Capability Presentation for SoURCE CODE

Christopher Taylor
Senior Research Engineer
Tactical Computing Labs

tactcomplabs.com

Tactical Computing Labs

- Research and Development Firm
  - Scientific Computing & RISC-V Hardware/Software Solutions
- Supercomputing/High Performance Computing (HPC)
  - Compilers/Runtime Systems
  - Hardware Simulation, Design/Testing/Evaluation/Implementation
- Numerical/Scientific Software
  - Machine Learning/AI

Tactical Computing Labs

- Founded in North Texas w/multiple CONUS Locations
  - In-House Data Center Facility
- Commercial Support for Structural Simulation Toolkit (SST) (Sandia)
- RISC-V Support for BLIS, HPX, OpenSHMEM, NVIDIA’s UCX, libfabric

SoURCE CODE

Tactical Computing Labs

Alignment with SoURCE CODE

- Tactical Computing Labs hosts an analytic capability called ***CIVA***
  - Compiler Integrated Vulnerability Analyzer (CIVA)
- CIVA is extensible to multiple programming languages
- CIVA is extensible to binary (compiled) application software

CIVA

Tactical Computing Labs

Compiler Integrated Vulnerability Analyzer (CIVA)

- Currently Detects Source Code Vulnerabilities
- CIVA can be extended to other applications and purposes
- Machine Learning Capability
- Currently Targets C Application Software (extensible to other languages)
- Originally developed under DARPA Cyber Fast Track (CFT)

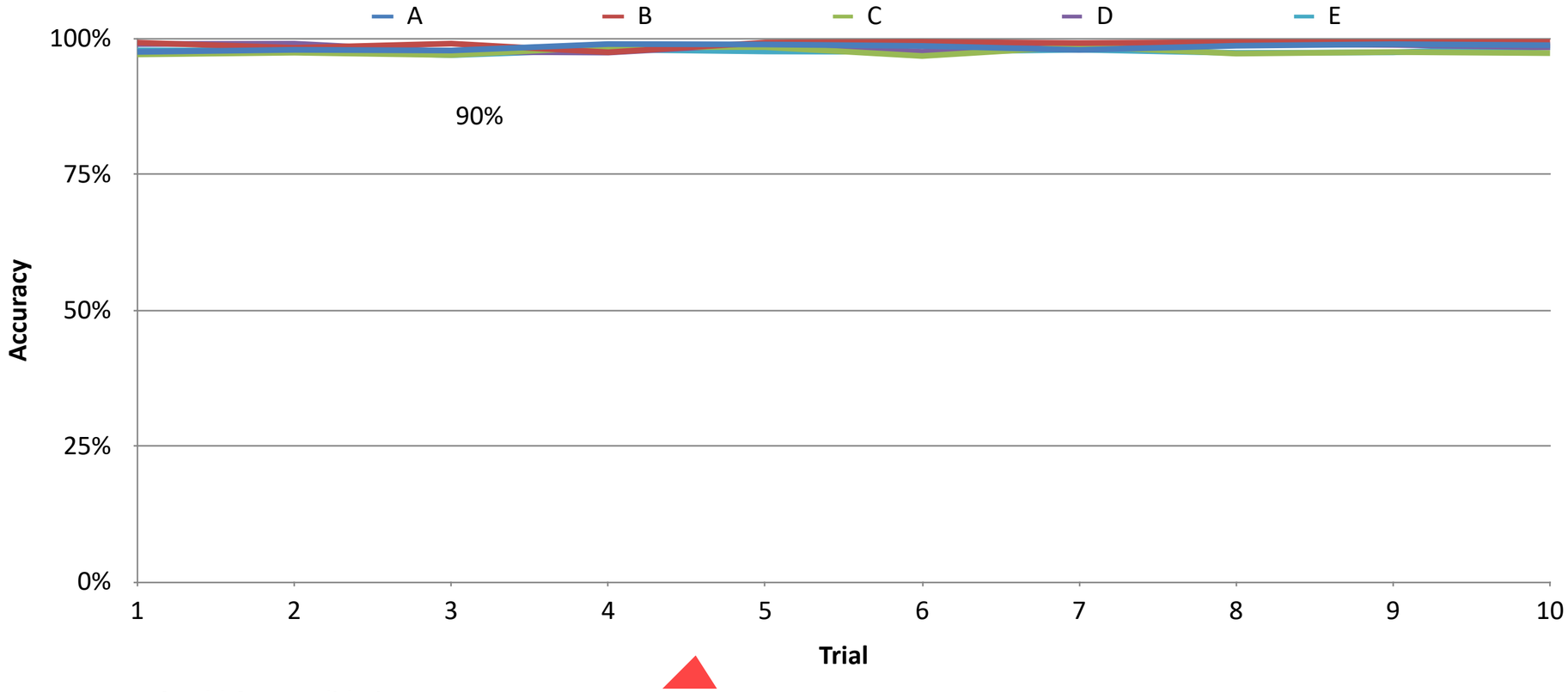
CIVA

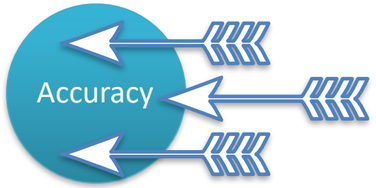
Tactical Computing Labs

Compiler Integrated Vulnerability Analyzer (CIVA)

- The following sides show performance for the following vulnerabilities:
  - A - stack-based buffer overflow
  - B - heap-based buffer overflow
  - C - integer overflow/wraparound
  - D - divide by 0
  - E - free of memory not on heap

CIVA - Accuracy





True Positives

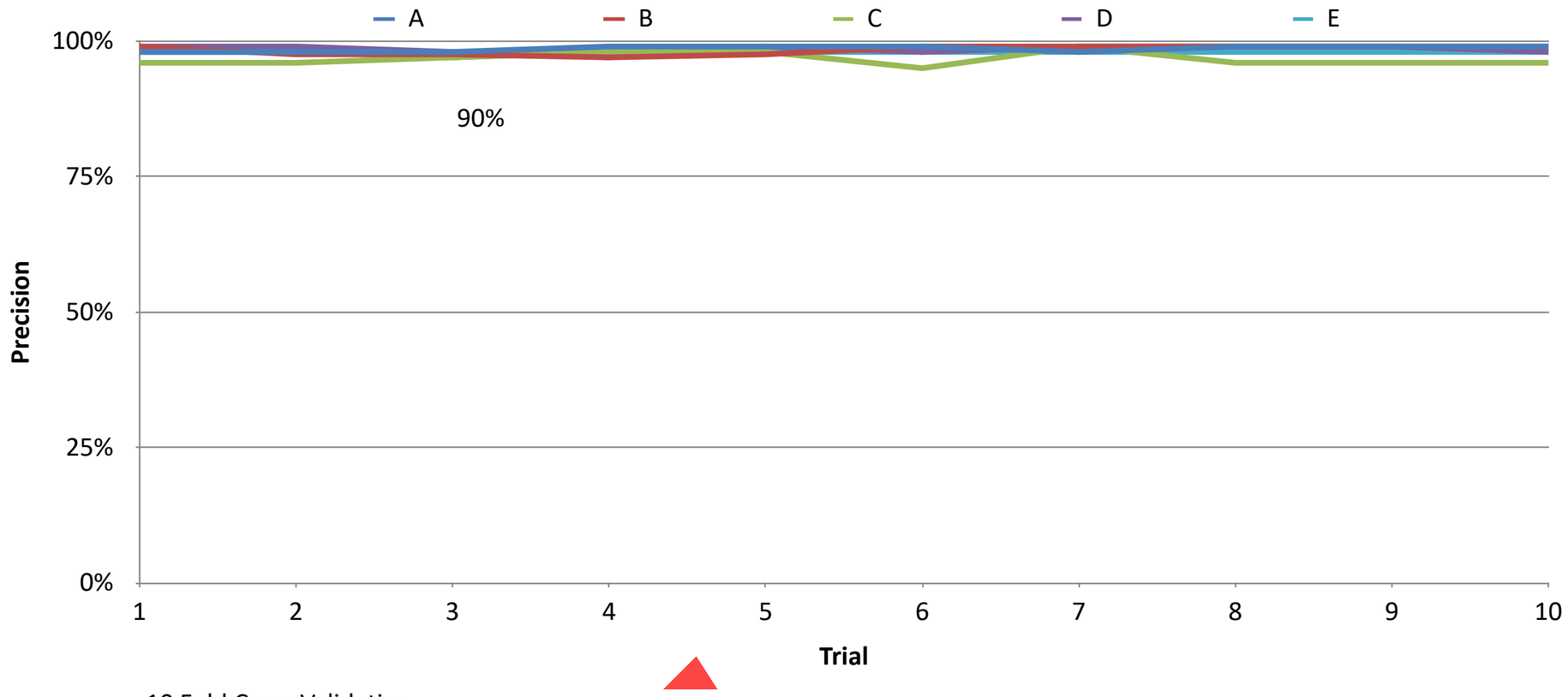
Total # Classifications

(True/False Positives & True/False Negatives)

10 Fold Cross Validation

Testing, Training, Evaluation Data Randomly Selected Each Fold

CIVA - Precision

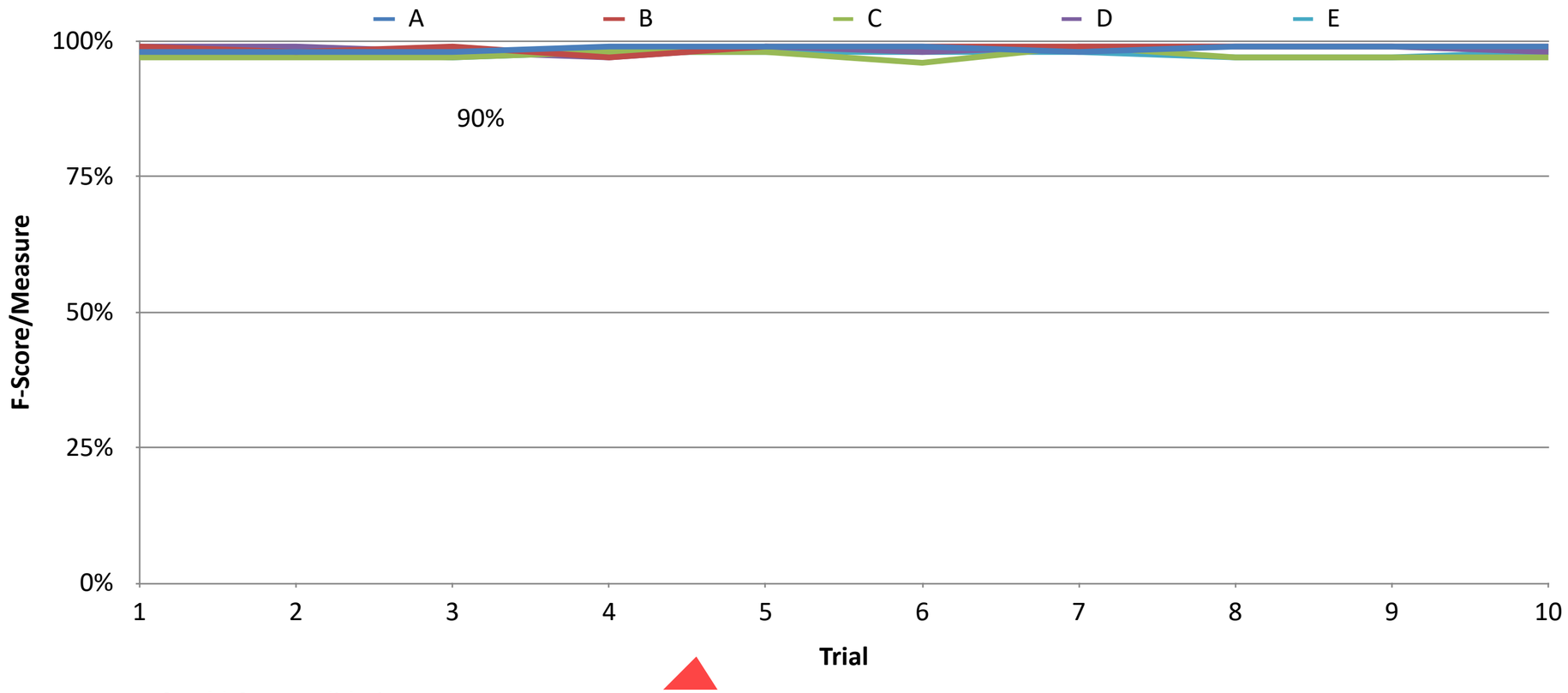


$$\frac{\#\text{ True Positives}}{\#\text{ True Positives} \;\&\; \text{False Positives}}$$

10 Fold Cross Validation

Testing, Training, Evaluation Data Randomly Selected Each Fold

CIVA - F-Score/Measure



F-Score/Measure

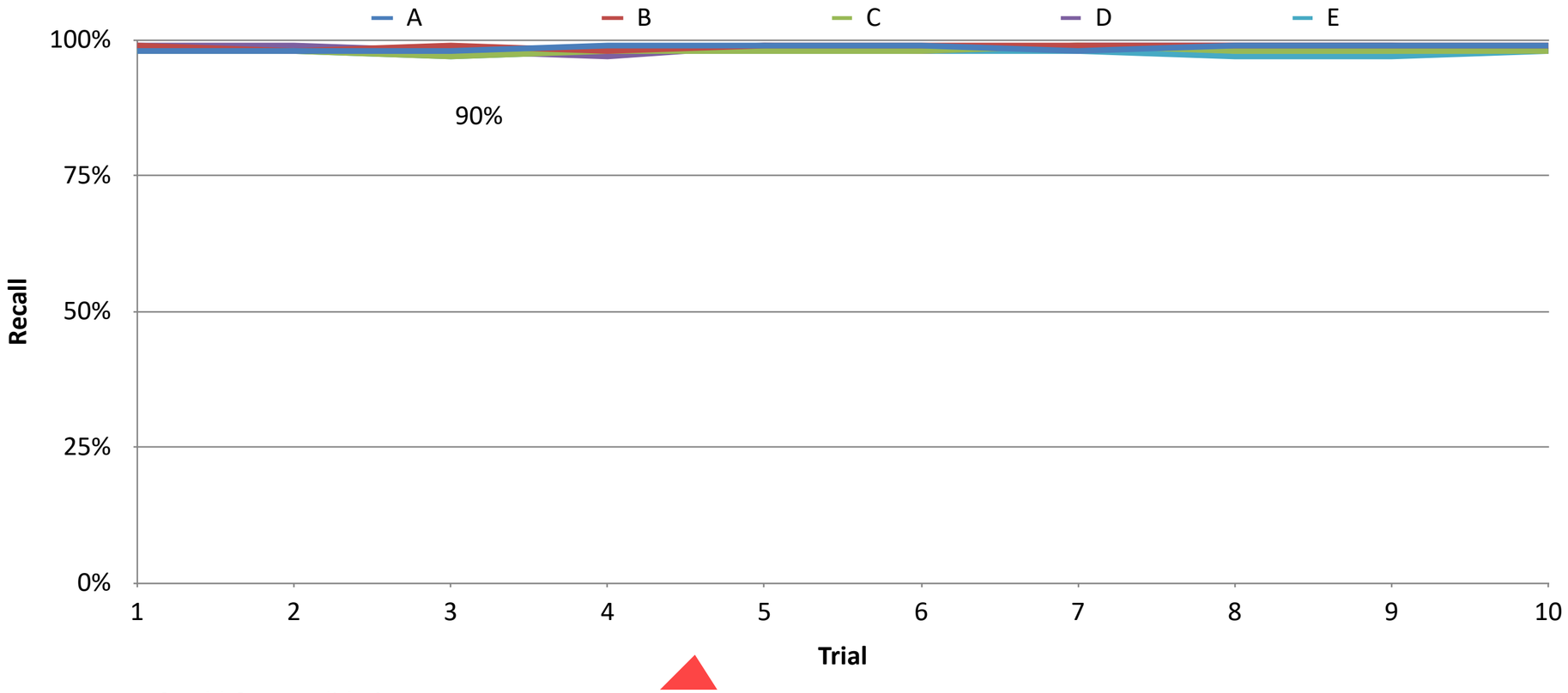
Precision x Recall

Precision + Recall

10 Fold Cross Validation

Testing, Training, Evaluation Data Randomly Selected Each Fold

CIVA - Recall



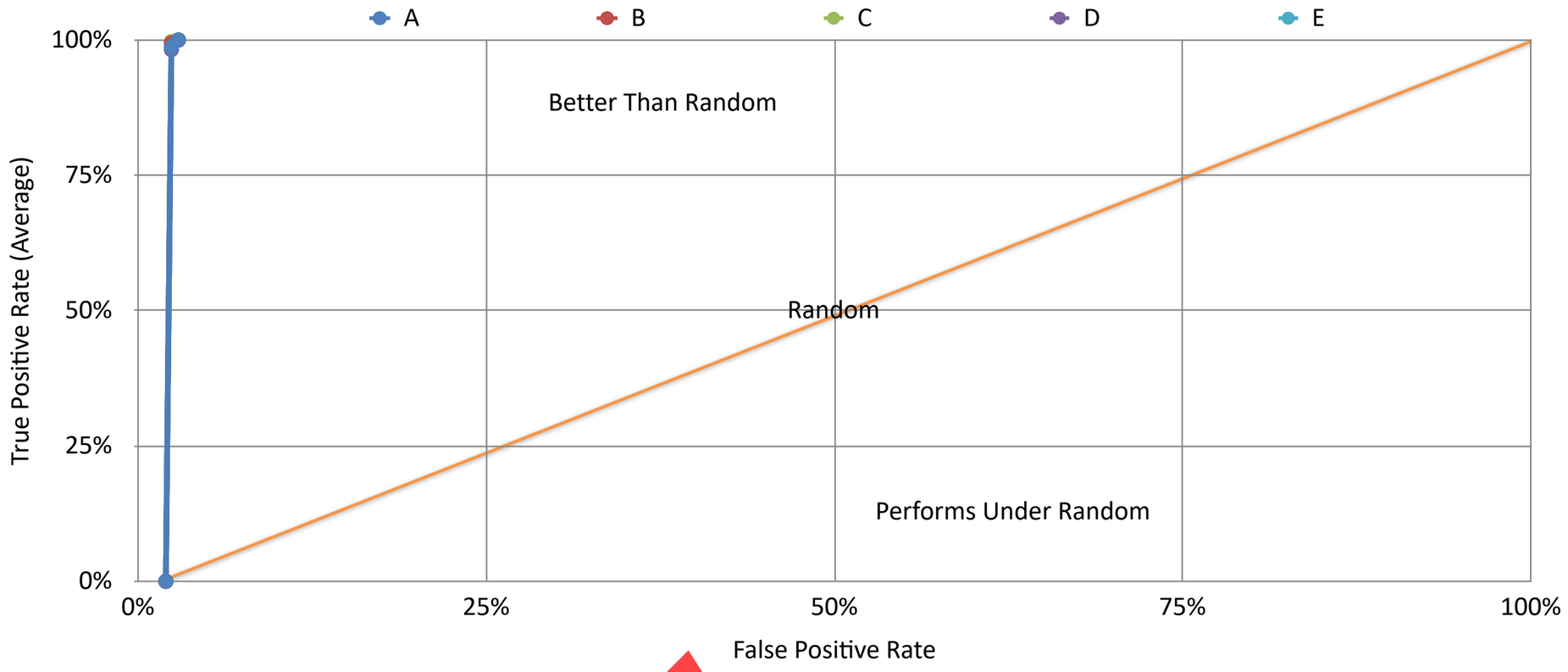
Recall

$$\frac{\text{\# True Positives}}{\text{\# True Positives \& False Negatives}}$$

10 Fold Cross Validation

Testing, Training, Evaluation Data Randomly Selected Each Fold

CIVA - ROC Curve



ROC Curve

True Positive Rate

False Positive Rate

Performance Better Than Random (Diagonal Orange Line)

ROC - Receiver Operating Characteristic Curve

Averaged over 10 Fold Cross Validation

Testing, Training, Evaluation Data Randomly Selected Each Fold

CIVA

Tactical Computing Labs

Compiler Integrated Vulnerability Analyzer (CIVA)

- Is CIVA overfitting? No, 10 fold cross validation
- Cross fold validation randomly splits collection of programs into 3 partitions
  - Training - a subset of lines of code from programs are used
  - Testing & Evaluation - all lines of code from programs are used
- The training data set omits lines of code
  - Other cross validation folds can select omitted lines of code

CIVA

Tactical Computing Labs

Compiler Integrated Vulnerability Analyzer (CIVA)

- Performance metrics demonstrate CIVA's ability to characterize software
- CIVA can be retargeted/repurposed for a variety of applications
- CIVA can be retargeted for new programming languages (uncompiled software) and compiled program formats (byte code, machine code, etc)

CIVA

Tactical Computing Labs

CIVA is Aligned with the SoURCE CODE effort

- CIVA is a novel characterization technology
- CIVA performs a forensic task
- CIVA can be repurposed to identify coding styles
- CIVA can be extended to support binary/machine code and source code (programming languages)

CIVA

CIVA is Aligned with the SoURCE CODE effort

- CIVA can be extended to perform similarity detection from known samples
- CIVA can accelerate malicious attack attribution for both public and private sector responses


TCL