A Math Perspective on Fusion Needs

KEN JARMAN, NAT BEAGLEY, DALE HENDERSON, TIM WHITE

Pacific Northwest National Laboratory

ADSA08 Workshop, October 24-25, 2012

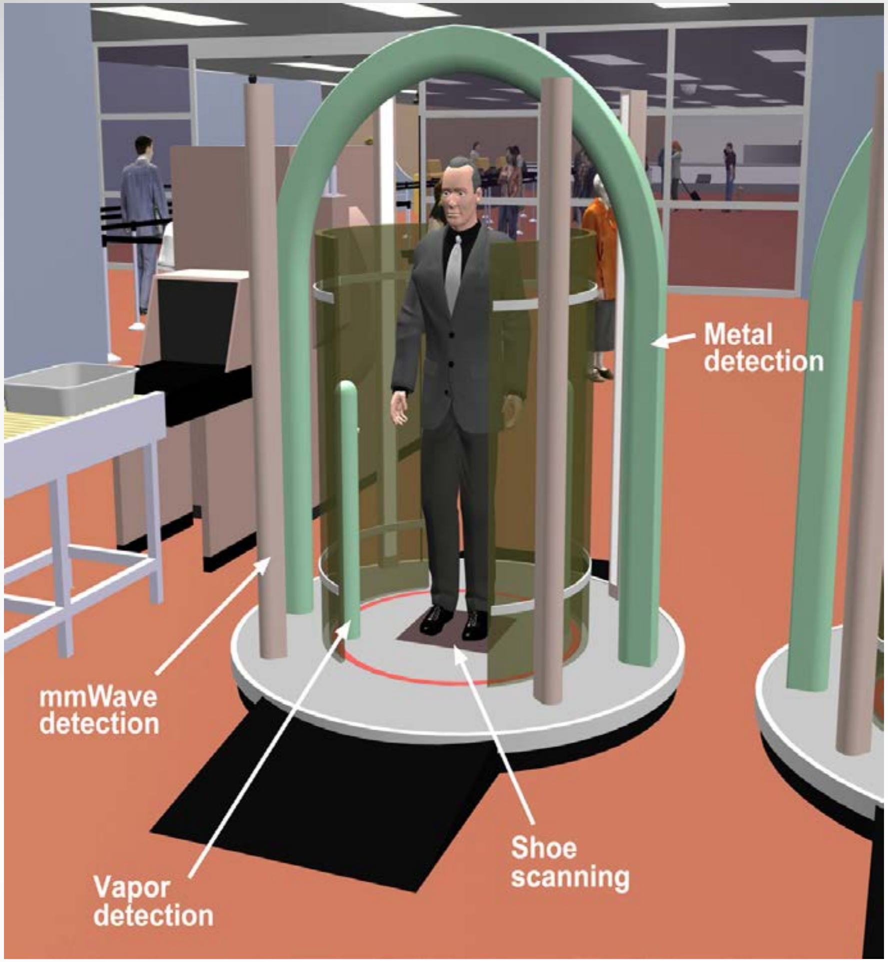
Conclusions

Pacific Northwest NATIONAL LABORATORY
Proudly Operated by **Battelle** *Since 1965*

- Need to study a variety of ATR fusion "models" (fuse at what step?)
  - "Deep" access to information produces better fusion—system developers need to study how much better, at what cost, what is feasible for specific system
- DHS S&T programmatic strategy is needed to evaluate and prioritize concepts for ATR fusion research investments
  - Define the **task**: problem space (threats, interferents, environments, …) and evaluation space (measures of performance and effectiveness, …)
  - Define standardized test scenarios and (large) data collections for fused system concept development, training, and evaluation
- DHS lab/industry/academia student incubators help solve "fusion challenge problems" with practical implications for explosives detection

Example Multi-Sensor System

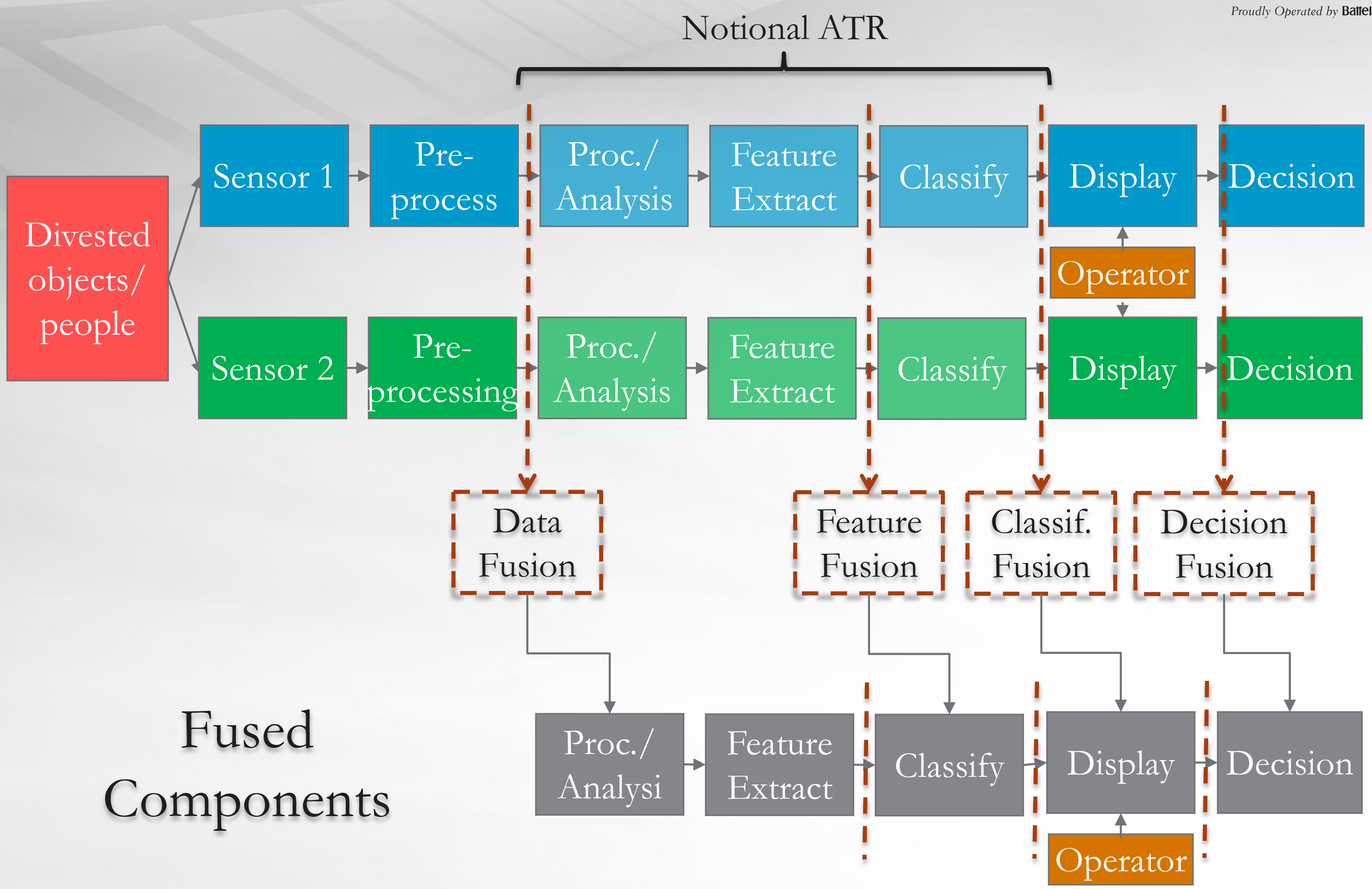
- Notional footprint-saving fusion example
- Consider mm-wave and metal detection
- Signatures
  - mm-wave: shape and dielectric constant
  - Metal detector: conductivity
- Task: detect explosives on person
  - Neither system directly sensitive to explosive material
  - Potential correlations in TP and FP spaces



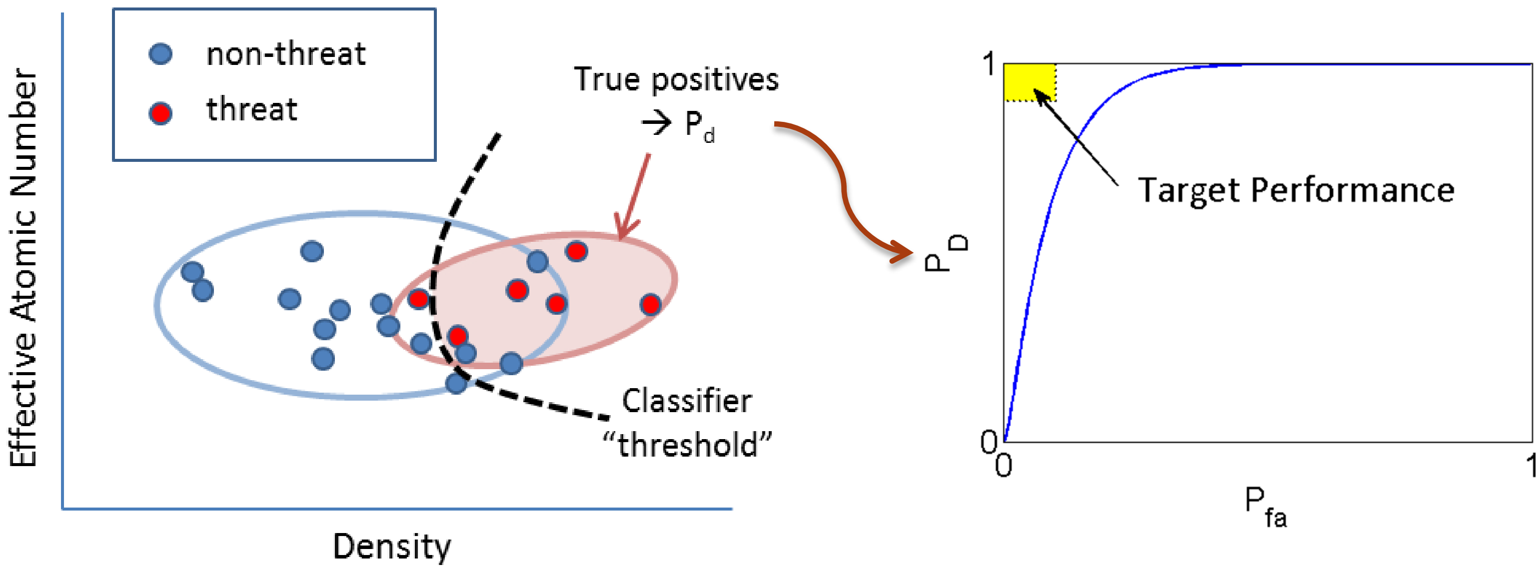
Modality	Plastic on Surface	Metal on Surface	Metal Below Surface
mm-wave	TP, FP	TP, FP	-
Metal Detector	-	TP, FP	FP

ATR and fusion – fuse at what step?

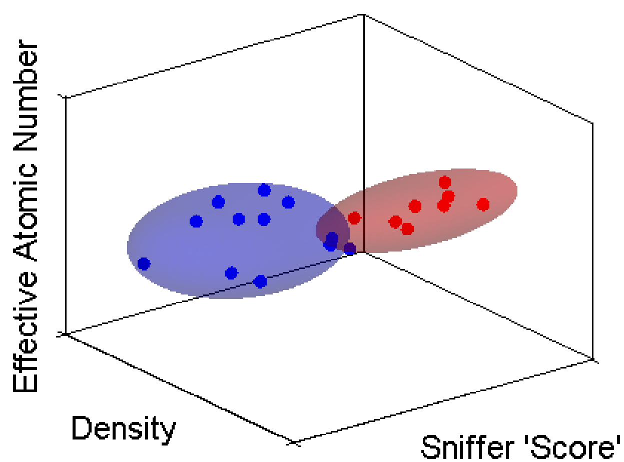
Pacific Northwest
NATIONAL LABORATORY
Proudly Operated by Battelle Since 1965



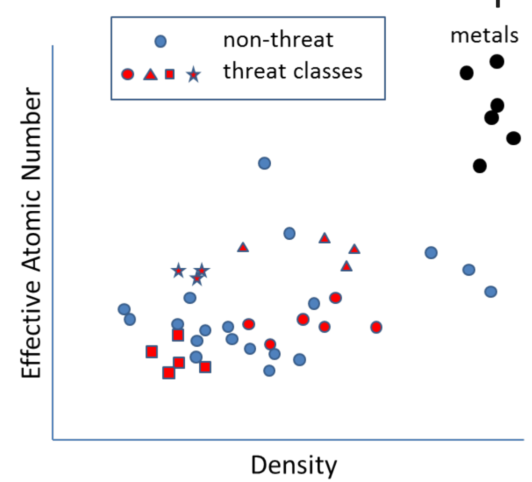
Detection Sensitivity and Specificity = Greater Separation in Feature Space



Increasing separation via complementary ("orthogonal") technology



More realistic classification problem



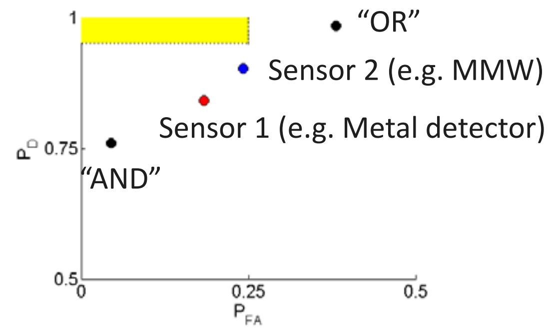
Good fusion needs “deep” info sharing

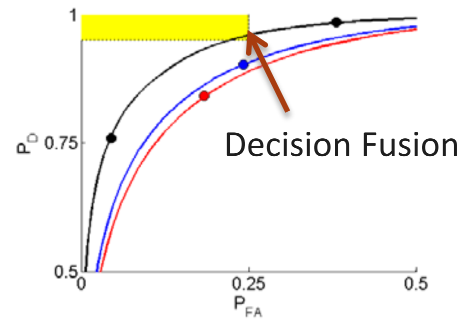
Increasing:

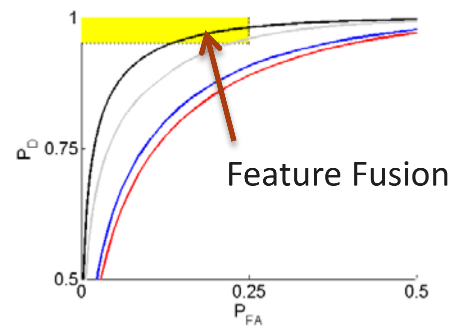
Information Sharing (and Info Security) --DICOS/DSFP?

Fusion Performance (and Fusion Complexity)

- Fixed P_D/P_{FA} for each sensor
- ROC curve for each sensor
- ROC curves plus correlation (modeled/estimated)
- Feature data/score, each sensor
- Feature data plus correlation (modeled/estimated)
- “Raw” data (only if fusion system developers are also experts at extracting features from the data)

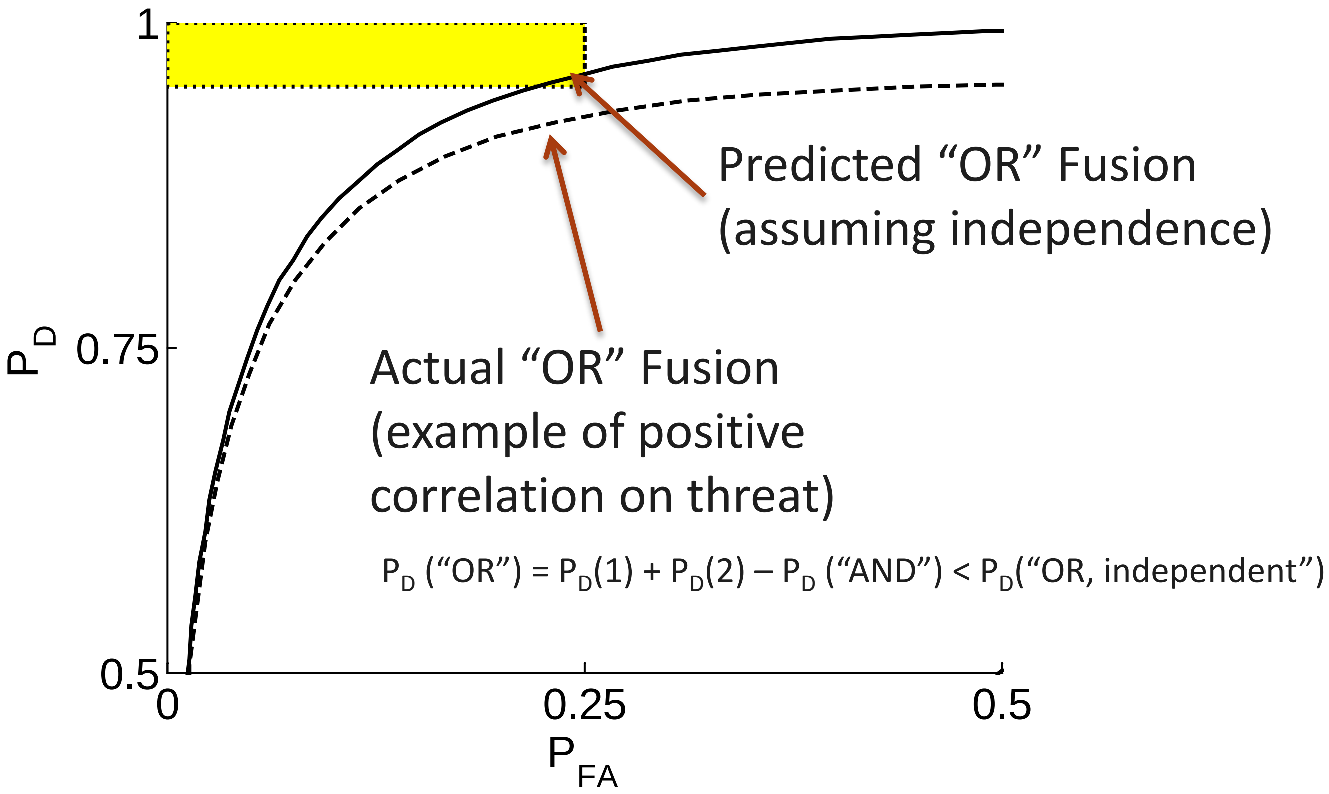






Good fusion needs "deep" info sharing

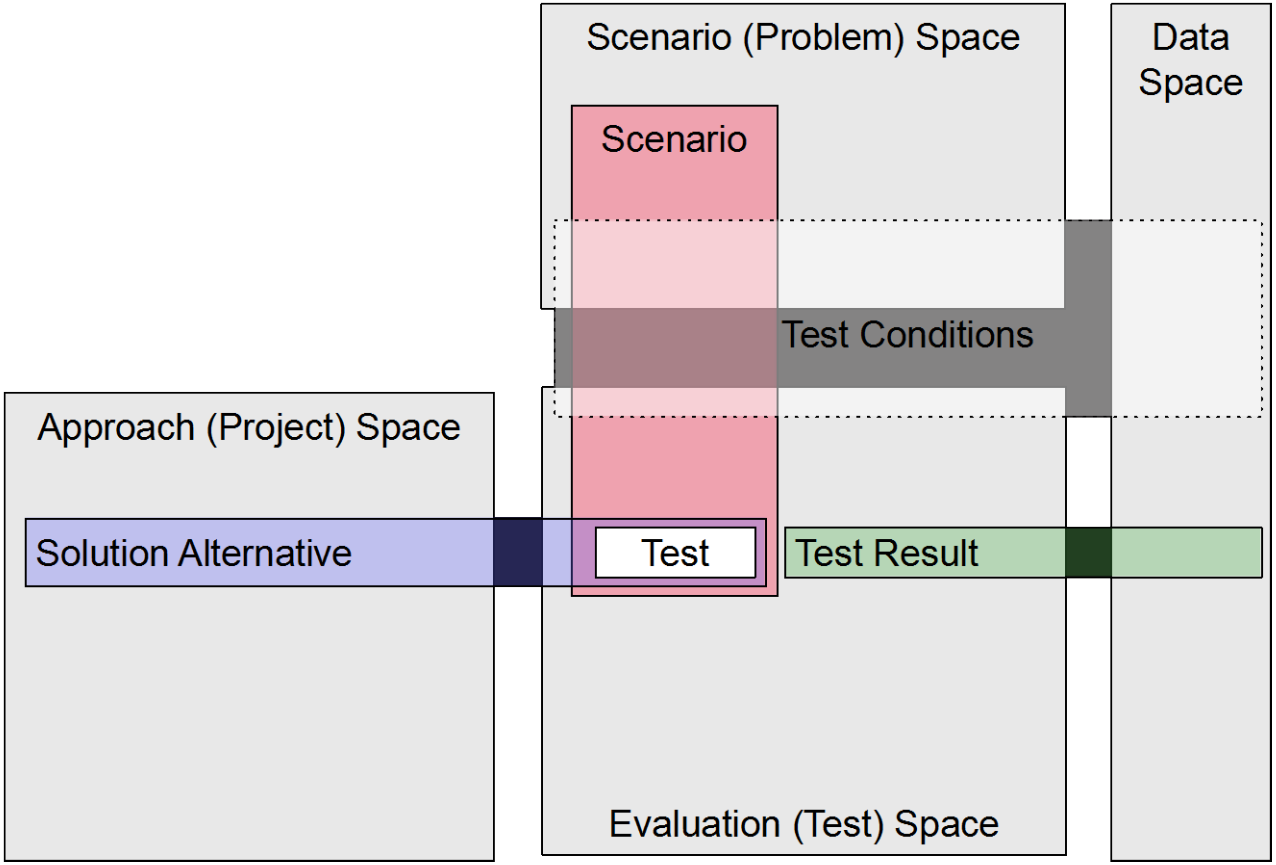
- So "ROC beats P_D/P_{FA}, features beat ROC, 'raw' data beats features" (maybe), and *neglecting potential correlation can lead to over-predicting (or under-predicting) performance*



Complementary Technology Programs

- Value:
  - The combination of signals through fusion algorithms or human interpretation can provide higher performance than the information provided by these signals taken independently.
- Conditions for Success:
  - Why and how is it anticipated that this solution will potentially improve system effectiveness?
  - Can the benefits of the solution be demonstrated on paper with synthetic or notional data against concrete measures?
  - How can we measure the impact of the solution in performance (MOP) and effectiveness (MOE)?
  - What is the TRL of this solution, and what is the plan to bring it to an operational level TRL?
  - What are the implications of the solution for the operational environment or under operational constraints?

Complementary Tech Program Spaces



Problem Space
- Threat & Materials
- Operational Conditions
- Scenarios & Test Conditions

Solution Space
- Approaches & Systems
- Orthogonal Concepts
- Prototype through Deployment

Evaluation Space
- Intersection between problem and solution
- Generates performance and effectiveness data

Data Space
- Test Results
- Test Cases

DHS S&T shares these spaces with strategic partners & key contributors. These spaces contain the elements of a research program strategy.

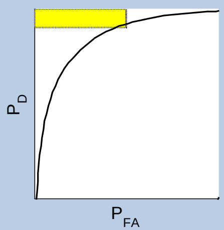
DHS/S&T EXD Student Incubator Projects 2012

A collection of projects focused on the mathematics of data fusion

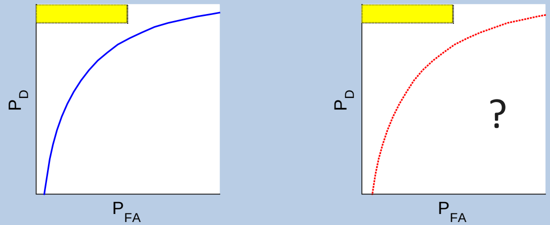
1) Alex Venzin
Mentor: Mark Oxley
Air Force Institute of Technology

ROC Curve Algebra
- Formal basis for augmenting a current system to achieve a desired system performance

Target fused performance

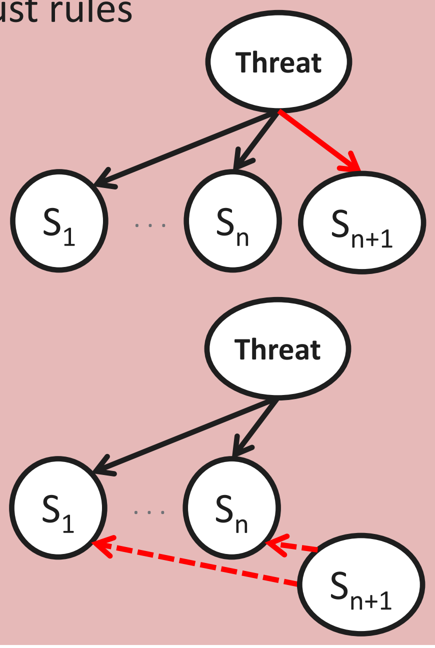


“minus” current → new sensor req



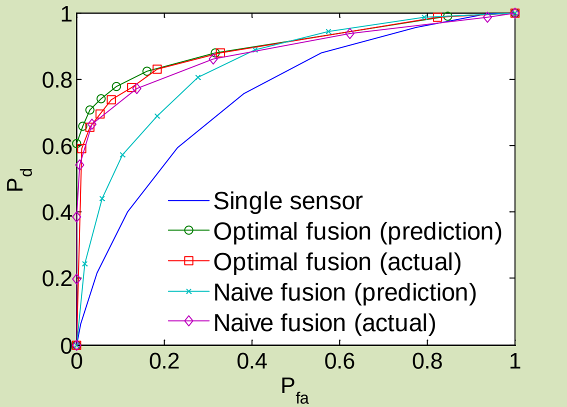
2) Matt Higger
Mentor: Deniz Erdogmus
Northeastern University

Fusion Robust to Sensor Failure
- Learning failed sensor characteristics; generating robust rules



3) Claire Longo
Mentor: Dale Henderson (PNNL)
University of New Mexico

Fusion sandbox library
- Numerical tool for exploring fusion concepts (incl. correlation and sensor failure)

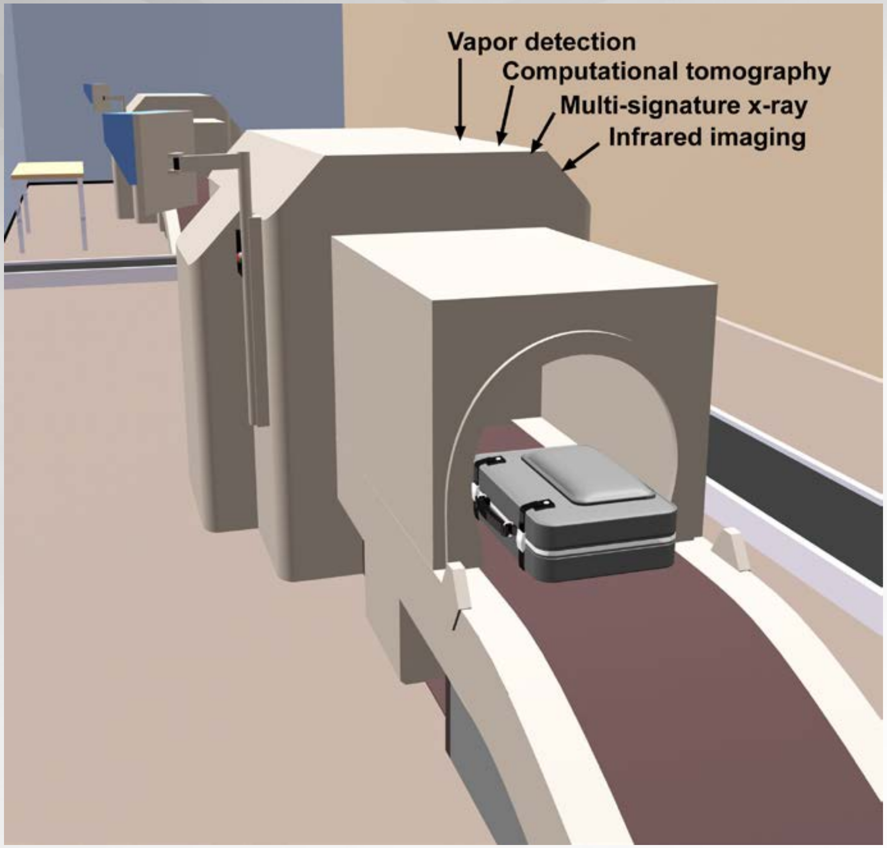


Conclusions

- Need to study a variety of ATR fusion "models" (fuse at what step?)
  - "Deep" access to information produces better fusion—system developers need to study how much better, at what cost, what is feasible for specific system
- DHS S&T programmatic strategy is needed to evaluate and prioritize concepts for ATR fusion research investments
  - Define the **task**: problem space (threats, interferents, environments, …) and evaluation space (measures of performance and effectiveness, …)
  - Define standardized test scenarios and (large) data collections for fused system concept development, training, and evaluation
- DHS lab/industry/academia student incubators help solve "fusion challenge problems" with practical implications for explosives detection

Additional Slides: OT Strategy

Example Multi-Sensor Systems



- Consider x-ray CT and IR imaging
- Signatures
  - CT: shape, density, Z_{eff}
  - IR: contamination of surface with explosive residue
- Task: detect explosives in bag
  - Presence of contamination may not be correlated with bulk explosives

modality	Bulk Explosive	Residue on Surface
X-ray CT	TP, FP	-
IR Imaging	-	TP, FP

Strategies for Complementary Technology

Pacific Northwest NATIONAL LABORATORY

Proudly Operated by **Battelle** *Since 1965*

Sponsored by DHS S&T Explosives Division

Focused on baggage and checkpoint screening for explosives

Objectives

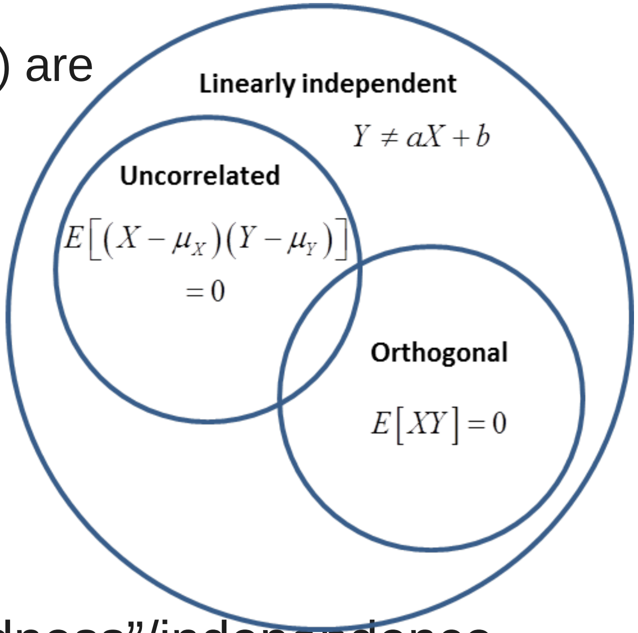
- Develop strategies for research in complementary technologies
  - Based on mathematical arguments and issues
  - Frame programmatic strategy for evaluating systems
- Initiate and oversee student "incubator" projects
- Outcomes: Briefing and two reports with recommendations pertinent to researchers, vendors, and funding agencies

Proposed Definitions

- A **signature** is a unique or distinguishing measurement, pattern or collection of information that indicates a phenomenon (e.g. object or event) of interest.
- A **technology** in this context is a practical application of knowledge, or a capability provided by such application of knowledge.
- A **sensor** is a type of technology that transmits information in response to a stimulus.
- **Fusion** in this context is the combination of output from multiple technologies to predict or estimate a potential threat state (e.g. presence of an object consistent with a type of threat).
- Technologies may be considered (partially) **complementary** if they either provide information related to different signatures of the same target object or are sensitive to different classes of target objects.

Orthogonality, Correlation, Independence

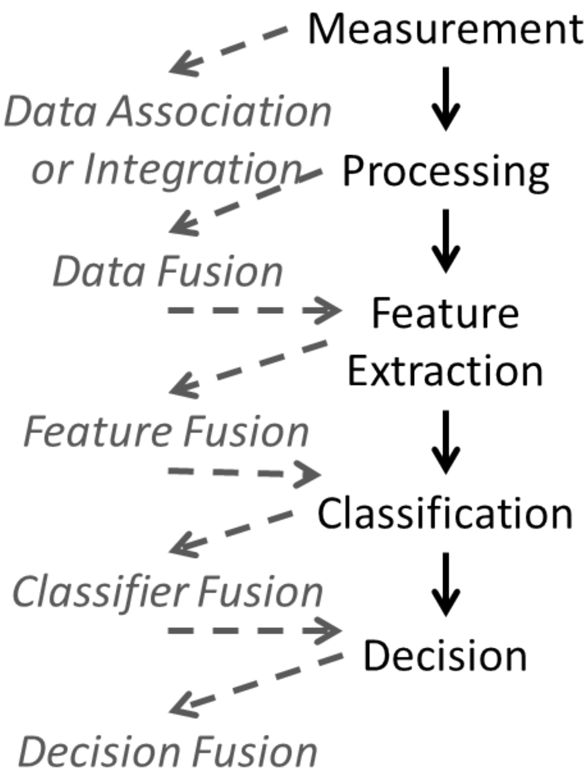
- *Mathematical* definitions:
  - Let X and Y be random variables (e.g. a spectral peak intensity from trace detection and density from CT)
  - Then X and Y (and corresponding technologies) are
    - Orthogonal if E[XY] = 0
    - Uncorrelated if E[XY] – E[X]E[Y] = 0
    - Linearly independent if Y ≠ a + bX for some scalar a, b
    - Independent if P[X<x and Y<y] = P[X<x]P[Y<y]



- BUT it's *conditional* orthogonality/"uncorrelatedness"/independence that concerns us
  - e.g. E[trace peak intensity × density | threat present] = 0

Fusion

- Categories
  - Combining sensor/classifier output directly
  - Primary/secondary
  - Adaptive; one sensor's output modifies operation or parameters of second
- Basic techniques
  - Heuristic/rule-based; voting
  - Pattern recognition
  - Bayesian, Dempster-Shafer, etc.
  - Hybrids
- Levels
  - ("raw") data fusion
  - Feature fusion
  - Classifier fusion
  - Decision fusion



Fusion Research and Data Needs

- DoD Wisdom
  - Fusion framework elements
  - Fusion methodology
  - Categorized “pitfalls”

Examples informing research/data needs

- Feature fusion beats decision fusion
- Ignoring (conditional) correlation is dangerous
- “Doubled” sensors provide a fusion performance baseline
  - Combining results of two “i.i.d.” sensors observing the same object improves performance, so any fused system should at least beat that (subject to cost, operational constraints)
- The “inverse” problem
  - Fusing current system with a new sensor, what new sensor performance is needed to boost from current system performance to a specified fused system performance
- The certification “gaming” problem
  - Achieving overall certification by gaming individual sensor performance