Technologies
Security & Detection Systems

Macro Security Revisited

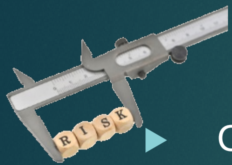
ANDREW FOLAND
ANDREW.FOLAND@L3T.COM

ADSA17 OCTOBER 17, 2017

Image: latimes.com

So What, Who Cares?

A. Foland



- Macro-Security system needs multiple system-level *lingua francas*
  - (Entity Assignment and Tracking, Threat, Policy, Decisional)
  - Informational *ℓf*: Probability updates put TSE's on common ground
- Common risk framework enables clear understanding of tradeoffs among cost, efficiency, and P_D
  - Passenger-level anomalies contain information: 2 TWL passengers on same flight is ~ as strong a signal as carrying a knife or other PI; 3 TWL in airport
- Certification procedure must reflect system-level priorities as much as TSE-level priorities
  - ROC curves vs. operating points
  - Rapid-response
- Crucial role of  flowing back from airports to system / TSE providers to utilize information

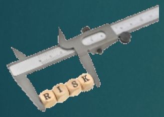
  - Create nonthreat model
  - Spot anomalies
  - Improve discrimination
- Whole-system design, with strong central leadership, will achieve cost and operational efficiency at system level; can be approached in steps

Talk motifs: data feedback from airports; consistently quantitative risk assesment

High-Level Goals Are Simply Stateable

A. Foland

- Move *X* passengers and belongings per hour across a security perimeter
- In a footprint of size *Y*
- Subject to
  - Constraint: cost / passenger < $C_{acceptable}$
  - Constraint: P(threat event)<$P_{acceptable}$
  - Soft Constraint: passenger experience



C_{today}~\$3.25 / passenger

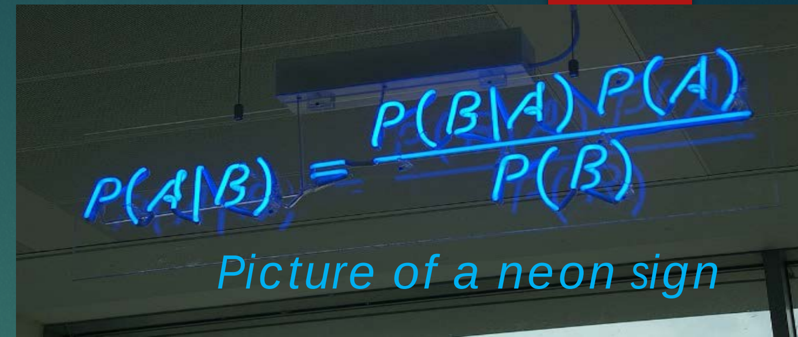
Talk uses mostly checkpoint for examples, but methods extend to checked bags

MacroSecurity is an Informational Approach

- Make best use of all available information
- Better info → fewer FA, higher PD → higher throughput, more targeted secondary inspection → better passenger experience, lower costs, better security
  - Clarify and motivate tradeoffs
- How to distribute limited resources to maximally cover the possibilities
  - Limited resources include passenger time and goodwill
- Accept that there is such a thing as $P_{acceptable}$
  - Practical range1E-10 to 3E-12
    - Low end is 1 bad event per 100 years of world air traffic volume
    - Comparing aviation today to 100 years ago, it will be completely different by then
      - Good odds that no events happen in current-era aviation

L3 Technologies
Security & Detection Systems

Require ℓf for System P(event)



- Initial estimate of P(event) at customer checkin
- Update P(event) at every data acquisition
  - Comparison to threat lists
  - Behavioral tracking
  - Bag scans
  - Body scans
  - Secondary screens
  - Tertiary screens
  - LEO actions

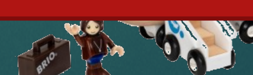
A probability lingua franca puts these on the same footing in a common system

- Continue acquisitions until one of
  - $P(\text{event}) < P_{acceptable}$
    - 1E-11?
  - $P(\text{event}) > P_{unacceptable}$
    - ~5E-7?
  - Cost > $C_{acceptable}$
    - Smallish multiple of $3.25?
  - No more data will be available

Passenger Checks In

Assign P_0(event)

Feature Vector
Probability
Classification Label



Combine Information for Low P_{FA}

$$f_{External} \stackrel{\text{def}}{=} r = \begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_{Nr} \end{pmatrix}$$

A. Foland

$$f_{X-ray} \stackrel{\text{def}}{=} x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{Nx} \end{pmatrix}$$

Decision Matrix	External Classification		
X-Ray Classification		Threat	Not-Threat
	Threat	Threat!	???
	Not-Threat	???	Not a Threat!

Combining Classification

$$f_{system} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{Nx} \\ r_1 \\ r_2 \\ \vdots \\ r_{Nr} \end{pmatrix}$$

System Classification

Threat

Not Threat

Combining Features

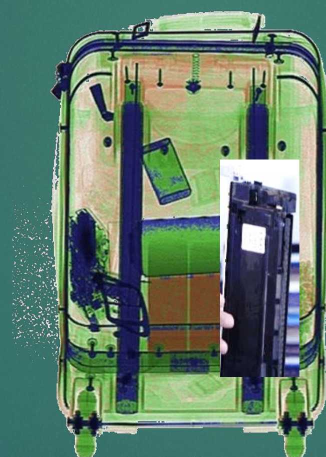
A. Foland

Not All Alarms are Created Equal

- Want TSE to report these three cases differently

wired.com

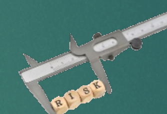
wired.com

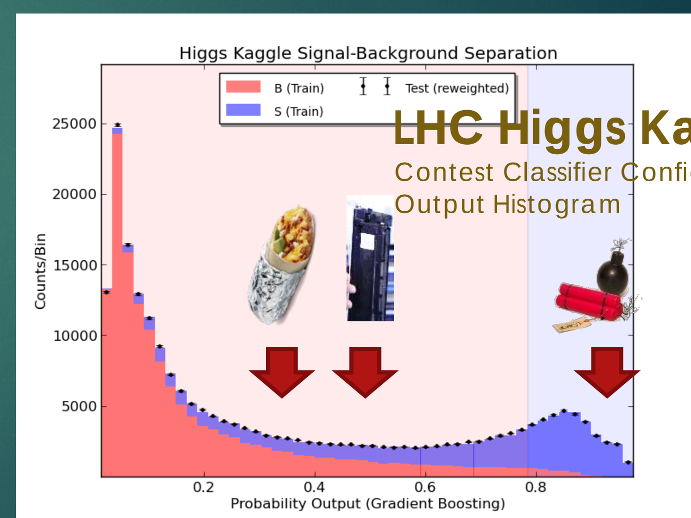


wired.com

- Today, generally report 1 bit of information (0 or 1, Clear or Alarm)

- TSE reports classification and confidence
  - Softmax over multiple classifications?
    - Including
      - "Nothing of Interest"
      - "I don't usually see this"





LHC Higgs Kaggle

Contest Classifier Confidence Output Histogram

Red: Confusants (~FA)

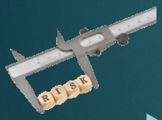
Blue: Signal to be detected

Macro Security Requires a Threat Model

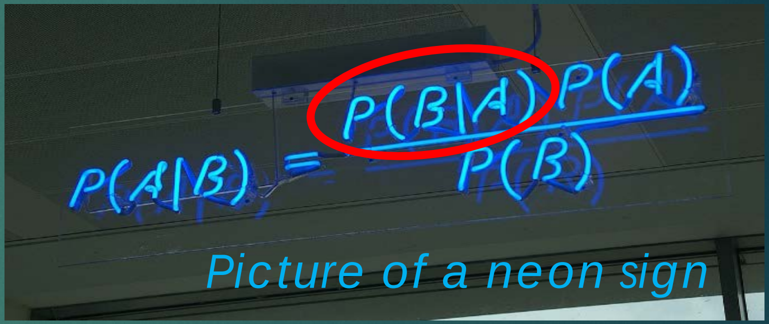
A. Foland

- Systematic approach requires estimates of
  - P(Detector Result | Threat)
  - For instance
    - Probability that a bad actor will have a prohibited item detected in their baggage
    - Probability that a bad actor will take an extra long time to get from check-in to security
    - *Probability that a bad actor will check-in onto the same flight as a separate high-threat-category passenger*

- *Crude models are numerically valuable*
  - Can baby-step to best models
- Model owned outside of TSE's
  - Best performance requires significant input from real-world data
  - Real-world data must be coupled with reasonable but numerically-explicit assumptions
  - TSE's report classifications and confidence (the detector findings); model turns those into probability updates



- *Existing security system today already makes such assumptions*
  - Implicitly rather than explicitly



Picture of a neon sign

The Role of Anomalies

A. Foland

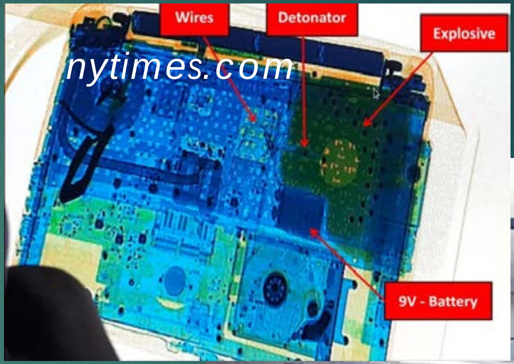
- It is to be expected that $P(\text{Anomaly} \mid \text{Threat}) >> P(\text{Anomaly})$
- System can in principle be set by fiat such that
  - Sufficiently anomalous observations are assigned to an anomaly category
  - “Sufficiently” anomalous can be defined as inducing an FA rate that is not operationally burdensome
  - “Anomaly” category model of $\frac{P(Anomaly \mid Event)}{P(Anomaly)}$ set high enough to trigger $P_{unacceptable}$ for most or all categories of passengers

- Challenge
  - Need sufficient data from airports for TSE’s to be able to recognize “I don’t normally see something like this”
  - Need protocol for TSE’s to report anomalous observation

Rapid Response

A. Foland

- Three Options For Responding to Events
  
  - Change initial P(threat) for some or all passenger classes
  - Add new detection actions to decision tree
    - "Is there a laptop?"
  
  - Change the Event Model
    - Increase P(Hat | Threat)



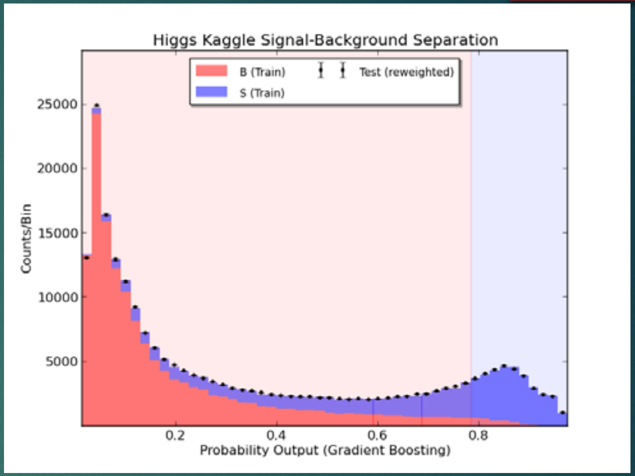
independent.co.uk

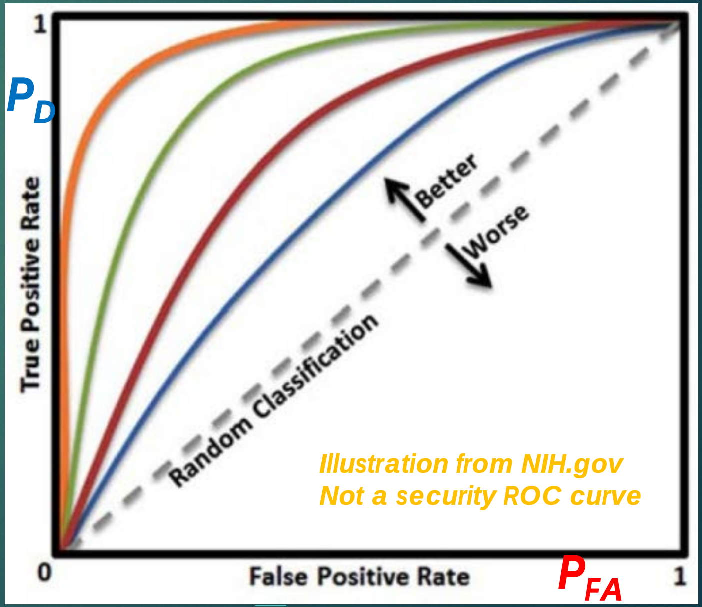
Implications for Certification

- Vendor strategy is driven by certification
- Explore ROC-curve based model?

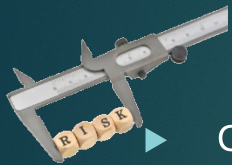
  - Algorithm outputs category-confidence values, not alarm/clear binary values
    - Current EDS cert effectively corresponds to one category
    - Internally to TSL: characterize P_D/P_{FA} at each threshold of confidence value
    - If there exists any threshold for which P_D/P_{FA} pass current cert requirements
      - Set the operating threshold in passing region, the machine is certified to current standards
- As standards evolve

  - Option to vary sensitivity / P_{FA} continuously
  - Add new category classifiers as needed to already-certified machines
    - Replay test to evaluate P_{FA} impact
- Balance rapid feedback with (appropriate) concerns about test-set transparency





So What, Who Cares?



- Macro-Security system needs multiple system-level *lingua francas*
  - (Entity Assignment and Tracking, Threat, Policy, Decisional)
  - Informational ℓf: Probability updates put TSE's on common ground
- Common risk framework enables clear understanding of tradeoffs among cost, efficiency, and P_D
  - Passenger-level anomalies contain information: 2 TWL passengers on same flight is ~ as strong a signal as carrying a knife or other PI
- Certification procedure must reflect system-level priorities as much as TSE-level priorities
  - ROC curves vs. operating points
  - Rapid-response
- Crucial role of  flowing back from airports to system / TSE providers to utilize information
  - Create nonthreat model
  - Spot anomalies
  - Improve discrimination
- Whole-system design, with strong central leadership, will achieve cost and operational efficiency at system level; can be approached in steps

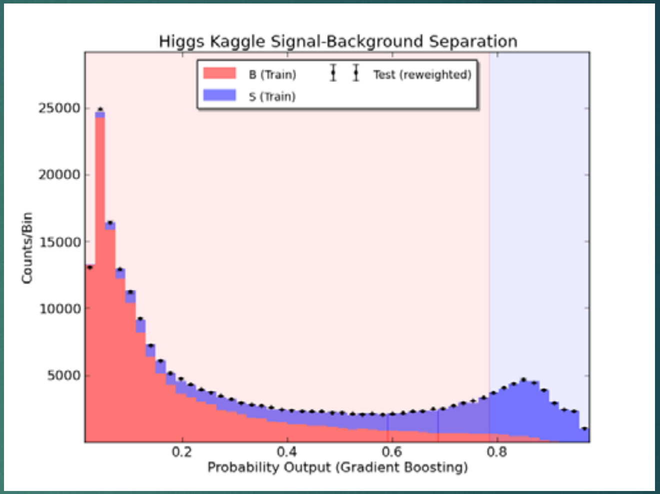
Talk motifs: data feedback from airports; consistently quantitative risk assesment

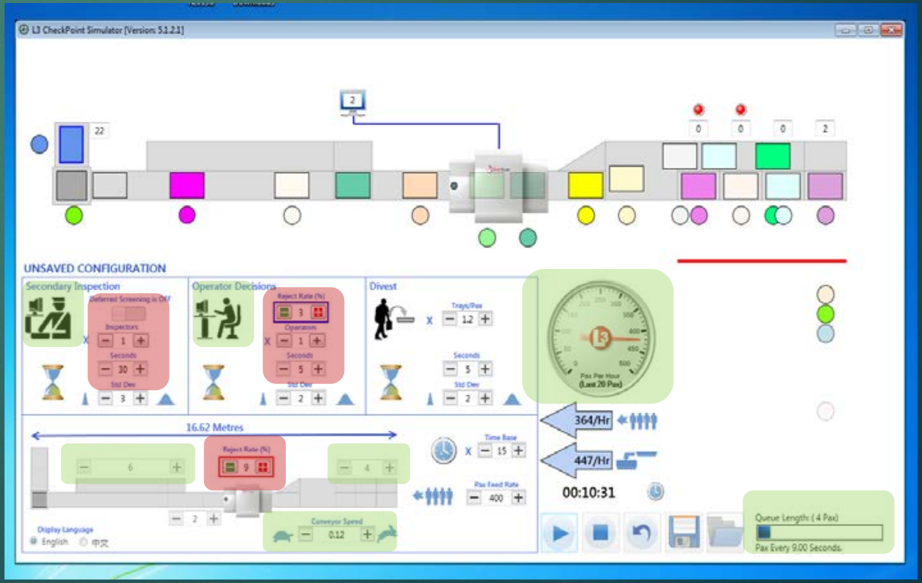
Cost and Throughput Depend on Discrimination

A. Foland

- *P*(threat) informational model depends on outcomes of ordered measurements
- Costs of the system depend on ability of each measurement to improve knowledge (i.e. its discrimination)
  - Related to, but not quite the same, as P_D/P_{FA}
  - More closely related to ROC curve

- Throughput of system depends on action of decision tree under normal conditions
  - Closely related to P_{FA}
- Cost, throughput, and security all depend on the discrimination of individual TSE's
  - Often in non-obvious ways
- This is where commonality and clarity will pay off
  - Can model tradeoffs in the MacroSecurity decision tree





P_0

1E-12 1E-11 1E-10 1E-7 1E-2

A. Foland

- BACKUPS

Some Nice Round Numbers

- 3.5B passenger-flights per annum in world
- 1B passenger-flights per annum in USA
- Following taken from Wikipedia rounded to one significant digit
- ~2000 guns / year in USA
- ~100,000 prohibited items / year in USA
- ~1000 Americans on no-fly list
- ~20000 non-Americans on no-fly list
- ~100,000 Americans on “terrorist watch list”
- ~2M non-Americans on TWL

- Some Calculations
- Averaged over all flights of last 10 years in USA
  - P(event)<1E-10
- Assuming P(PI present | threat)=0.5
  - Update factor for finding PI is ~5000 (= 0.5 / (100000/1B))
  - P(updated)=5E-7
- Assuming P(gun | threat) = 0.1
  - Update factor for finding gun is ~50000 (= 0.1 / (2000/1B))
  - P(threat updated)=5E-6
- Assuming P(On TWL | threat)=0.2
  - Initial P for TWL should be 6E-8=1E-10*0.2/(100k/300M)
- Assuming P(Comrade On TWL | threat)=0.2
  - Updated P after finding a second person on same flight on TWL: 4E-7

L3 Technologies
Security & Detection Systems

Strong Centralization Required

- Challenges not addressed here
  - Entity tracking
  - Networking of TSE’s
  - Intelligence Input
  - New hardware for passenger ID, tracking
- Define protocols for
  - Initial passenger assignment (communication with external databases)
  - Detector networking and reporting to system
  - Intelligence input to system threat model
  - Defining a measurement decision tree
  - Updating probability estimates
  - Certification of TSE’s
- Define and own
  - Measurement set
  - Decision tree
  - Threat model

Eric Houser, TSA ORCA Deep Learning Exhcange, 8/17

Aviation Security Evolution

Aviation security must evolve to effectively and efficiently support a higher commercial demand while detecting a wider range of threats.

Present	Future
Transportation Security Officers (TSO) review x-ray images of every carry-on bag	Enhanced Automated Threat Recognition (ATR) of explosives, weapons, and contraband
Passengers divest liquids, aerosols, gels (LAG), laptops, bulky outer garments, and shoes	Minimal divestiture of LAG, laptops, and clothing increases throughput
Passengers stop and pose for Advanced Imaging Technology (AIT)	Passengers move through checkpoint at a walking pace in parallel with carry-on items
High system Probability of False Alarm (Pfa) leads to labor intensive screening/reduced throughput	Reduced Pfa to increase screening efficiency
Passengers are screened at Standard or Pre✓ Lanes	Risk Based Security (RBS) enables dynamic screening, more efficient allocation of resources
Transportation Security Equipment (TSE) software, algorithms, and data managed locally	TSE securely networked and communicating via Security Technology Integrated Program (STIP)
Variation among TSE user interfaces increase complexity and training requirements	Common Graphical User Interface (GUI) yields consistent user experience across TSE fleet
Unique TSE designs and interfaces result in long capability development lead times	Open Architecture and Application Program Interfaces (APIs) enable modular "plug and play"

Keith Goll, TSA ORCA Deep Learning Exchange, 8/17

The Benefits of System Architecture

SA supports both TSA and the industry by developing innovative solutions, resulting in the following benefits:

Enables Modularity

Introduces modular components by **defining system infrastructure and interfaces** enabling **plug-&-play functionality** and increasing **system flexibility**

Enhances Innovation

Drives standardization and modularity **to foster greater competition** at sub-system levels**, expand industry base,** and **reward modular implementation** via incentive-based procurement

Advances Risk-based Security

Enables RBS by developing a **common data model** and the **infrastructure** required for the **masking of sensitive information and use of threat data** to expedite the screening process

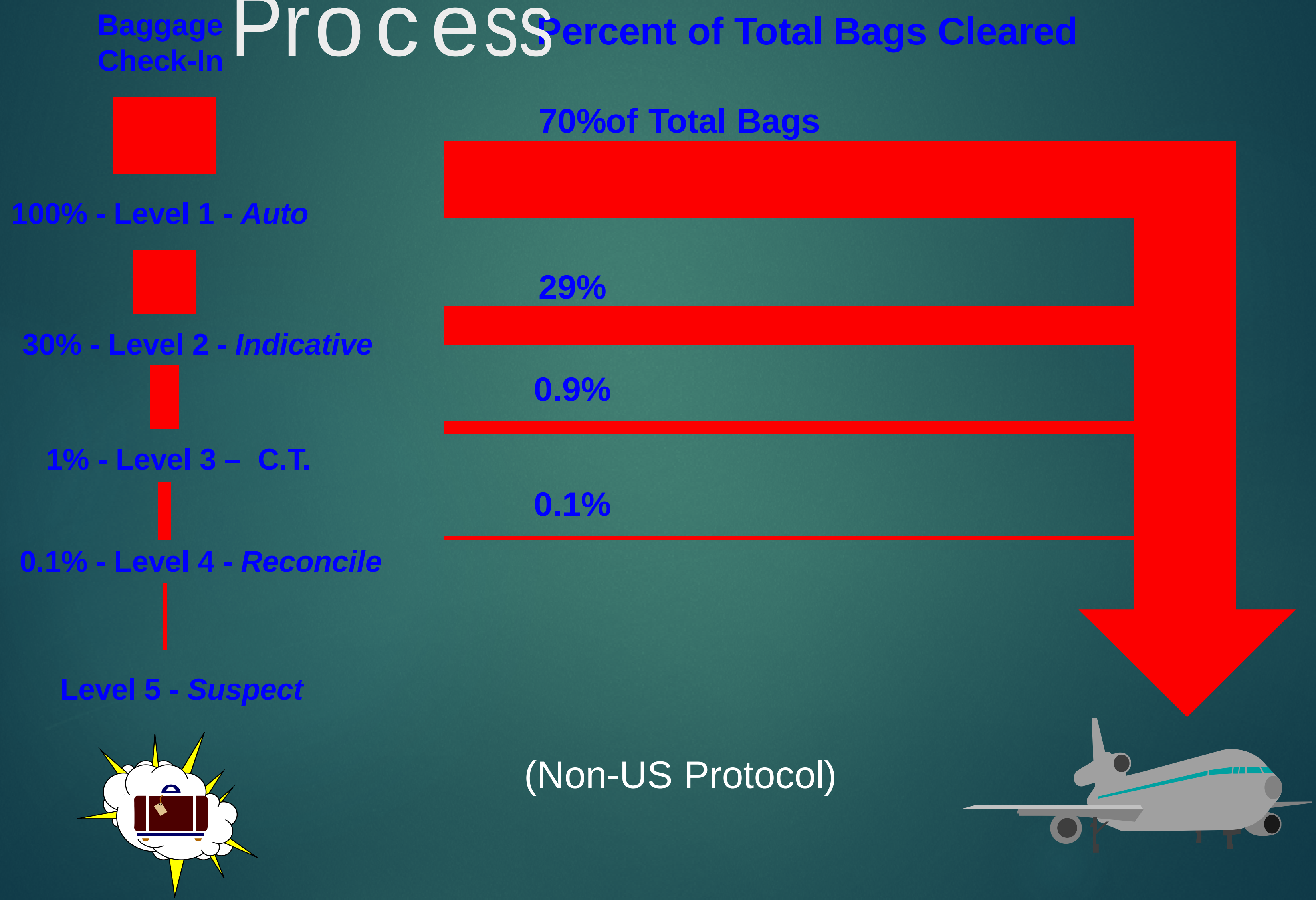
Reduces Costs

Promotes interoperability and incremental upgrades to reduce **duplicative development** and **testing requirements**

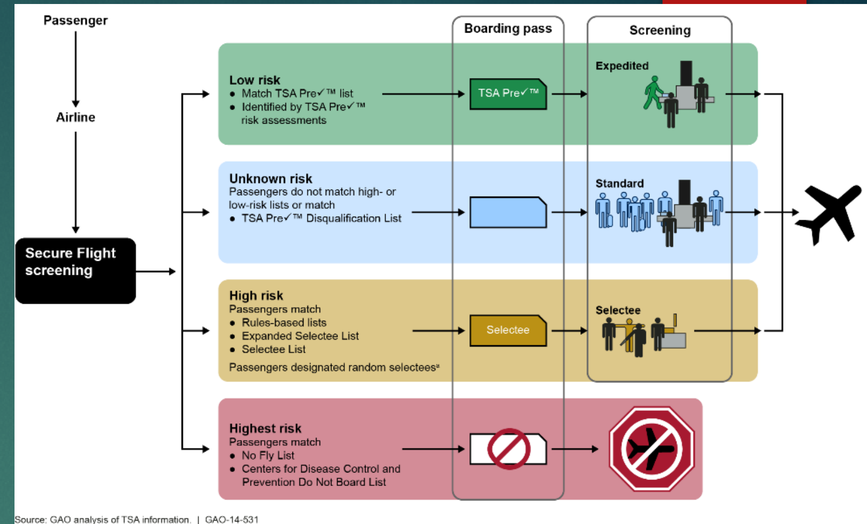
Expedites Delivery of Capabilities

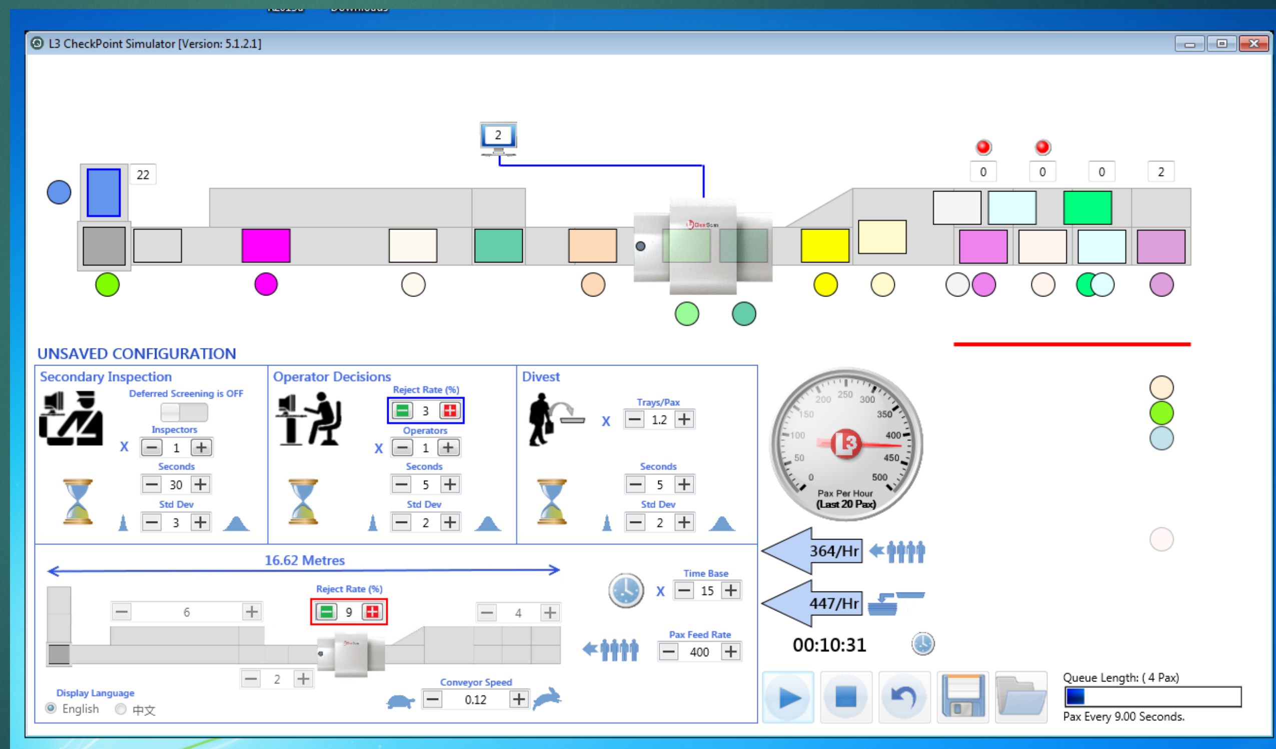
Reduces the timespan between the **inception and delivery of a capability** by providing vendors with **well-defined open standards**

Multi-Level Screening Process



Whole System Design Is Required





Whole System Design Is Required

- Every little decision has impacts throughout system
  - ASL vs Standalone
  - How many divest stations
  - How deep a secondary queue
  - How long operator review is
  - Reconstitution
  - Ratio of secondary : primary
- Holistic design only possible with strong centralizers
  - ROC curves
  - P(threat | what's known)
  - Replay / rapid deployment

Certification Is Central To Development Strategy

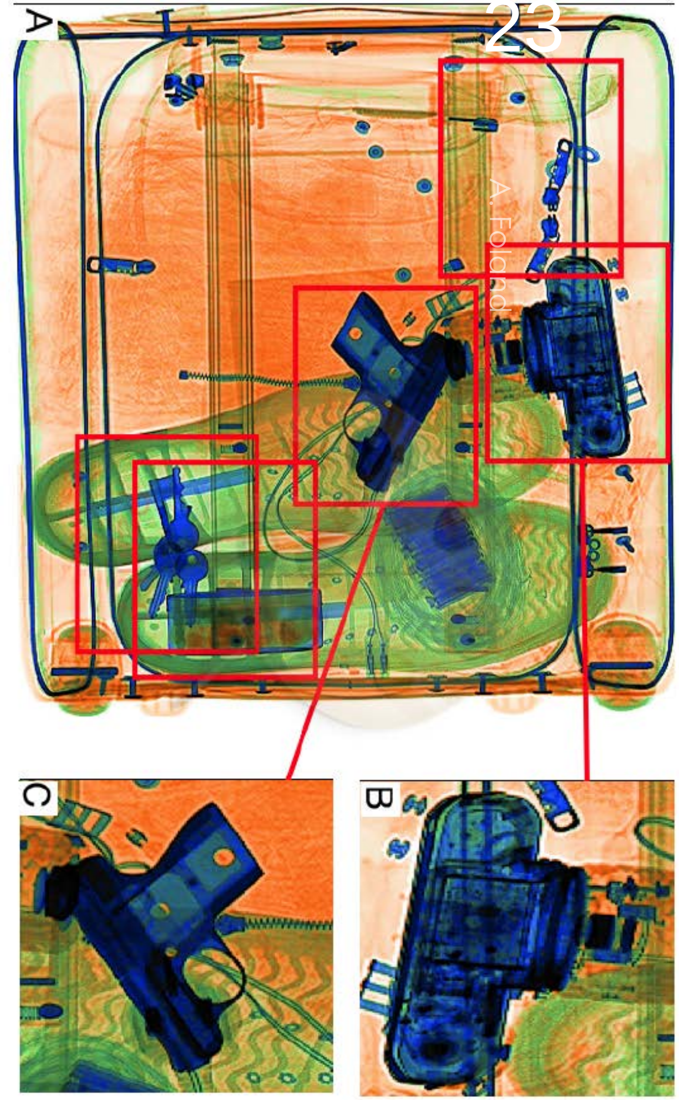
A. Foland

- Machine development NRE is a risk by vendors of tens of millions of dollars
- Development decisions, design decisions, and roadmaps are driven in large part by requirement to achieve certification
- Any significant shifts to TSA development thinking must be accompanied by "what (if any) changes to certification procedure are required by this shift?"
  - Otherwise, unintended consequences
- Replay / rapid deployment

We Can Better Use Operators' Time By Reducing Cognitive Load

ocregister.com

[Akcay et al, 2016]



A. Foldes

L3 Technologies

Transmission X-ray Rules for a Reason

- Cheap, fast, effective
- Can leverage off medical experience
- Engineering highly optimized
- Easy to train humans to use
- From security point of view: the easy part of the 80/20 tradeoff
  - Corollary of 80/20: progress from here costs 16x per unit of performance

1897

Operational Costs

A. Foland

- From testimony on FY17 budget
  - https://www.tsa.gov/news/testimony/2016/03/01/hearing-fy17-budget-request-transportation-security-administration
  - $3.1B in operational expenses related to TSO activities
  - $200M in equipment expenses
- 949M passengers annually
- Broadly: the US spends about $3.25 in operational costs per passenger
  - About $0.21 in equipment
- European cost models are broadly similar

Cost Models

- Broadly: the US spends about $3.25 in operational costs per passenger
  - Dominant costs are
    - Threat review by operators at checkpoint
    - Secondary resolution of false alarms at checkpoint
  - Lesser costs are
    - Secondary resolution of false alarms in checked bags
    - Tertiary+ resolution of false alarms in checked bags
- About $0.21 in equipment

Throughput Models

A. Foland

- There's more to throughput than belt speed
- All processing systems reach an equilibrium where they are gated by the slowest throughput stream
- In airport checkpoints today
  - Near tie between primary review and secondary resolution
    - Both much slower than scanner throughputs
- Many ways to address
  - Parallelize primary review
  - Speed up secondary resolution
  - Parallelize secondary resolution

System-Level Design

- All processing systems reach an equilibrium where they are gated by the slowest throughput stream
  - Cannot buy TSE's in isolation
  - Intelligent flow, fan-in/fan-out, throughput matching required to get smooth system at peak input