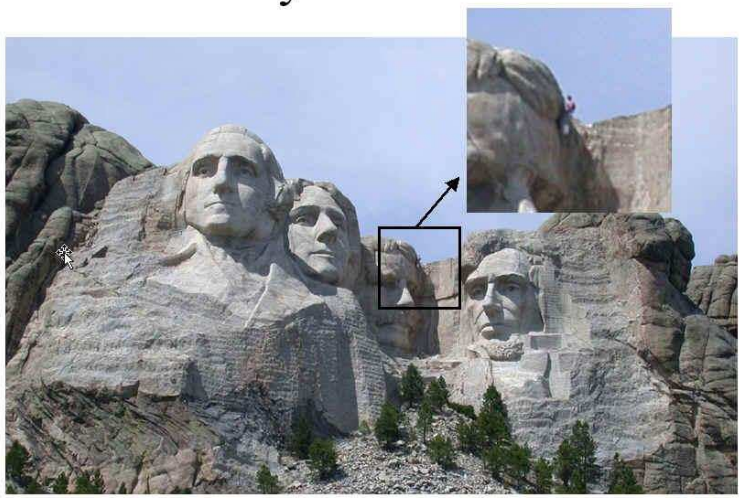
Selected Application Highlights

Machine Learning Team

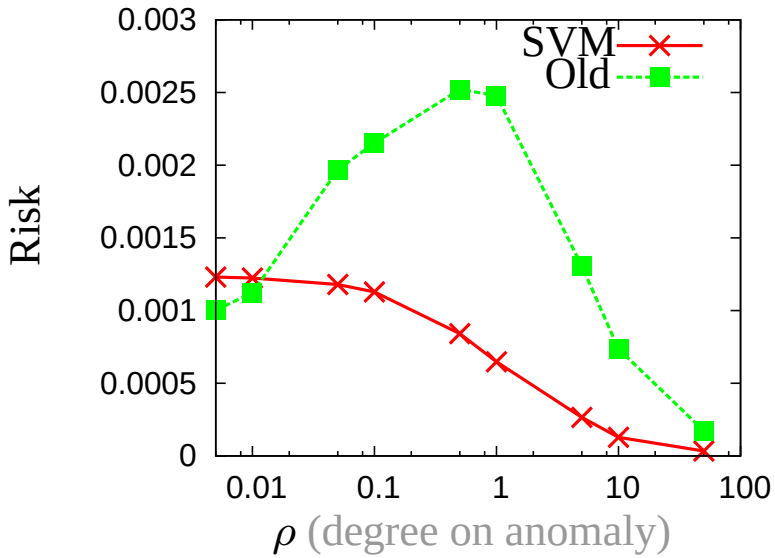
LA-UR-08-1661

Anomaly Detection

Anomaly -- "Rushmore"

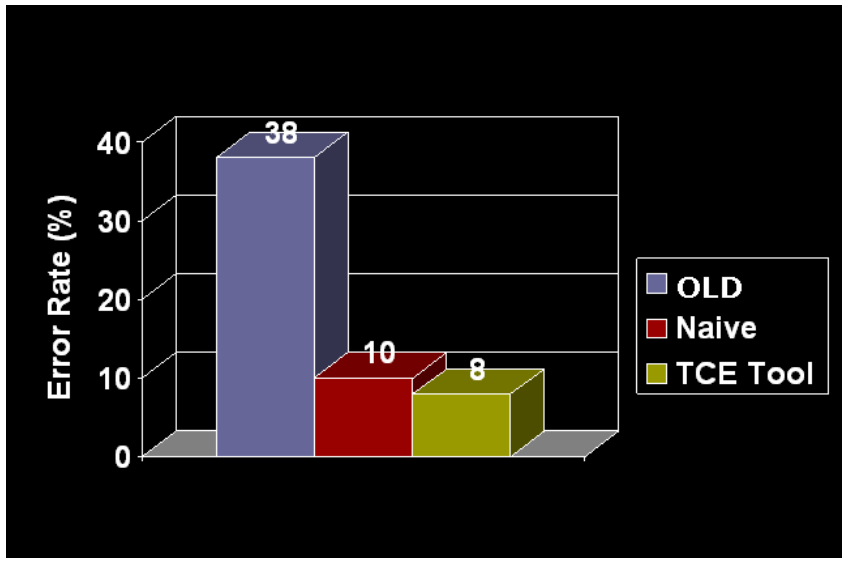


AD for Computer Network Traffic



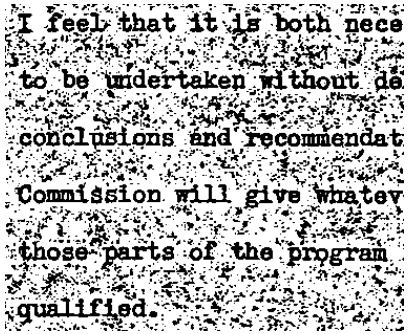
- Problem: detect anomalous points (i.e. points with low probability density)
- Challenge: validating the detection error
- Resolution: discovered a risk function that is calibrated to detection error and can be reliably estimated from field data
- Additional Contributions: new class of solution methods that minimize the calibrated risk *directly* (circumventing the need for density estimation)

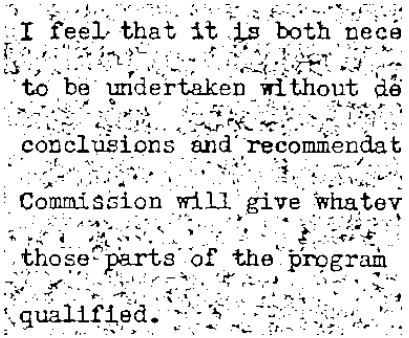
TCE Detection

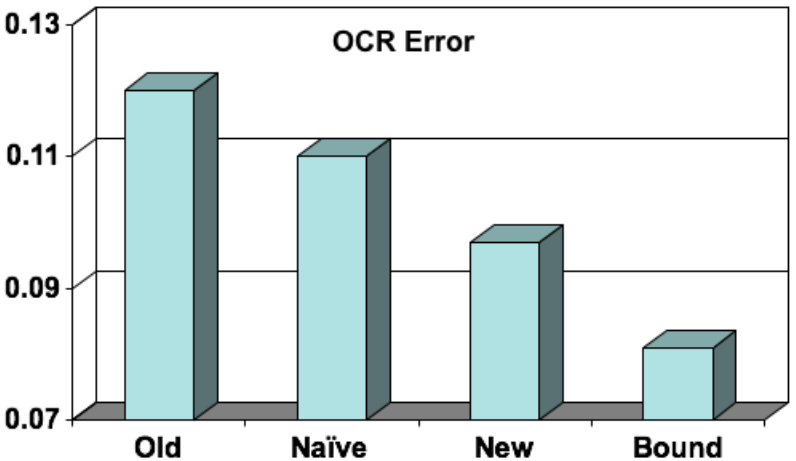


- Problem: design a system that uses the Tufts Artificial Nose to detect trichloroethylene (TCE).
- Challenge: fraction of TCE varies from one deployment to the next resulting in poor performance from traditional detector design methods
- Resolution: developed and implemented a *detector design method* that is *robust* to changes in the fraction of TCE (optimizes performance over a range of TCE fractions)

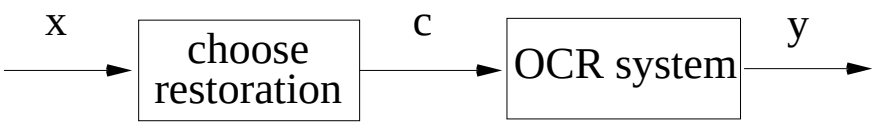
Document Processing



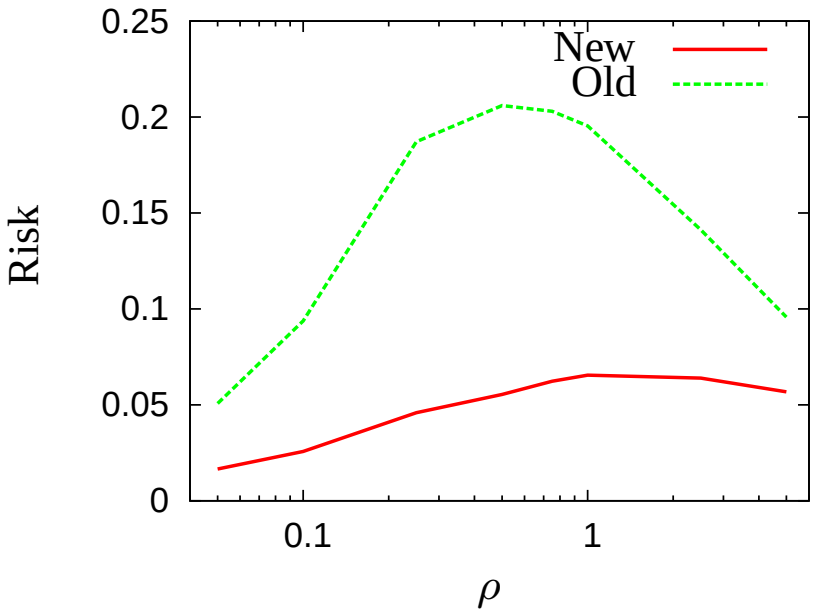




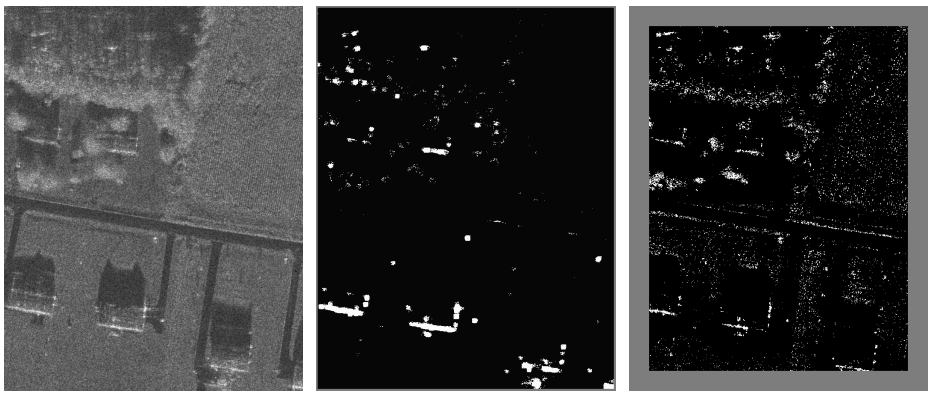
- Problem: restore images so that their OCR score is improved
- Challenge: different types of noise require different restoration methods
- Resolution: formulated and solved as a non-traditional multi-class problem (where the correct labels c are unknown)



SAR Image Segmentation



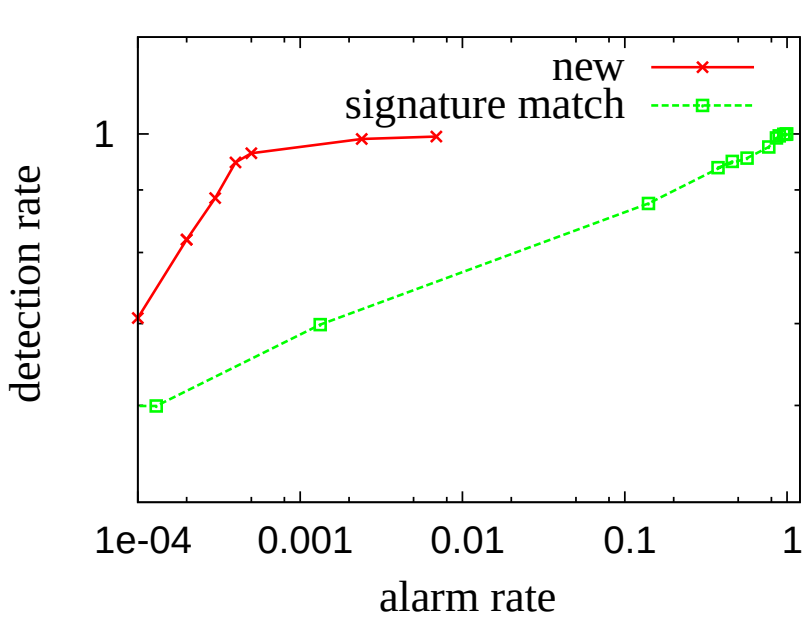
- Problem: identify *pixels of interest* in SAR
- Challenges:
  - validating the deployed error rate
  - different statistics for each deployment
- Resolutions:
  - discovered a risk function that is calibrated to error rate and can be reliably estimated from field data
  - new solution methods that minimize the calibrated risk *directly*, and *adaptively* (in-the-field)

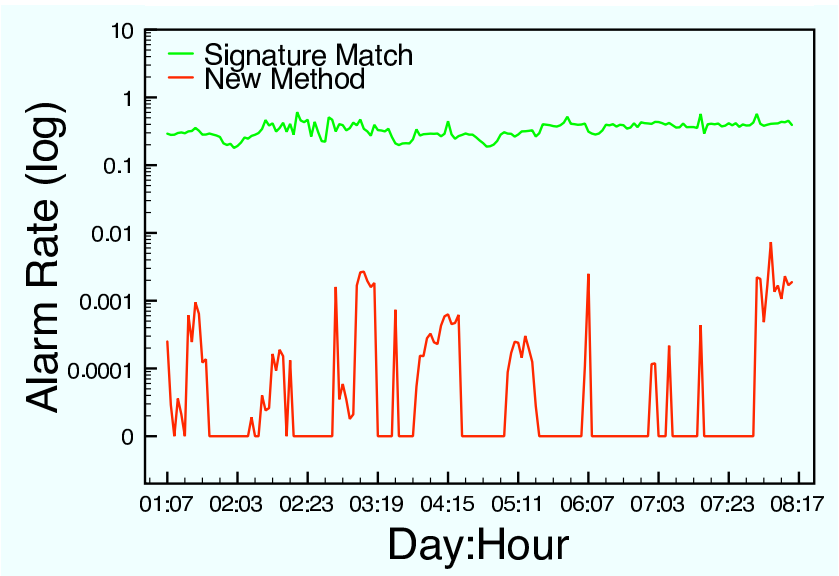


SAR Image New Old

Tank New Old

Network Monitoring for Cybersecurity



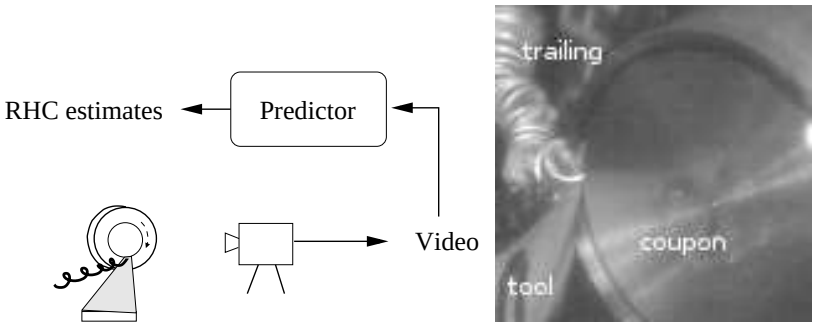


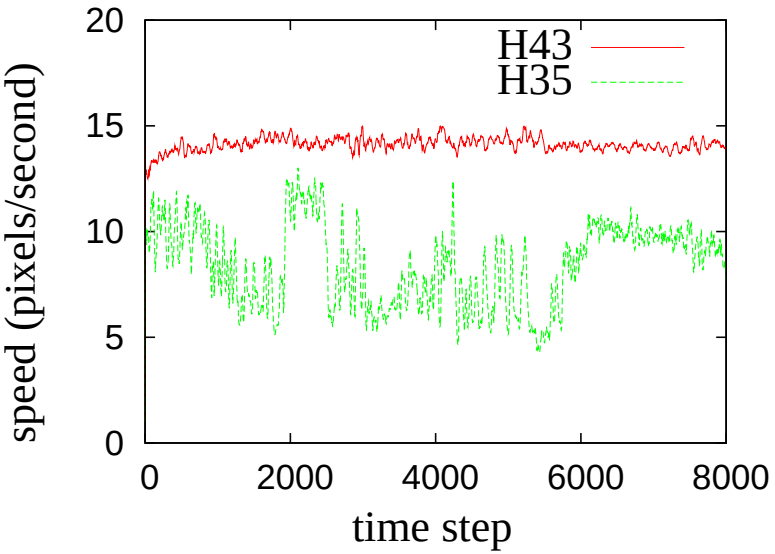
- Problem: detect *tunneled-CHAT* in *encrypted* network traffic
- Challenges:
  - limited information (due to encryption)
  - validating the deployed error rate
  - changing statistics (traffic patterns)
- Resolutions:
  - extract *current* CHAT info from *current* unencrypted traffic *meta-data*:

Packet Sizes	132, -122, 43, 28, -27, 23
Wait Times	-0.081, 0.003, -0.183, 0.002

  - discovered a risk function that is calibrated to error rate and can be reliably estimated from field data
  - new solution methods that minimize the calibrated risk *directly*, and *adaptively* (in-the-field)

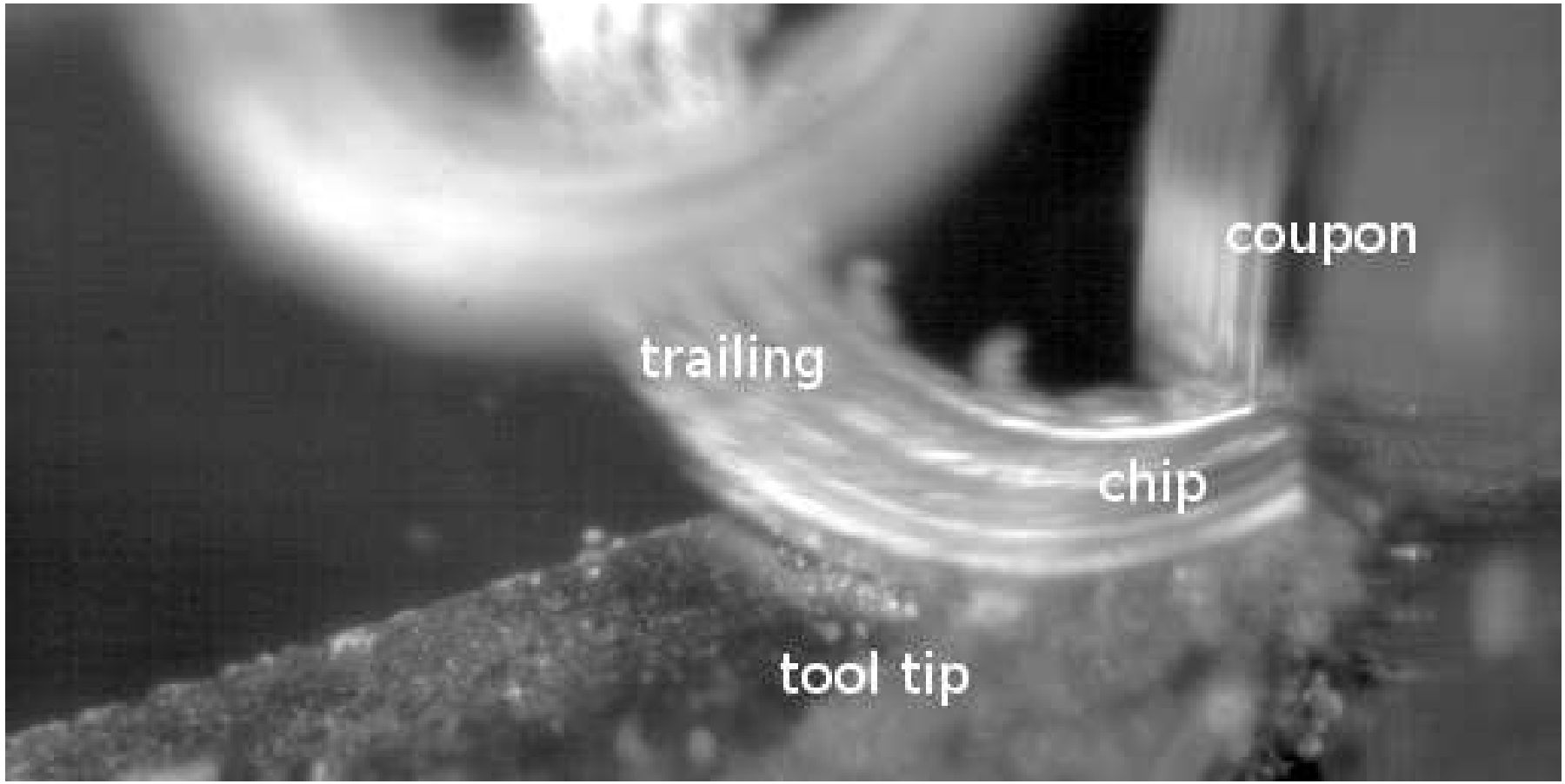
Predicting Material Hardness





Predicted Hardness Accuracy		
Uncalibrated	Calibrated	Naive
± 1.9	± 0.8	± 3.9

- Problem: use high-speed video of the cutting process to predict the RHC hardness value *on-the-fly*
- Challenges:
  - limited hardness information
  - invariant representation of video info
  - noise (severe lighting variations)
- Resolutions:
  - discovered/validated a physics-based relation between hardness and *chip speed* (see next slide)
  - designed multi-stage noise suppression system
  - discovered unexpected variability in the machining process (calibration)


coupon
trailing
chip
tool tip