Optimization of Containers Inspection at Port-of-Entry

E. A. Elsayed, Christina Young, Yada Zhu
Department of Industrial and Systems Engineering

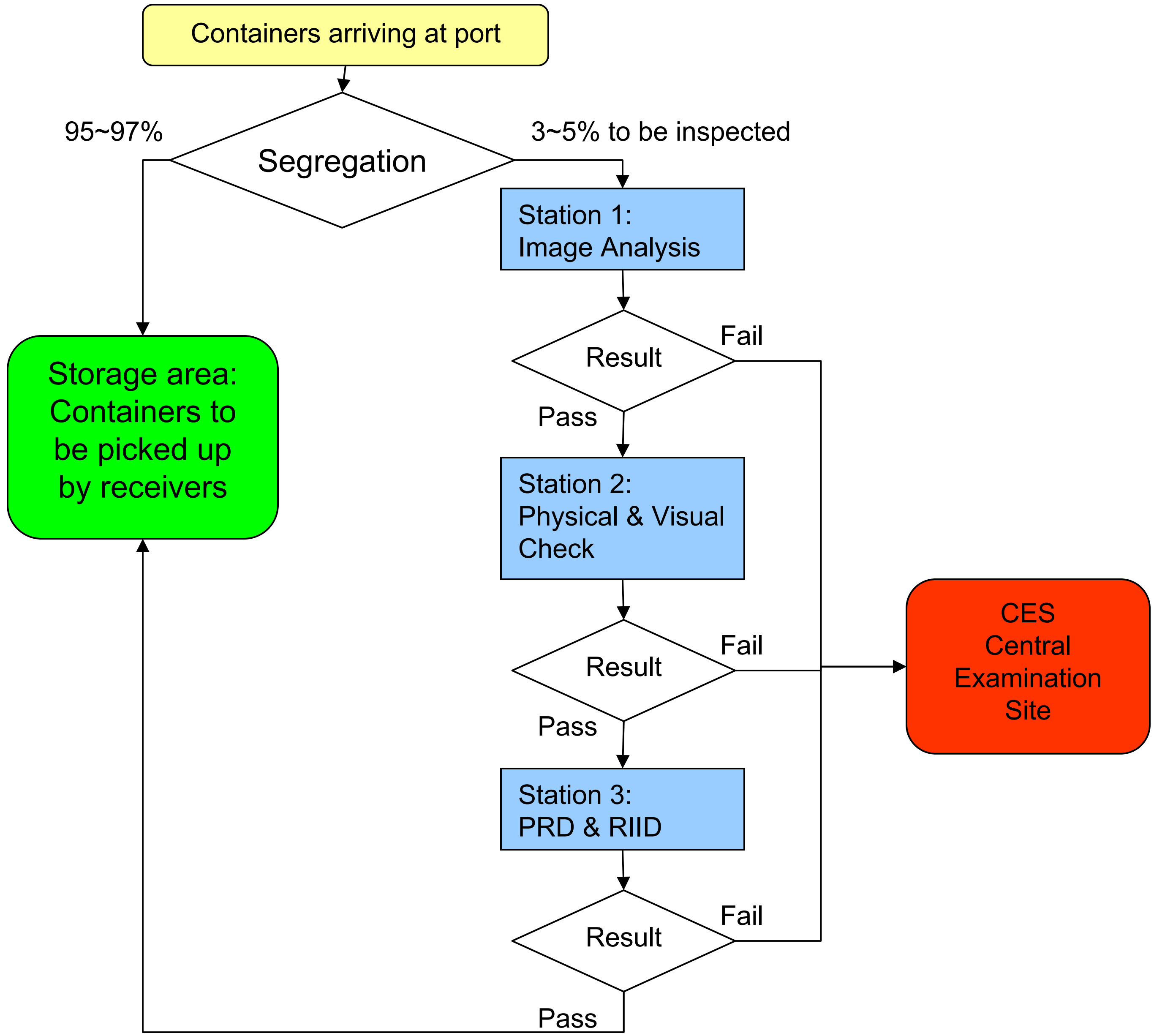
Mingyu Li, Minge Xie
Department of Statistics

Rutgers University

Introduction

- Containers arrive at a port-of-entry for inspection
    - *n* attributes tested independently
- Sensors used to classify attributes of a container as safe (d=0) or suspicious (d=1) based on selected threshold values (T_i for station i)
- Overall accept/reject decision based on specified Boolean function of station decisions

Simple Container Inspection Procedure



Inspection Problem Description

- Threshold levels affect the decisions and probabilities of misclassification at each station
- Sequence of inspection stations affects the expected cost and time of inspection per container
- Inspection policy specifies sequence (S) and T_i values

Objective

Minimize expected cost of inspection, cost of container misclassifications, and expected inspection time

Decision Variables

Sensor threshold values and sequence of stations

Modeling Approach

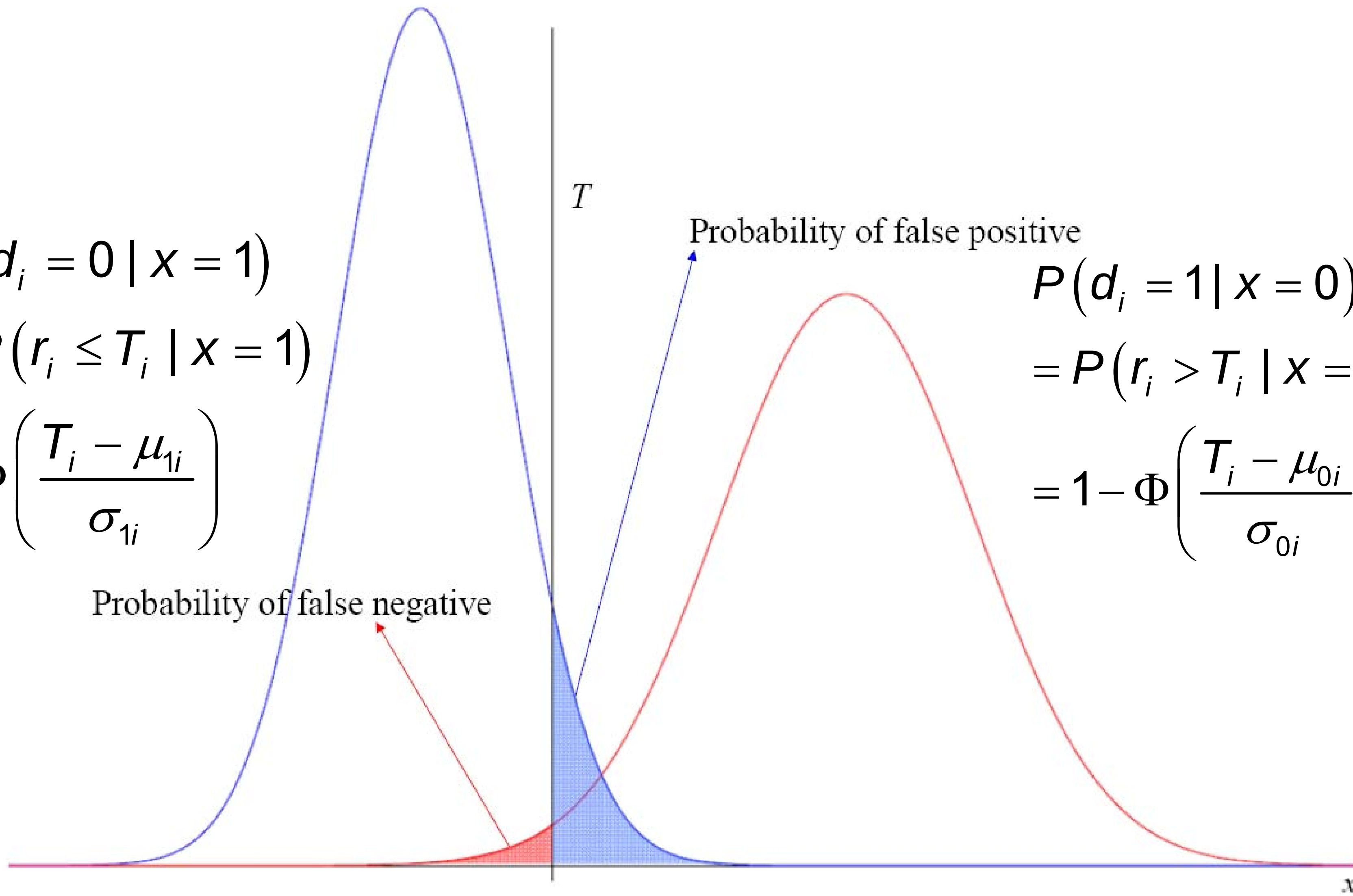
- Assume the unit's true state (*x*) is 0 or 1
- A container attribute value comes from a mixture of two normal distributions, depending on the unit's true state
- Let *r* represent the sensor reading returned from a unit, assumed equal to the attribute value
  - $r \mid x = 0 \sim Norm(\mu_0, \sigma_0^2)$
  - $r \mid x = 1 \sim Norm(\mu_1, \sigma_1^2)$
- For some value T_i at station *i*
  - If measurement r_i > threshold level (T_i), then decision for station *i*, $d_i = 1$
  - If $r_i \leq T_i$, $d_i = 0$

Modeling Approach

- Assume prior distribution of suspicious containers is known: $\pi = P(x=1) = 1 - P(x=0)$
- Subscript *i* is used to indicate association with station *i*
- Assume parameters of the two normal distributions are known: μ_{0i} , μ_{1i} , σ_{0i} , σ_{1i}

Probabilities of Error

$$P\left(d_i = 0 \mid x = 1\right) = P\left(r_i \le T_i \mid x = 1\right) = \Phi\left(\frac{T_i - \mu_{1i}}{\sigma_{1i}}\right)$$



$$P\left(d_i = 1 \mid x = 0\right) = P\left(r_i > T_i \mid x = 0\right) = 1 - \Phi\left(\frac{T_i - \mu_{0i}}{\sigma_{0i}}\right)$$

Cost of Misclassification

- The individual results of stations are combined according to defined system Boolean function to reach an overall inspection decision to accept or reject a container
- This system decision *D* may or may not agree with the container's true status
- Probability of false accept, $PFA = P(D = 0 \mid x = 1)$
- Probability of false reject, $PFR = P(D = 1 \mid x = 0)$
- Two sources of container misclassification cost:
  - c_{FA} = cost of false acceptance (undesired cargo)
  - c_{FR} = cost of false rejection (manual unpack)
- Total cost: $C_F = \pi \; PFA \; c_{FA} + (1 - \pi) PFR \; c_{FR}$

Cost of Inspection

- c_i = cost of using i^{th} sensor
- Expected cost of inspection depends on probability of passing (or failing) sensor i :
- $$p_i = P(d_i = 0) = P(d_i = 0 \mid x = 0)(1-\pi) + P(d_i = 0 \mid x = 1)\pi$$
$$= (1-\pi)\Phi\left(\frac{T_i - \mu_{0i}}{\sigma_{0i}}\right) + \pi\Phi\left(\frac{T_i - \mu_{1i}}{\sigma_{1i}}\right)$$
- $$q_i = 1 - p_i = P(d_i = 1 \mid x = 0)(1-\pi) + P(d_i = 1 \mid x = 1)\pi$$
$$= (1-\pi)\left\{1 - \Phi\left(\frac{T_i - \mu_{0i}}{\sigma_{0i}}\right)\right\} + \pi\left\{1 - \Phi\left(\frac{T_i - \mu_{1i}}{\sigma_{1i}}\right)\right\}$$

Optimum Inspection Sequence

- The sequence in which stations are visited affects the expected cost of inspection
  - A sequence which minimizes this cost is known as an optimum sequence
- ***Theorem 1:*** For a ***series*** Boolean decision function, inspecting attributes $i,\ i = 1,2,\ldots,n$ in sequential order is optimum (minimizes expected inspection cost) if and only if: $c_1/q_1 \le c_2/q_2 \le \ldots c_n/q_n$.
- A container is suspicious if decision for any station *i is 1. In other words:*

$$F(d_1, d_2, \ldots, d_n) = \left(d_1 \vee d_2 \vee \ldots d_n\right)$$

Optimum Inspection Sequence

- ***Theorem 2:*** For a ***parallel*** Boolean decision function, inspecting attributes $i,\ i = 1, 2, \ldots, n$ in sequential order is optimum (minimizes expected inspection cost) if and only if:

$$c_1 / p_1 \leq c_2 / p_2 \leq \ldots c_n / p_n$$

In other words:

$$F(d_1, d_2, \ldots, d_n) = \left(d_1 \wedge d_2 \wedge \ldots d_n\right)$$

Total Expected Cost

- Example: parallel Boolean decision function
- Cost of misclassifications:

$$C_F = \quad \text{(False acceptance cost)} \quad + \quad \text{(False rejection cost)}$$

$$= \pi c_{FA}\left[1-\prod_{i=1}^{n}\left\{1-\Phi\left(\frac{T_i-\mu_{1i}}{\sigma_{1i}}\right)\right\}\right]+\left(1-\pi\right)c_{FR}\prod_{i=1}^{n}\left\{1-\Phi\left(\frac{T_i-\mu_{0i}}{\sigma_{0i}}\right)\right\}$$

Total Expected Cost

- Example: parallel Boolean decision function
- Cost of inspection:

$$C_I = c_1 + \sum_{i=2}^{n}\left[\prod_{j=1}^{i-1} q_j\right] c_i$$

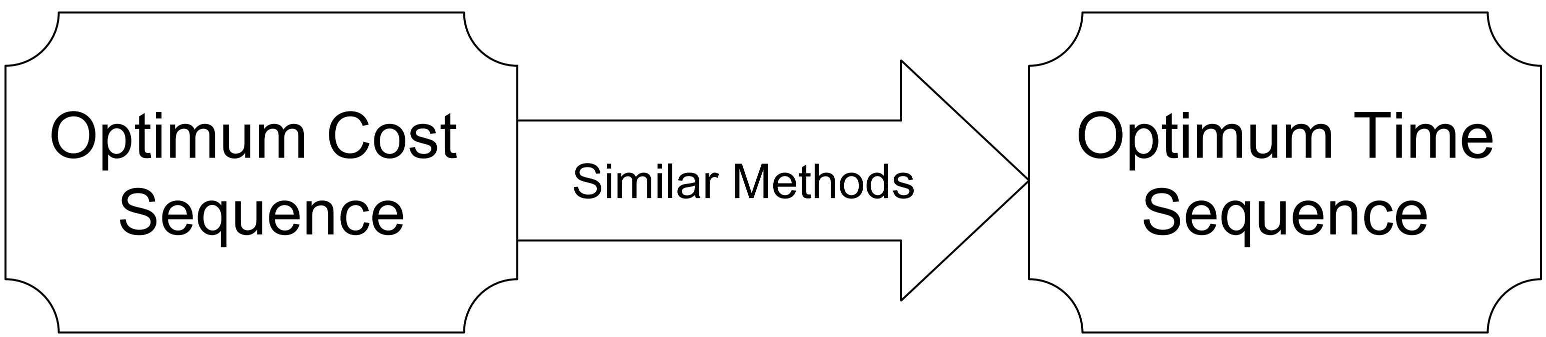
$$= c_1 + \sum_{i=2}^{n} c_i \prod_{j=1}^{i-1}\left[(1-\pi)\left\{1-\Phi\left(\frac{T_j-\mu_{0j}}{\sigma_{0j}}\right)\right\} + \pi\left\{1-\Phi\left(\frac{T_j-\mu_{1j}}{\sigma_{1j}}\right)\right\}\right]$$

- Total expected cost:

$$c_{total} = C_I + C_F$$

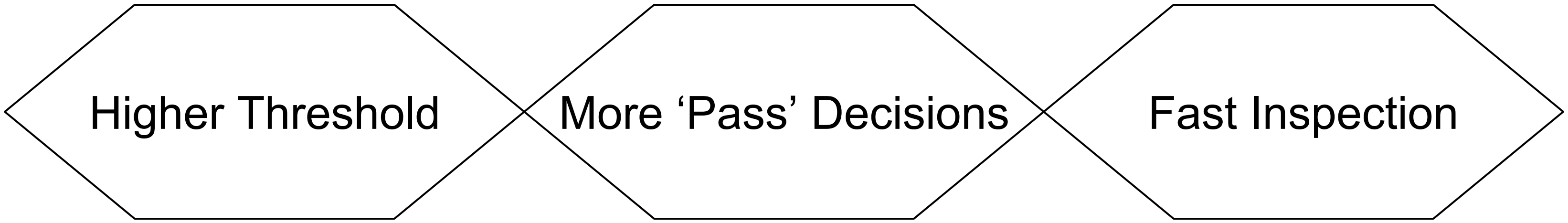
Inspection Time

- Time for a container to complete inspection at station *i* denoted t_i
- Optimum sequence with regard to total expected inspection time can be found with similar method to cost



Inspection Time

t_i could be a function of threshold T_i, approximated from data



Multi-Objective Problem

- Objectives $\min_{Sequence,Threshold} \{c_{total}, t_{total}\}$
  - Minimize the total expected cost including inspection cost and misclassification cost, $c_{total} = C_I + C_F$
  - Minimize the total expected inspection time t_{total}
- Some trade-off between objectives
- Goal: find solutions located along Pareto front
- Find min of weighted objective function $f_{w_1,w_2}(S,T) = w_1 c_{total} + w_2 t_{total}$, $w_2 = 1\text{-}w_1$ for various weights
  - Each solution is a Pareto optimal point for multi-objective problem
- Take advantage of optimal sequence theorem to improve efficiency of algorithms

Optimum Sequence for Weighted Objective

- Given fixed weights, optimum sequence theorem can be adapted
  - Change objective from c_i to $w_1 c_i + w_2 t_i$
- For parallel Boolean, minimum sequence condition:

$$\frac{w_1 c_1 + w_2 t_1}{p_1} \le \frac{w_1 c_2 + w_2 t_2}{p_2} \le \ldots \le \frac{w_1 c_n + w_2 t_n}{p_n}$$

- Condition for series Boolean:

$$\frac{w_1 c_1 + w_2 t_1}{q_1} \le \frac{w_1 c_2 + w_2 t_2}{q_2} \le \ldots \le \frac{w_1 c_n + w_2 t_n}{q_n}$$

Modified Weighted Sum Algorithm

- Computationally expensive to solve minimization of weighted objective function
  - Easier if sequence is known
- Apply optimum sequence given thresholds to compute $f_{w_1,w_2}(T) = \min_S f_{w_1,w_2}(S,T)$
- Solve $\min_T f_{w_1,w_2}(T)$
- Avoiding consideration of all potential sequences improves the efficiency of the algorithm

Solution Methods

- Three methods developed and implemented to compare optimality of results
- Grid search- complete enumeration of discrete threshold values
- Two methods involve repetitions with various weights to solve $\min_{T} f_{w_1,w_2}(T)$ and generate Pareto-optimal solutions
  - Matlab *fmincon* function
  - Genetic algorithm
- Output graphed (Pareto frontier)
  - Time vs. cost expectations

Numerical Example

- Three station system using parallel Boolean decision function
- Cost fixed for each station, $c_i = 1$
- Prior $\pi = 0.0002$
- Distribution parameters

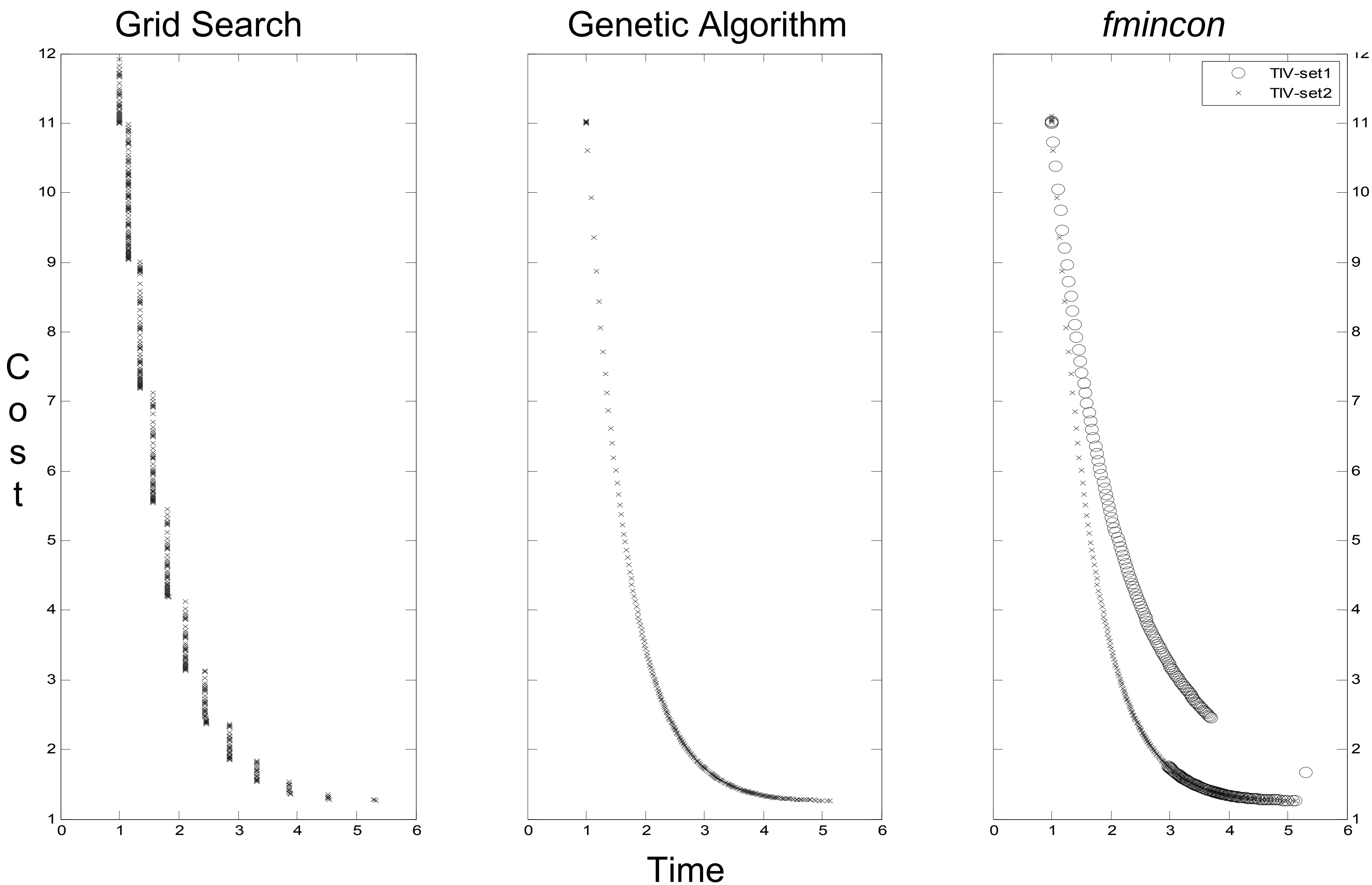
  $\mu_0 = [0\ 0\ 0]$ $\mu_1 = [1\ 1\ 1]$

  $\sigma_0 = [0.16\ 0.2\ 0.22]$ $\sigma_1 = [0.3\ 0.2\ 0.26]$
- Cost parameters

  $c_{FA} = 100000$ $c_{FR} = 500$
- Time related

  $a = [20\ 20\ 20]$ $b = [-3\ -3\ -3]$

Comparison of Three Solution Methods



Pareto-Optimal Solutions

- Output:
  - Graph of time vs. cost trade-off curve
  - Optimal sequence of sensors
  - Threshold level for each sensor
- Examples of solutions

T_1	T_2	T_3	Sequence	Cost	Time
0.2	0.75	0.35	2-3-1	9.03	1.16

Conclusions

- Port-of-entry container inspection problem was formulated to determine optimum threshold levels of sensors by minimizing total expected cost and time
  - Estimate threshold-dependent probabilities of false accept and false reject to calculate expected cost of false classification
  - Sequence of inspection affects expected cost and time of inspection but not probabilities of error
- Compare three approaches to multi-objective problem
  - GA provides dependable set of Pareto solutions

Acknowledgment and Website

- This research is conducted with partial support from ONR grant numbers N00014-05-1-0237, N00014-07-1-029 and NSF grant number NSFSES 05-18543 to Rutgers University
- For further information and related papers, please visit the DIMACS website

  http://dimacs.rutgers.edu/Workshops/PortofEntry