# Port-of-Entry Inspection: Sensor Deployment Policy Optimization

Elsayed A. Elsayed, Christina Schroepfer, Minge Xie, Hao Zhang, and Yada Zhu

***Abstract*—This paper considers the problem of container inspection at the port-of-entry. Containers are inspected through a specific sequence to detect the presence of nuclear materials, biological and chemical agents, and other illegal shipments. The threshold levels of sensors at the inspection stations affect the probabilities of incorrectly accepting or rejecting a container. In this paper we present several optimization approaches on how to select sensor threshold levels under considerations of misclassification errors, total cost of inspection, and budget constraint. Examples of the use of the approaches in different sensor arrangements are demonstrated.**

***Index Terms*—Boolean function, probability of false accept, probability of false reject, sensor threshold levels, receiver operating characteristic curve**

## I. Introduction

THE trade globalization and outsourcing of manufacturing goods have caused significant increases in the number of cargo containers being transported internationally. For example, each year more than 100 million cargo containers which constitute about 90 percent of the entire world's cargo crisscross international sea lanes and more than 95 percent of the non–North American foreign trade arrives into US ports by ship. Slowing the flow long enough to inspect either all or a statistically significant random selection of imports would be economically intolerable (Loy and Ross 2002). The emphasis on improving security of such containers prompted the development and installation of a wide range of inspection machines and sensors. These machines or sensors have different capabilities of detecting the smuggling of nuclear materials, biological agents, drugs, and hazardous and illegal shipments.

One of the most widely used techniques of non-invasively "seeing" into a container (James *et al.* 2002) is based on the use of electromagnetic (EM) waves (radio waves, light, X-rays, $\gamma$-rays, etc.). This technique is utilized in VACIS (vehicle and cargo inspection system) device which combines two formerly separate scanning techniques. Gamma rays look for suspicious images, while radiation detection tracks radioactive signatures.

The task of definitely identifying the contents of the container through this system of inspection stations is very difficult regardless of the X- or $\gamma$-radiation source being used due to the fact that these current scanners are based on the digital radiography and the images obtained are just projection images. All goods in a container are overlapped on the image, and the gray scale is dependent on the total mass thickness along the radiation beam (An *et al*. 2003). Moreover, the penetration capability is one of the main factors which condition the efficiency of an X-ray inspection facility for freight.

The penetration is dependent on two factors of equal importance: (1) The energy of the X-ray beam, which determines the attenuation factor of the beam through a given cargo and (2) The dynamics of the detectors, which corresponds to the ratio between the largest and the smallest signals they can detect. Most authors tend to forget the second point and the only argument taken into account in the design of X-ray systems is the beam energy (Bennett *et al*. 1992). Such reasoning is meant to justify the use of high energy X-rays of up to 10 MV. Indeed such energies are necessary in order to obtain a good penetration when the X-ray detectors have a dynamics of the order of $10^4$, as is the case with traditional technologies where a scintillating material is coupled to a photodiode (Gaillard 1996). In many cases, it becomes necessary to increase the energy of the beam in order to obtain more information about the content of the container. This in turn reduces the probability of falsely identifying the type of the cargo in the container. The threshold level has a direct impact on the classification of the container and the probability of making the "wrong" decision. Therefore, it is important to determine the optimum threshold level that minimizes the overall cost of inspection and making the "wrong" decision (Liao *et al.* 2006).

Identifying the type of the cargo and classifying the container accordingly as acceptable (no suspicious material) or not and the consequence of such decision in terms of Type I error, also known as an "error of the first kind", an $\alpha$ error, or a "false accept" (where a container that has suspicious cargo is accepted) and Type II error, also known as an "error of the second kind", a $\beta$ error, or a "false reject" (where a container is rejected or goes through extensive manual

This research is conducted with partial support from ONR grant number N00014-05-1-0237 and NSF grant number NSFSES 05-18543 to Rutgers University.

E. A. Elsayed, C. Schroepfer, H. Zhang, and Y. Zhu are with the Department of Industrial and Systems Engineering, Rutgers University, Piscataway, NJ 08854-8018 USA (e-mail: elsayed@rci.rutgers.edu)

M. Xie is with the Department of Statistics, Rutgers University, Piscataway NJ 08854 (e-mail: mxie@stat.rutgers.edu).

examination when in fact it has no suspicious contents). Either the Bayesian or the Neyman-Pearson type criterion can be used to model systems with these two types of errors. Thomopoulos *et al*. 1989, for instance, considered a Neyman-Pearson formulation where one assumes a bound on the global probability of false alarm, the goal is to determine the optimum local and global decision rules that minimize the global probability of miss (or equivalently maximize the global probability of detection). When the inspection stations are deployed so that their observations are conditionally independent, one can show that these decision rules are threshold rules based on likelihood ratios (Thomopoulos *et al*. 1989). The problem now becomes one of determining the optimal thresholds of the sensors at each station. While this task is quite non-trivial, it can still be done for a reasonably small number of inspection stations using iterative techniques or using complete enumeration (Zhang *et al*. 2006, Stroud and Saeger 2003, Elsayed 2003, Elsayed and Zhang 2006) ).

In this paper, we decompose the port-of-entry inspection problem into two sub-problems. One problem deals with the determination of the optimum sequence of inspection or the structure of the inspection decision tree in order to achieve the minimum expected inspection cost. This problem can be formulated and investigated using approaches parallel to those used in the optimal sequential inspection procedure for reliability systems as described in Butterworth (1972), Halpern (1974a, 1974b, 1977) , Ben-Dov (1981), Cox *et al*. (1989, 1996), and Azaiez *et al*. (2004). The other problem deals with the determination of the optimum thresholds of the inspection stations so as to minimize the cost associated with false reject and false accept. This paper gives an overview of the solution to the first (sequence) problem and applies those results in obtaining an overall inspection policy solution by determining the optimum threshold levels at inspection stations for a given optimum sequence of inspection.

## II. Problem Description

### A. Port-of-Entry Container Inspection System

The inspection of a container at the port-of-entry is performed sequentially at stations that form the inspection system. Containers are inspected and classified according to observations made regarding their attributes. Suppose there are $n$ inspection stations in the system; one sensor (equipment) at each station is used to identify one attribute of the container being inspected, for example presence of radiation or uncharacteristic X-ray readings. There are several categories into which we seek to classify the containers. In the simplest case, these are negative and positive, 0 or 1, with "0" designating containers (entities) that are considered "acceptable" and "1" designating entities that raise suspicion and require special treatment. After each inspection, we either classify the entity as acceptable or subject it to another inspection process.

The classification is thought of as a *decision function F* that assigns to each binary string of attributes $(a_1, a_2, ..., a_n)$ a category, i.e. $F(a_1, a_2, \ldots, a_n) = 0$ indicates negative class and that there is no suspicion with the container and $F(a_1, a_2, \ldots, a_n) = 1$ means positive class and that additional inspection is required (usually manual inspection). In this paper, we focus on the case where there are two categories, 0 or 1. Thus, $F$ is a *Boolean function*.

By definition, for instance, a series Boolean function is a decision function $F$ that assigns the container class "1" if any of the attributes is present, i.e. $a_i = 1$ for any $i \in \{1, 2, \ldots, n\}$, and a parallel Boolean function is a decision function $F$ that assigns the container the class "1" if all of the attributes are present, i.e. $a_i = 1$ for all $i \in \{1, 2, \ldots, n\}$.

### B. Sensor Performance and Inspection Threshold Levels

Let $X$ represent a randomly selected container for inspection. There are two possibilities: $X = 0$ (representing that this container is acceptable) and $X = 1$ (representing that this container is unacceptable). Suppose that information about the probability of a container's true classification is obtained from inspection history. Denote $\pi = P(X = 1) = 1 - P(X = 0)$. An inspection system consists of multiple sensors or inspection stations. Let $r_i$ be the measurement taken by the $i^{th}$ sensor (inspection station) for this item $X$. Suppose that two normal distributions for each attribute $i$ are assumed or estimated from previous inspections such that $r_i \mid X = 0 \sim N(0, \sigma_{o,i}^2)$ and $r_i \mid X = 1 \sim N(1, \sigma_{1,i}^2)$. Each measurement $r_i$ is compared against a given threshold value $T_i$. Without loss of generality, we assume that the $i^{th}$ station rejects this item ($d_i = 1$) if the reading $r_i$ is higher than $T_i$ and accepts it ($d_i = 0$) if the reading is less than $T_i$.

There are potential errors in making this type of decision. Given an entity with $X = 1$, there is a conditional probability $\alpha_i^0$ that a decision $d_i = 0$ could be made at the $i^{th}$ inspection; this is a type I error (falsely accepting a bad item $X = 1$) $\alpha_i^0 = P(d_i = 0 \mid X = 1) = P(r_i \le T_i \mid X = 1) = \Phi\left(\dfrac{T_i - 1}{\sigma_{1i}}\right)$. There is also a type II error (rejecting a good item $X = 0$) with the conditional probability $\alpha_i^1$ of the decision $d_i = 1$ given $X = 0$

$$\alpha_i^1 = P(d_i = 1 \mid X = 0) = P(r_i > T_i \mid X = 0) = 1 - \Phi\left(\frac{T_i}{\sigma_{0i}}\right).$$

### C. System Inspection Policy

In this paper we consider how different parameters of the inspection system affect the costs associated with performing inspection and misclassification of containers. It becomes clear that the performance of the inspection system is determined by both the sequence in which inspection stations are visited and the threshold levels used at those stations, which we denote collectively as the inspection policy. Therefore the goal of this paper is to formulate the expected cost of inspection and classification errors (false positive and false negative) and use this information to generate a policy for the system's optimum performance. The optimization method is illustrated for inspection systems with decision functions of series, parallel, series-parallel, and parallel-series Boolean functions.

## III. Optimization Approaches

### A. Cost of Wrong Decisions

At the system level, there are also two types of misclassification errors: falsely reject a container that should be cleared and falsely accept a container that should be rejected. Related to these two types of errors in a system, the probability of false reject (PFR) is defined as the probability of false rejection in the overall system. The probability of false accept (PFA) is defined as the probability of false acceptance in the overall system. The complementary probabilities of these two errors are true reject (PTR) and true accept (PTA). Denote by $D$ the decision of the entire inspection system of sensors where $D=1$ means to reject, and $D=0$ to accept. The four probabilities are listed as follows:

$PFR = P(D=1 \mid X=0)$, $PTA = P(D=0 \mid X=0) = 1 - PFR$,

$PFA = P(D=0 \mid X=1)$, and $PTR = P(D=1 \mid X=1) = 1 - PFA$.

The inspection decision of a system $D$ depends on the inspection results of its sensors and the system Boolean function. Some examples (Elsayed 1996) are given below.

Example 1. Series System

The PFR and PFA for the $k$-series system are given by

$$PFR_{series}^{[k]} = 1 - \prod_{j=1}^{k} P(d_j = 0 \mid X = 0) = 1 - \prod_{j=1}^{k} \Phi\left(\frac{T_j}{\sigma_{0,j}}\right) \text{ and}$$

$$PFA_{series}^{[k]} = \prod_{i=1}^{k} P(d_i = 0 \mid X = 1) = \prod_{i=1}^{k} \Phi\left(\frac{T_i - 1}{\sigma_{1,i}}\right)$$

Example 2. Parallel System

The PFR and PFA for the $k$-parallel system are given by

$$PFR_{parallel}^{[k]} = \prod_{j=1}^{k} P(d_j = 1 \mid X = 0) = \prod_{j=1}^{k} [1 - \Phi\left(\frac{T_j}{\sigma_{0j}}\right)] \text{ and}$$

$$PFA_{parallel}^{[k]} = 1 - \prod_{i=1}^{k} P(d_i = 1 \mid X = 1) = 1 - \prod_{i=1}^{k} [1 - \Phi\left(\frac{T_i - 1}{\sigma_{1i}}\right)]$$

Example 3. Parallel-series System

The PFR and PFA for the ($n$, $m$) parallel-series system (shown in Figure 2) are given by

$$PFR_{parallel-series}^{[n,m]} = \prod_{i=1}^{n} [1 - \prod_{j=1}^{m} P(d_{ij} = 0 \mid X = 0)] = \prod_{i=1}^{n} [1 - \prod_{j=1}^{m} \Phi\left(\frac{T_{ij}}{\sigma_{o,ij}}\right)]$$

$$PFA_{parallel-series}^{[n,m]} = 1 - \prod_{i=1}^{n} [1 - \prod_{j=1}^{m} P(d_{ij} = 0 \mid X = 1)] = 1 - \prod_{i=1}^{n} [1 - \prod_{j=1}^{m} \Phi\left(\frac{T_{ij} - 1}{\sigma_{1,ij}}\right)]$$

Example 4. Series-parallel System

The PFR and PFA for the ($n$, $m$) series-parallel system (shown in Figure 3) are given by

$$PFR_{series-parallel}^{[n,m]} = 1 - \prod_{i=1}^{n} [1 - \prod_{j=1}^{m} P(d_{ij} = 1 \mid X = 0)] = 1 - \prod_{i=1}^{n} \{1 - \prod_{j=1}^{m} [1 - \Phi\left(\frac{T_{ij}}{\sigma_{o,ij}}\right)]\} \text{ and}$$

$$PFA_{series-parallel}^{[n,m]} = \prod_{i=1}^{n} [1 - \prod_{j=1}^{m} P(d_{ij} = 1 \mid X = 1)] = \prod_{i=1}^{n} \{1 - \prod_{j=1}^{m} [1 - \Phi\left(\frac{T_{ij} - 1}{\sigma_{1,ij}}\right)]\}$$

Ideally, we seek a set of threshold values of the sensors and system configurations in a system based on which both these errors are minimized. Unfortunately, this simultaneous minimization of both PFR and PFA is not possible, and reducing one error is likely to increase the other. Therefore, we need to define an optimal policy that balances such a trade off.

One feasible approach is through defining the expected cost of the system making a wrong decision. Let $c_{FA}$ be the cost of the system accepting a "bad" container and $c_{FR}$ be the cost of the system rejecting a "good" container. Then the total cost of the system making a wrong decision is

$$C_F = \pi \; PFA \; c_{FA} + (1 - \pi) \; PFR \; c_{FR} .$$

The set of optimal threshold values is the one that minimizes the expected cost of the system making a wrong decision $C_F$ over all possible combinations of sensors in the system and all possible threshold values:

$$\{T_1, T_2, ..., T_k\} = \arg\min C_F .$$

Although it may not be easy to assign a specific value, measuring the cost of false rejects $c_{FR}$ is relatively straightforward. The cost of FR is the cost of additional tests. In the practice of port-of-entry inspection, these additional tests mean inspecting the contents manually. This is quite expensive since it might involve several workers for several hours and delays in completing the inspection and reduction in the inspection system throughput. Measuring the costs of false acceptance $c_{FA}$ is even more challenging. Indeed, the cost of missing a container that contains illegal drugs is not comparable to the cost of missing a container that holds a "dirty bomb". One way is to assign a large cost value, say a few hundred -or even more- times the cost of a false reject, however these choices are very subjective.

A more flexible approach that avoids assigning exact values to the misclassification costs is the so called Receiver Operating Characteristics (ROC) curve. In diagnostic situations, an ROC curve provides a useful graphical representation of the tradeoff between the probability of false accept (PFA) and probability of false reject (PFR). Typically, we plot the probability of true reject (PTR = 1 - PFA) against the probability of false reject (PFR) while varying the threshold parameters and the arrangement of sensors in a system which results in an ROC curve as shown in Figure 1.

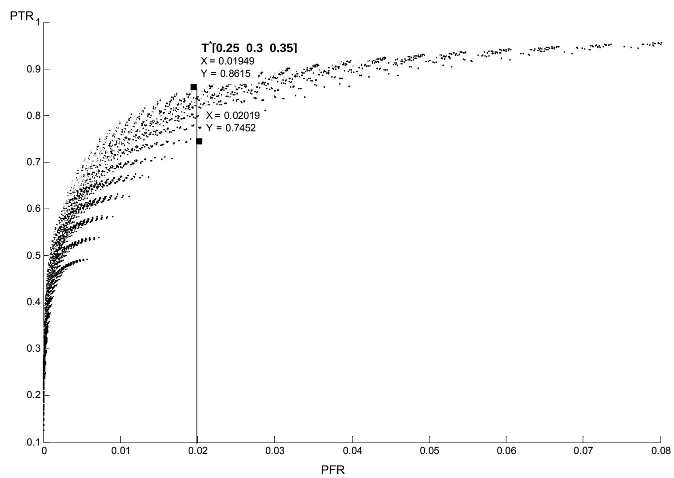


Figure 1. ROC Curve for Parallel System

All of the points are in a [0,1] by [0,1] square box. The most upper-left points form a curve which is referred to as the ROC curve. The ROC curve is in fact the optimal curve in the sense of Pareto optimization; it consists of the best choices of threshold values under different conditions. In particular, for any choices of ($c_{FA}, c_{FR}$), there is a point on the ROC curve that corresponds to the aforementioned optimization problem. The ROC curve always passes through two points (0, 0) and (1, 1) in extreme cases. The point (0, 0) corresponds to $c_{FA}$=0 and $c_{FR}$=∞, where the classifier finds no positives (detects no alarms). In this case, it always classifies the negative cases correct but it classifies all positive cases wrong. The point (1, 1) corresponds to $c_{FA}$=∞ and $c_{FR}$=0, where all containers are classified as positive. So all positive cases are correctly classified but all negative cases are wrongly classified (i.e. it raises a false alarm on each negative case).

The ROC Curves are flexible and useful tools in decision-making. In practice, we often choose an operating point, a fixed point on the ROC curve, where a set of threshold levels can be determined. For a given condition, the best operating point might be chosen so that the classifier gives the best trade off between the costs of failing to detect positives against the costs of raising false alarms. In our port-of-entry problem, for example, we may be able to set a small tolerance level for the FA (so that the PTR will be always be constrained above this level) choosing a set of threshold values that minimizes the PFR. In this case, we can identify the operating point by drawing a vertical line at the tolerance level, where the interception of the ROC curve and this vertical line is the operating point; see Figure 1. The set of threshold values that corresponds to this operating point is the optimal choice that we are seeking.

### *B. Expected Inspection Cost and Optimal Sequence*

In addition to the cost of making false decisions, there is also the cost of inspection itself. There are many possible ways to calculate the cost of obtaining a sensor reading. For instance, we can break down the cost of obtaining a sensor reading into two components: unit variable cost and fixed cost. The unit variable cost is just the cost of using the sensor to inspect one container, and the fixed cost is the cost of the purchase and deployment of the sensor itself. In many cases, the primary cost is the unit variable cost since many inspections are very labor intensive. The fixed cost is usually a constant and often does not contribute to the optimization functions, so for simplicity we disregard the fixed cost. Thus, the inspection cost is basically the expected cost of making observations for a container. Note that depending on the system configuration, a container may or may not be inspected by all sensors. The arrangement of the sensors is closely related to the inspection cost.

Denote $p_i$ and $q_i$ by

$$p_i = P(d_i = 0) = \sum_{j=0}^{1}[P(d_i = 0 \mid X = j)P(X = j)]$$

$$= (1-\pi)\Phi\left(\frac{T_i}{\sigma_{0i}}\right) + \pi\Phi\left(\frac{T_i - 1}{\sigma_{1i}}\right).$$

and

$$q_i = 1 - p_i = (1-\pi)\{1 - \Phi\left(\frac{T_i}{\sigma_{0i}}\right)\} + \pi\{1 - \Phi\left(\frac{T_i - 1}{\sigma_{1i}}\right)\}$$

They are functions of threshold values $T_i$. Let $c_i$ be the inspection cost of sensor *i*. Zhang *et al* (2006) prove the following theorem.

Theorem 1: *(a) For a series Boolean decision function, inspecting attributes* $i = 1, 2, \ldots, n$ *in sequential order is optimum, minimizes expected inspection cost, if and only if:* $c_1 / q_1 \le c_2 / q_2 \le \ldots c_n / q_n$ *(condition 1a).*

*In this case, the expected inspection cost is given by*

$$C_I = c_1 + \sum_{i=2}^{n}[\prod_{j=1}^{i-1} p_j]\, c_i \, .$$

*(b) For a parallel Boolean decision function, inspecting attributes* $i = 1, 2, \ldots, n$ *in sequential order is optimum, minimizes expected inspection cost, if and only if:* $c_1 / p_1 \le c_2 / p_2 \le \ldots c_n / p_n$ *(condition 1b).*

*In this case, the expected inspection cost is given by*

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

We generalize these results to systems with arrangements of Parallel-Series and Series-Parallel sensors, given in Theorem 2.

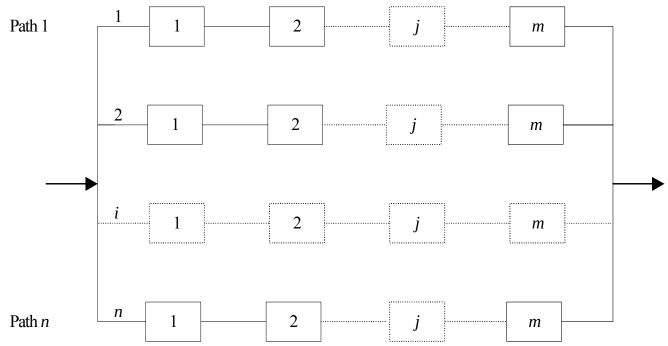


Figure 2. Conceptual Network of Parallel-Series System

Theorem 2: *(a) Consider a parallel-series decision function with n paths of m sensors each (see Figure 2). If an inspection system with attributes* $i = 1, 2, \ldots, n$ *and* $j = 1, 2, \ldots, m$ *arranged in parallel-series is optimal, it satisfies the following conditions: the inspection sequence of the series of sensors within each path should be arranged in the order of* $c_{i,1}/q_{i,1} \le c_{i,2}/q_{i,2} \le \ldots c_{i,m}/q_{i,m}$*, and the inspection sequence of parallel paths should be arranged in the order of* $C_1/P_1 \le C_2/P_2 \le \ldots \le C_n/P_n$ *(condition 2a). Here,* $C_i$ *and* $P_i$ *are the (minimal) inspection cost and the probability of acceptance of the* $i^{th}$ *path:*

$$C_i = c_{i1} + \sum_{j=2}^{m} [\prod_{k=1}^{j-1} p_{ik}]\, c_{ij}$$

$$= c_{i1} + \sum_{j=2}^{m} c_{ij} \prod_{k=1}^{j-1} [(1-\pi)\Phi\left(\frac{T_{ik}}{\sigma_{0,ik}}\right) + \pi\Phi\left(\frac{T_{ik}-1}{\sigma_{1,ik}}\right)]$$

*and* $P_i = P(D_i = 0) = \prod_{j=1}^{m} p_{ij}$ *. In this case, the minimal inspection cost is:*

$$C_I = C_1 + \sum_{i=2}^{n} [\prod_{j=1}^{i-1} (1-P_j)]\, C_i = C_1 + \sum_{i=2}^{n} C_i \prod_{j=1}^{i-1} (1 - \prod_{k=1}^{m} p_{jk}).$$

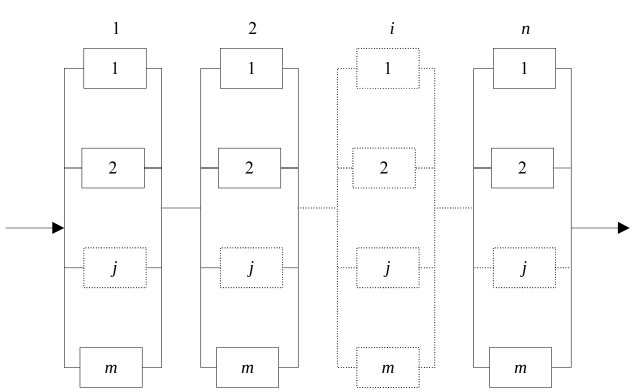


Figure 3. Conceptual Network of Series-Parallel System

*(b) Consider a series-parallel decision function that has n subsystems in series with m units in parallel in each subsystem (see Figure 3). If an inspection system with attributes* $i = 1, 2, \ldots, n$ *and* $j = 1, 2, \ldots, m$ *arranged in series-parallel is optimal, it satisfies the following conditions: the inspection sequence of the series of each subsystem should be arranged in the order of* $c_{i,1}/p_{i,1} \le c_{i,2}/p_{i,2} \le \ldots c_{i,m}/p_{i,m}$*, and the inspection sequence of parallel paths should be arranged in the order of* $C_1/Q_1 \le C_2/Q_2 \le \ldots \le C_n/Q_n$ *(condition 2b). Here,* $C_i$ *and* $Q_i$ *are the (minimal) inspection cost and the probability of rejection of the* $i^{th}$ *subsystem:*

$$C_i = c_{i1} + \sum_{j=2}^{m} [\prod_{k=1}^{j-1} q_{ik}]\, c_{ij}$$

$$= c_{i1} + \sum_{j=2}^{m} c_{ij} \prod_{k=1}^{j-1} [(1-\pi)\{1-\Phi\left(\frac{T_{ik}}{\sigma_{0,ik}}\right)\} + \pi\{1-\Phi\left(\frac{T_{ik}-1}{\sigma_{1,ik}}\right)\}]$$

*and* $Q_i = P(D_i = 1) = \prod_{j=1}^{m} (1-p_{ij})$ *. In this case, the minimal inspection cost is*

$$C_I = C_1 + \sum_{i=2}^{n} (\prod_{j=1}^{i-1} P_j)\, C_i = C_1 + \sum_{i=2}^{n} C_i \prod_{j=1}^{i-1} \{1 - \prod_{k=1}^{m} (1-p_{jk})\}.$$

The optimal arrangement of sensors depends on the values of $p$'s and $q$'s, which are functions of the threshold values. We formulate the optimization problem of choosing thresholds to minimize the inspection cost as follows:

$$\{T_1, T_2, ..., T_k\} = \arg\min C_I$$

where the optimization is over threshold values that satisfy the constraints given by the sufficient and necessary conditions 1a, 1b, 2a, and 2b stated in the two theorems.

C. *Expected Total Cost- Combined Optimization*

In some situations, it is conceivable that we may consider the combined cost of making wrong decisions and inspection cost. In this case, the total expected cost is the sum of the expected inspection cost and the cost of wrong decision: $C_{total} = C_I + C_F$ .

The total cost $C_{total}$ is calculated from the results of the previous sections. To facilitate the computation for different systems with a large number of sensors, we provide a induction methods which calculate the cost (both $C_I$ and $C_F$ ) in Appendices 1-4.

The optimization problem now becomes finding a set of threshold values $\{T_1, T_2, ..., T_k\}$ that minimizes the total expected cost:

$$\{T_1, T_2, ..., T_k\} = \arg\min C_{Total}$$

among the sets of threshold values that satisfy the constraints in Section III.B.

The optimal set of threshold values from this optimization may be different from those obtained from optimization by minimizing cost of misclassification errors or minimizing inspection cost. In the case of port-of-entry inspection, the optimal solution of the cost-combined optimization may be very close to the solution obtained by the former if $C_I$ is much smaller than the cost of the system making a wrong decision. Note that in practice in the port-of-entry inspection, the cost of false positives is often the cost of additional testing, such as opening the container and manually inspecting its contents. This is quite expensive since it might involve several workers for hours, delays in completing the inspection, and reduction in the inspection system throughput as stated earlier in the paper. In comparison to routine inspection cost of unit testing such as neutron or gamma emissions detection, this FR cost is relatively high. The FA cost would be even greater, including a huge potential social or economic impact.

### D. *Optimization with Budget Consideration*

At a given port of entry of inspection station, the inspection practice is often constrained by budget. It is not possible to open and manually inspect every container or every cargo, which is by far the most accurate but also extremely costly inspection method. If budget allows, we may want to allow more containers to be manually inspected, which in turn affects the sensor inspection process. For example, if the budget is large, it is possible to set low threshold levels to increase the PTR of the sensor system, and flag more containers for further manual inspection.

In our formulation, the total budget in an inspection station covers both the initial cost of the inspection system and additional manual inspection cost. Therefore, the budget is defined by

$$budget = C_I + C_{manual} = C_I + c_{unpack}[(1-\pi)PFR + \pi PTR],$$

where $C_I$ is the cost of initial system inspection, and $c_{unpack}$ is the unit cost of additional manual inspection (unpacking the container).

Under the budget constraint, we maximize the probability of properly classifying suspicious cargos passing through the entire inspection system, including the sensor inspection system and manual inspections. So, the optimization problem can be described as:

$$\{T_1, T_2, ..., T_k\} = \arg\max PTR$$
$$\text{subject to: Budget} < B_0$$

where $B_0$ is maximum available budget for the inspection station, and the $\{T_1, T_2, ..., T_k\}$ is selected from possible threshold level values. We can formulate the budget constraint optimization problem similarly for other considerations. For example, minimization of the cost of making wrong decisions can be obtained by finding the argument of the minimum of $C_F$ defined in Section III.A.

This optimization problem can be presented by a graphical technique similar to the ROC curve, especially when we want to investigate the impact of the budget constraint. For instance, it is informative to investigate the relation between the chance of missing a suspicious cargo or dirty bomb and the budget. In this case, we plot the *PTR = 1 – PFA* against the total budget $B_0$ while varying the threshold values and the combination of sensors in an inspection system; see Figure 7. The most upper-left points form a curve. This curve consists of points corresponding to optimal threshold values and best combination of sensors at different budget levels.

## IV. System Analysis with Numerical Examples

This section includes computations for the Boolean functions: parallel, series, parallel-series, and series-parallel. Numerical examples, graphs, and results are also presented.

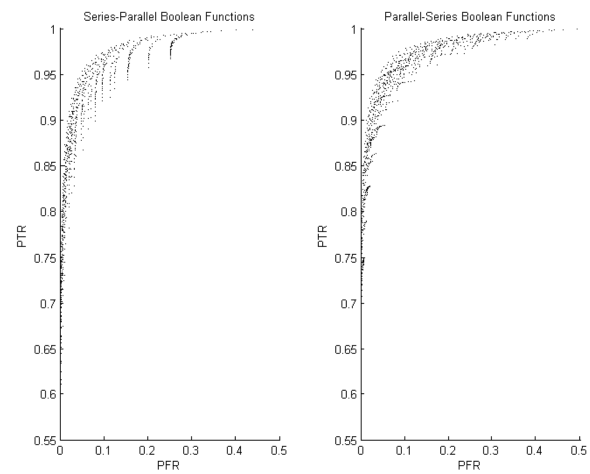


Figure 4. ROC Curves for Series-Parallel vs. Parallel-Series

Figure 5 presents the minimum total cost of an inspection policy given the following system information: parallel Boolean decision function, unit misclassification penalty costs $c_{FR} = \$500$ and $c_{FA} = \$100{,}000$, unit inspection cost $c_1 = c_2 = c_3 = 1$, and distribution parameters $\mu_{0i} = 1$, $\mu_{1i} = 2$, and $\sigma_{0i} = \sigma_{1i} = 0.5$ (i=1,2,3). The results are arranged by varying $T_1$ values along the horizontal axis and each point represents an optimal combination of threshold values and sequence, with the total cost along the vertical axis. The data series are the result of varying the prior distribution.

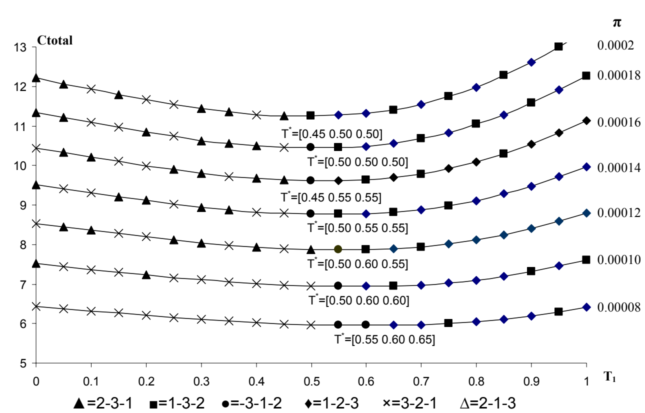


Figure 5. Minimum Cost Curves for Parallel System: Effect of $T_1$ and Prior

Figure 5 also illustrates that for given parameters and prior distribution, there is an optimal sequence and threshold values that correspond to the minimum total cost. For example, given $\pi = 0.0002$, the combination of threshold values $\{T_1, T_2, T_3\} = T^*[0.45, 0.50, 0.50]$ in the inspection sequence $2-3-1$ results in the minimum total cost for a system implementing a parallel Boolean decision function. The variation of the prior probability value can influence the optimal inspection sequence but does not have much influence on the optimal threshold values in this example.

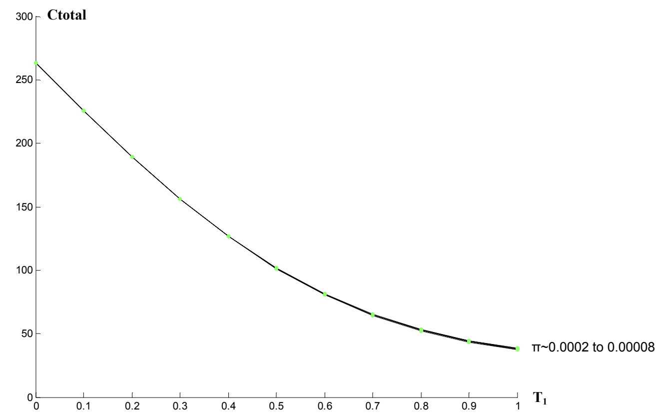


Figure 6. Minimum Cost Curves for Series System: Effect of $T_1$ and Prior

Figure 6 presents the results for a series Boolean decision function. The parameter values and presentation of results are similar to the parallel Boolean example. In comparison to the results for the parallel Boolean decision function, the inspection sequence and threshold values for the series Boolean are not sensitive to small prior values. As seen in Figure 6, the minimum cost curves for various prior values overlap. In general, the total cost of the series system is higher than that for parallel for the given parameters. This is due to the relative increase in the expectation of a costly false acceptance.

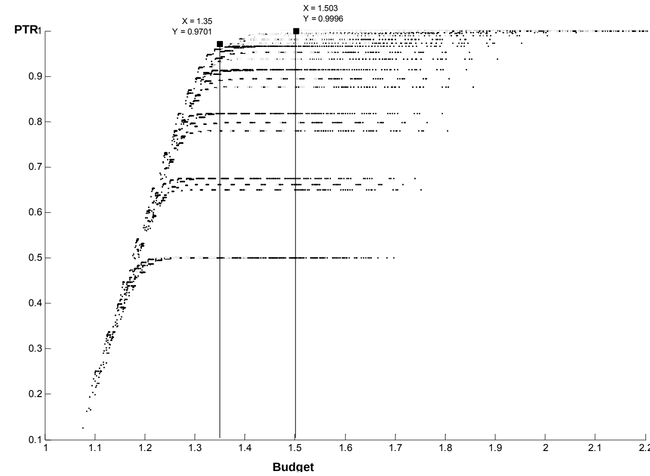


Figure 7. Budget Curve for Parallel System

With the same parameters in the series system, we plot the relationship between the budget level and the probability of true reject; see Figure 7. If $B_0$ = 1.35 on the horizontal axis, the optimal threshold values are $T^*=\{0.75 \;\; 0.05 \;\; 0.75\}$ and the probability of missing a suspicious cargo in the optimal case is about $1-0.8=0.2$. If we increase the budget from $B_0$ = 1.35 to $B_0$ = 1.50, the optimal $T^*=\{0.5 \;\; 0.4 \;\; 0.6\}$ and the probability of missing a suspicious cargo decreases to less than 0.1. Clearly, the 11% increase of the budget results in a significant increase in the detection of unacceptable containers. Now with $B_0$ = 1.50, an additional budget increase of the same amount results in little change in the probability of missing a suspicious cargo. It is not cost effective to apply the additional amount of inspection. Such information may help decision makers in assigning appropriate budgets to the Port-of-entry inspection stations.

## V. DISCUSSION

In this paper we investigated the POE problem with a small number of inspection stations. Complete enumeration of all possible threshold levels for each sensor resulted in determining the optimum threshold levels for the sensors such that the total inspection cost is minimized. This has been done for series, parallel, series-parallel and parallel-series sensors configurations. The key factor that has a direct effect on the determination of the sensors threshold levels is the cost of misclassification of the container with type I error. This cost is difficult to estimate as it is a function of many unknowns but its effect could be catastrophic. Clearly, tightening the threshold levels will minimize the type I error but may increase the cost of delaying the container. This has not been considered in this research but raises an important consideration of not only the cost of container inspection but also the cost of delay incurred in the system. Hence, we have two conflicting objectives. This is a multi-objective optimization problem which will be investigated in the future. Likewise, the optimum inspection sequencing problem has not been addressed in this paper and it demands further work. Finally, the determination of the optimum threshold levels of sensors arranged in non standard arrangements such as a general network of sensors or k-out-of-n arrangement warrants further investigation.

## APPENDICES

### A.1 Induction Formula for the Parallel System

Suppose the $(k+1)^{th}$ sensor is added to the *k*-parallel sensors with $c_1/p_1 \le \ldots \le c_K/p_K \le c_{k+1}/p_{k+1}$. For this system of (*k*+1) parallel sensors, the cost of inspection can be calculated by the following induction formula:

$$C_I^{[1]} = c_1 \text{ and } C_I^{[k+1]} = C_I^{[k]} + c_{k+1}\prod_{j=1}^{k} q_j, \text{ for } k = 1,2,\ldots$$

The cost function of making a false decision is:

$$C_F^{[k+1]} = \pi c_{FA}(1 - A^{[k+1]}) + (1-\pi)c_{FR}B^{[k+1]},$$

where $A^{[k+1]}$ and $B^{[k+1]}$ can be computed by the following induction formulas:

$$A^{[1]} = \overline{\Phi}\left(\frac{T_1 - 1}{\sigma_{11}}\right),\ B^{[1]} = \overline{\Phi}\left(\frac{T_1}{\sigma_{01}}\right),\ A^{[k+1]} = A^{[k]}\overline{\Phi}\left(\frac{T_{k+1}-1}{\sigma_{1,k+1}}\right),$$

$$\text{and } B^{[k+1]} = B^{[k]}\overline{\Phi}\left(\frac{T_{k+1}}{\sigma_{0,k+1}}\right),\ \text{for } k = 1,2,\ldots \text{ and } \overline{\Phi} = 1-\Phi.$$

### A.2 Induction Formula for the Series System

Suppose the $(k+1)^{th}$ sensor is added to the *k*-series sensors with $c_1/q_1 \le \ldots \le c_k/q_k \le c_{k+1}/q_{k+1}$. For this system of (*k*+1) series sensors, the cost of inspection can be calculated by the following induction formula:

$$C_I^{[1]} = c_1 \text{ and } C_I^{[k+1]} = C_I^{[k]} + c_{k+1}\prod_{j=1}^{k} p_j,\ \text{for } k = 1,2,\ldots$$

The cost function of making a false decision is $C_F^{[k+1]} = \pi c_{FA}A^{[k+1]} + (1-\pi)c_{FR}(1 - B^{[k+1]})$,

where $A^{[k+1]}$ and $B^{[k+1]}$ can be computed by the following induction formulas:

$$A^{[1]} = \Phi\left(\frac{T_1 - 1}{\sigma_{11}}\right),\ B^{[1]} = \Phi\left(\frac{T_1}{\sigma_{01}}\right), A^{[k+1]} = A^{[k]}\Phi\left(\frac{T_{k+1}-1}{\sigma_{1,k+1}}\right),$$

$$\text{and } B^{[k+1]} = B^{[k]}\Phi\left(\frac{T_{k+1}}{\sigma_{0,k+1}}\right),\ \text{for } k = 1,2,\ldots$$

### A.3 Induction Formula for the Parallel-Series System

**Step 1:** Add $(m+1)^{\text{th}}$ sensor to each branch:
(*n*, *m*)→(*n*, *m*+1)

If $c_{i,1}/q_{i,1} \le \ldots \le c_{i,m}/q_{i,m} \le c_{i,m+1}/q_{i,m+1}$ for all branches, Section III-B gives the minimum inspection cost as

$$C_I^{[m+1]} = C_{I,1}^{[m+1]} + \sum_{i=2}^{m} C_{I,i}^{[m+1]}\prod_{j=1}^{i-1}[1 - \prod_{k=1}^{m+1} p_{j,k}],$$

where the inspection cost of the $i^{\text{th}}$ path of *m*+1 sensors in series $C_{I,i}^{[m+1]}$, for $i = 1,2,\ldots,n$, can be calculated by the induction formula in A.2.

From the results from Section III-A, the total expected cost for wrong decision is

$$C_F^{[m+1]} = \pi c_{FA}[1 - PM^{[m+1]}] + (1-\pi)c_{FR}PN^{[m+1]},$$

where $PM^{[m+1]} = \prod_{i=1}^{n}\{1 - M_i^{[m+1]}\}, PN^{[m+1]} = \prod_{i=1}^{n}\{1 - N_i^{[m+1]}\}$, and $M_i^{[m+1]}$ and $N_i^{[m+1]}$ can be updated by the following induction formulas

$$M_i^{[m+1]} = M_i^{[m]}\Phi\left(\frac{T_{i,m+1}-1}{\sigma_{1,(i,m+1)}}\right) \text{ and } N_i^{[m+1]} = N_i^{[m]}\Phi\left(\frac{T_{i,m+1}}{\sigma_{0,(i,m+1)}}\right).$$

**Step 2:** Add the $(n+1)^{\text{th}}$ branch with *m* sensors:
(*n*, *m*) →(*n*+1, *m*)

If $C_{I,1}/Q_1 \le \ldots \le C_{I,n}/Q_n \le C_{I,n+1}/Q_{n+1}$, Section III-B gives the minimum inspection cost as:

$$C_I^{[n+1]} = C_I^{[n]} + C_{I,n+1}G_n$$

where $G_n$ can be updated by induction:

$$G_1 = Q_1, \text{ and } G_{j+1} = G_jQ_j,$$

$$\text{with } Q_j = 1 - \prod_{s=1}^{m}[(1-\pi)\Phi\left(\frac{T_{js}}{\sigma_{o,js}}\right) + \pi\Phi\left(\frac{T_{js}-1}{\sigma_{1,js}}\right)].$$

From the result of Section III-A, the total expected cost for wrong decision is

$$C_F^{[n+1]} = \pi c_{FA}[1 - PA^{[n]}A_{n+1}] + (1-\pi)c_{FR}PB^{[n]}B_{n+1}$$

where $PA^{[n+1]}$ and $PB^{[n+1]}$ can be updated by induction formula, for j =1, 2, …, N,

$$PA^{[1]} = A_1, PA^{[j+1]} = PA^{[j]}A_j,$$

$$PB^{[1]} = B_1, PB^{[j+1]} = PB^{[j]}B_j,$$

and, for j =1, 2, …, *n*+1,

$$A_i = 1 - \prod_{j=1}^{m}\Phi\left(\frac{T_{ij}-1}{\sigma_{1,ij}}\right) \text{ and } B_i = 1 - \prod_{j=1}^{m}\Phi\left(\frac{T_{ij}}{\sigma_{o,ij}}\right).$$

### A.4 Induction Formula for the Series-Parallel System

**Step 1**: Add the $(m+1)^{\text{th}}$ sensor in each subsystem:
(*n*, *m*) → (*n*, *m*+1)

If $c_{i,1}/p_{i,1} \le \ldots c_{i,m}/p_{i,m} \le c_{i,m+1}/p_{i,m+1}$ for all the branches, Section III-B gives the minimum inspection cost as

$$C_I^{[n,m+1]} = C_{I,1}^{[m+1]} + \sum_{i=2}^{n} C_{I,i}^{[m+1]}\prod_{j=1}^{i-1}[1 - \prod_{s=1}^{m+1}(1 - p_{js})]$$

where the inspection cost of the $i^{\text{th}}$ path of *m*+1 sensors in series $C_{I,i}^{[m+1]}$, for $i = 1,2,\ldots,n$, can be calculated by the induction formula in A.1.

From the result of Section III-A, the total expected cost for wrong decision is

$$C_F^{[m+1]} = \pi c_{FA}(PA^{[m+1]}) + (1-\pi)c_{FR}(1 - PB^{[m+1]})$$

where $PA^{[m+1]} = \prod_{i=1}^{n} \{1 - A_i^{[m+1]}\}, PB^{[m+1]} = \prod_{i=1}^{n} \{1 - B_i^{[m+1]}\}$, and

$A_i^{[m+1]}$ and $B_i^{[m+1]}$ can be updated by induction formula:

$$A_i^{[m+1]} = A_i^{[m]} \overline{\Phi}\left(\frac{T_{i,M+1} - 1}{\sigma_{1,(i,M+1)}}\right),\ B_i^{[m+1]} = B_i^{[m]} \overline{\Phi}\left(\frac{T_{i,m+1}}{\sigma_{0,(i,m+1)}}\right), \text{with } \Phi = 1 - \overline{\Phi}.$$

**Step 2:** Add the ($n$+1)th subsystem with $m$ sensors:
$(n, m) \rightarrow (n+1, m)$

If $C_{I,1} / Q_1 \leq ... \leq C_{I,n} / Q_n \leq C_{I,n+1} / Q_{n+1}$, Section III-B gives the minimum inspection cost as:

$$C_I^{[n+1]} = C_I^{[n]} + C_{I,n+1} G_n$$

where $G_n$ can be updated by induction:

$G_1 = P_1$, and $G_{j+1} = G_j P_j$,

with $P_j = 1 - \prod_{s=1}^{m} [(1-\pi)\Phi\left(\frac{T_{js}}{\sigma_{o,js}}\right) + \pi\Phi\left(\frac{T_{js} - 1}{\sigma_{1,js}}\right)].$

From the result of Section III-A, the total expected cost for wrong decision is

$$C_F^{[n+1]} = \pi c_{FA} PA^{[n]} A_{n+1} + (1-\pi) c_{FR} PB^{[n]} B_{n+1}$$

where $PA^{[n+1]}$ and $PB^{[n+1]}$ can be updated by induction formula, for j =1, 2, …, n,

$PA^{[1]} = A_1, PA^{[j+1]} = PA^{[j]} A_j$,

$PB^{[1]} = B_1, PB^{[j+1]} = PB^{[j]} B_j$,

and, for j =1, 2, …, $n$+1,

$$A_i = 1 - \prod_{j=1}^{m} \overline{\Phi}\left(\frac{T_{ij} - 1}{\sigma_{1,ij}}\right), B_i = 1 - \prod_{j=1}^{m} \overline{\Phi}\left(\frac{T_{ij}}{\sigma_{o,ij}}\right), \text{with } \Phi = 1 - \overline{\Phi}.$$

## ACKNOWLEDGMENT

This research is conducted with partial support from the Office of Naval Research and the National Science Foundation. The authors wish to thank Fred Roberts, Paul Kantor, and other members of the DIMACS Port-of-Entry research group for their comments and input throughout the research.

## REFERENCES

[1] An, J., Xiang, X., Wu, Z., Zhou, L., Wang, L. and Wu, H., (2003), "Progress on developing $^{60}$Co container inspection systems," *Applied Radiation and Isotopes* 58 (2003) 315–320
[2] Azaiez, N., Bier, V. M., (2004), Optimal resource allocation for security in reliability systems, *CREATE Report*.
[3] Ben-Dov, Y., (1981), "Optimal testing procedures for special structures of coherent systems", *Management Science*, 27(12):1410-1420.
[4] Bennett, G., Gay, A.M. and Waters, D.D., (1992), "The design of X-ray systems for cargo inspection," *Contraband and Cargo Inspection Technology International Symposium*, Washington, D.C.
[5] Butterworth, R. W., (1972), "Some reliability fault-testing models", *Operations Research*, 20, 335-343.
[6] Cox, L., Qiu, Y., Kuehner, W., (1989), "Heuristic least-cost computation of discrete classification functions with uncertain argument values", *Annals of Operations Research,* 21 1-30.
[7] Cox, L., Chiu, S., Sun, X., (1996), Least-cost failure diagnosis in uncertain reliability systems, *Reliability Engineering and System Safety*, 54, 203-216.
[8] Elsayed, E. A., (1996), *Reliability Engineering*, Addison Wesley.
[9] Elsayed, E. A., (2003), Distributed sensing for quality and productivity improvement, (Panelist) *INFORMS*, Atlanta, GA.
[10] Elsayed, E. A., and Zhang, H., (2006), "Optimum Threshold Level of Degraded Structures Based on Sensors Data," *Proceedings of the 12th ISSAT International Conference on Reliability and Quality in Design*, Chicago, Illinois, USA, August 3-5, pages 187-191.
[11] Gaillard, G., (1996), "Accelerator based X-ray facilities applied to freight control," *Nuclear Instruments and Methods in Physics Research B* 1 I3, I28-133.
[12] Halpern, J., (1974), "Fault-testing of a k-out-n system," *Operations Research*, 22, 1267-1271.
[13] Halpern, J., (1974), "A sequential testing procedure for a system's state identification", *IEEE Transactions on Reliability*, Vol R-23, No 4, 267-272.
[14] Halpern, J., (1977), "The sequential covering problem under uncertainty", *INFOR*, Vol. 15, 76-93.
[15] James, K. A., de Sulima-Przyborowski, J. C. and Haydn, N. T. A., (2002), Non-invasive means of investigating container contents for customs agents at port, University Southern California.
[16] Liao, H., Elsayed, E. A., and Ling-Yau Chan, (2006) "Maintenance of Continuously Monitored Degrading Systems," *European Journal of Operational Research*, Vol. 75, No. 2, 821-835.
[17] Loy, J. M. and Ross, R. G., (2002), Global Trade: America's Achilles' Heel, Defense Horizon, Center for Technology and National Security University.
[18] Stroud, P.D. and Saeger, K.J., (2003) "Enumeration of Increasing Boolean Expressions and Alternative Digraph Implementations for Diagnostic Applications", In: *Proceedings Volume IV, Computer, Communication and Control Technologies: I*, (eds. H. Chu, J. Ferrer, T. Nguyen, Y. Yu), pp. 328-333.
[19] Thomopoulos, S. C. A., Viswanathan, R. and Bougoulias, D. K., (1989), "Optimal distributed decision fusion," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 25, pp. 761-765.
[20] Zhang, H., Schroepfer, C. and Elsayed E. A., (2006) "Sensor Thresholds in Port-of-Entry Inspection Systems," *Proceedings of the 12th ISSAT International Conference on Reliability and Quality in Design*, Chicago, Illinois, USA, August 3-5, 2006, pages 172-176.