RUTCOR
RESEARCH
REPORT

# DECEPTIVE DETECTION METHODS FOR OPTIMAL SECURITY WITH INADEQUATE BUDGETS: THE SCREENING POWER INDEX

Paul Kantor[a] Endre Boros[b]

ERROR! AUTOTEXT ENTRY NOT DEFINED.7
OCTOBER 2007

RUTCOR
Rutgers Center for
Operations Research
Rutgers University
640 Bartholomew Road
Piscataway, New Jersey
08854-8003
Telephone: 732-445-3804
Telefax: 732-445-5472
Email: rrr@rutcor.rutgers.edu
http://rutcor.rutgers.edu/~rrr

[a] SCILS, Rutgers University, kantor@scils.rutgers.edu
[b] RUTCOR, Rutgers University, boros@rutcor.rutgers.edu

RUTCOR RESEARCH REPORT

# DECEPTIVE DETECTION METHODS FOR OPTIMAL SECURITY WITH INADEQUATE BUDGETS: THE SCREENING POWER INDEX

Paul Kantor Endre Boros

**Abstract.** Detection of contraband depends on countermeasures, some of which involve examining cargo containers and/or their associated documents. Documents screening is the least expensive physical methods, such as gamma ray detection are more expensive, and definitive manual unpacking is most expensive. We cannot apply the full array of methods to all incoming cargoes, for budgetary reasons. We study the problem using principles of Game Theory, (and find that for optimal detection the available budget is allocated between screening and definitive unpacking using a mixture of strategies,) which maximizes detection rate and, further, serves to deceive opponents as to specific tests to which contraband will be subjected. This yields increases of as much as 100% in detection, with essentially no increase in inspection cost.

**Acknowledgements:** The support of the Office of Naval Research, Contract Number N0014-07-1-0299, of the NSF, under Contract Number 05-18543, The National Science Foundation and DHS under award CBET-0735910 and DHS Contract Number N000-14-07-1-0150 under ONR are all most gratefully acknowledged. Thanks to Fred Roberts, Alex Kogan, and Sheldon Jacobsohn for stimulating conversations.

# 1 Introduction:

**Inspection of containers involves inexpensive tests and expensive unpacking procedures.** Securing our ports and harbors is a major component of protecting our homeland. The possibility that terrorists might introduce catastrophic weapons, such as chemical or biological agents, or nuclear devices requires high vigilance. The definitive examination of each container (that is unpacking it and exploring it) is prohibitively expensive. Screening methods have been proposed for deciding which containers to unpack. Screening steps based on documents are relatively inexpensive. Other screens involve more expensive scanning or imaging tests. Estimates (Martonosi, Ortiz, & Willis, 2006) have been made of the costs of providing totally secure inspection. This goal seems out of reach even for container traffic through ports and harbors. When we consider the further problem of protecting cities and high value industrial, cultural and human targets, complete "screening" is clearly not feasible. Previous work has distinguished among "scans", "screens" and "unpacking". For clarity in presentation we consider only two steps here, which we take to be "testing" and "unpacking". Thus "testing" may include some examples of what are elsewhere called "scans" and "screens". Related work is reviewed in (Boros, Fedzhora, Kantor, Saeger, & Stroud, 2006; Martonosi et al., 2006). Saeger and Stroud (Stroud & Saeger, 2003) have addressed the problem as a binary decision tree problem. Madigan, Mittal and Roberts (Madigan, Mittal, & Roberts, 2007) implemented an efficient local search algorithm to identify the best tree, with a non-linear search included to select the best thresholds for each tree. Boros et al. (Boros et al., 2006) have formulated it as a very large linear programming problem, for the case of tests that produce continuous readings representing the risk, or some other characteristic. Related issues arise in the study of passenger screening for airport security, as discussed by [McLay et al. (L. McLay, Jacobson, & Kobza, 2005; L. A. McLay, Jacobson, & Kobza, 2006)], who have formulated integer programming models of the problem. Virta, Jacobson and Kobza (Virta, Jacobson, & Kobza, 2003) and Jacobson, Karnani and Kobza (Jacobson, Karnani, & Kobza, 2005) recognized the benefits of reducing the use of expensive passenger baggage screening at airports by employing much cheaper and faster prescreening. While they argue that the *effectiveness* of the prescreening procedure is the most dominant factor in the effectiveness of the entire system, we find that the best choice of prescreening methods depends on a complex *relation between effectiveness and cost*. We find that achieving the fill benefit may require an element of randomness in pre-selecting, prior to prescreening itself. All of these methods involve considerable computational sophistication, although they find ways to place the problem within the reach of modern optimization software and powerful contemporary computers. The present note addresses a remarkable simplification arising when the available budgets are simply inadequate to implement the optimal strategies found by these methods.

Remarkably, we find that the cost-effectiveness of any particular screening test may be summarized by a single number, the *Screening Power Index* (SPI). This index depends on the sensitivity and specificity (or operating characteristics) of the test, on known cost information, and on estimated probabilities. These numbers are, in reality, difficult to find, or closely held. However, once they are determined, it is easy to compute the index. The method can therefore be applied by operators of terminals, and sensitive information need not be shared with researchers. The SPI applies precisely when budget limitations dictate that *not all containers can be* *screened*. In this situation randomization strategies will improve the detection rate (Boros et al., 2006). Such an approach, in the terminology of game theory, is called a *mixed strategy*. In addition to its optimality properties, it has the virtue of deception, as, properly implemented, it thwarts an opponents efforts to circumvent it.

**Simple tests have fixed thresholds.** The results of this paper stem from the fact that screening test or indicators (such as the documentation, shipper, country of origin, factory of origin, etc.) are not perfect indicators of whether a container should be unpacked. This uncertainty is expressed in terms as the probability that a particular test will "flag" a container with contraband and the probability that it will flag an innocent container, as shown in Table 1. The number represented by $d$ is the "conditional probability of detection," while $f$ is the "conditional probability of false alarm" (and as probabilities, these numbers must be between 0 and 1).

**Table 1:** Probability that a container of the given type will be flagged.

| | Contraband | Innocent |
|---|---|---|
| Flagged | $d$ | $f$ |
| Not Flagged | $1-d$ | $1-f$ |

The information given in Table 1 may be summarized in a graph of the "ROC" or "(Radar|Receiver) Operating Characteristic" (Egan, 1975; Swets, 1973; Zhou, ) as shown in Figure 1(a). The case of a fixed threshold (cf. Figures 1(a), 1(b)), will clarify the key ideas.

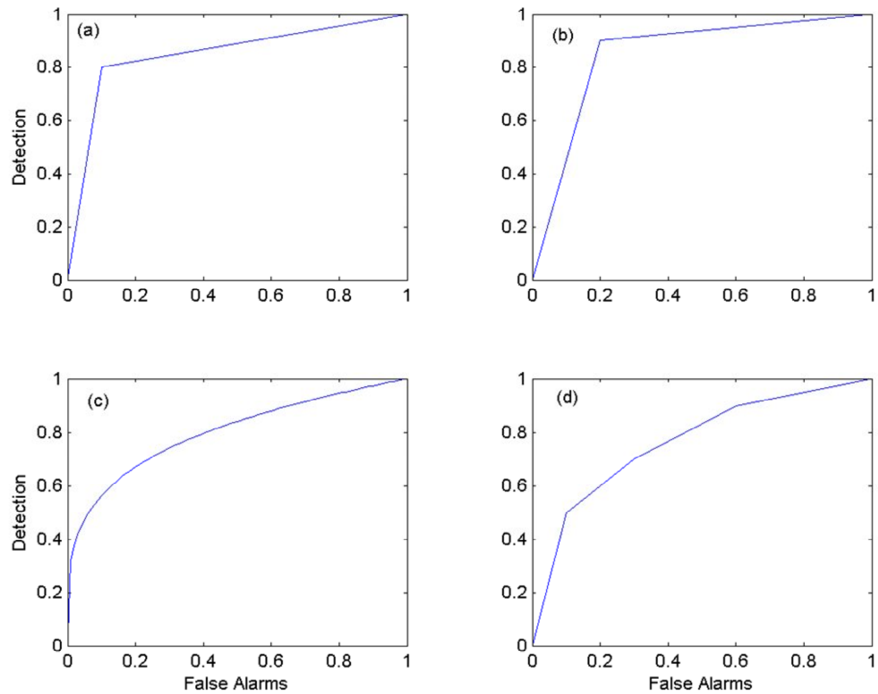


*Figure 1. The operating characteristic (ROC) shows the fraction of dangerous cargo detected, as a function of the fraction of harmless cargo that is tagged for costly inspection. If the system has a single threshold fixed in advance, the curve has two linear segments. Several models provide for a continuous dependence of detection on false alarms, as a threshold is varied. More complex discrete processes, such as document checks, may yield a curve with multiple linear segments. Operating characteristics are shown for sensors with: (a) (f=.1, d=.8) ; (b) (f=.2,d=.9); (c)* $d = f^{0.25}$ *; (d) {(d,f)}={(.1, .5); (.3, .7); (.6,.9)*

This case arises when the screening test has no adjustable parameters. Optimization depends on three costs: (a) the cost, $C_T$, of testing all the containers arriving in some agreed upon unit of time such as a month or a year; (b) The (much larger) cost $C_U$ cost of unpacking all those containers, to verify that they contain no contraband; (c) The costs of impact on commerce, $C_I$ of unpacking (and thus delaying) all the containers. We also need to know how many are expected to fall into each of the two categories: harmful (or contraband) and harmless. The *a priori* probability of a container being harmful is represented by $\pi$.

In principle, we would like to screen everything, and then open everything that is flagged as suspicious. The resulting total cost of operation $b^*$ (per budget period) can be computed from the above parameters, as given by Equation (1):

$$b_T{*} = C_T + C_U \pi d + (C_U + C_I)(1-\pi) f \qquad (1)$$

We refer to $b_T{*}$ as the "natural operating budget" required to fully utilize test T. The first term is the cost of screening everything by test T. The second term is the expected or average cost of unpacking the detected dangerous containers, and the third term is the (wasted) expense of unpacking and delaying harmless containers.

We remark that for a complete description we would need to add a fourth term "$+C_B\pi(1\text{-}d)$" representing the expected cost of not detecting a dangerous content, where $C_B$ denotes the cost of a smuggled nuclear device being blown up inside our country. The true value of $C_B$ is however highly speculative, and as the World Trade Center attacks of 2001 show, it is simply impossible to cast into a single number the true costs to the economy, to society and to our future. It is clear however that this fourth term would dominate our cost analysis, no matter how conservatively we try to estimate $C_B$. Thus, our only hope for minimizing the combined total cost is to maximize the detection rate d, which is the same goal, regardless of the specific enormous value of $C_B$. Therefore, in subsequent steps of our analysis we simply disregard the unknowably large fourth term, and focus on the "real" cost of inspection as expressed by (1) and on maximizing the detection rate *d*.

We further note that while we speak of a single "test T", in fact T could involve a number of different screening technologies in an array of a specific inspection policies. For all useful screening technologies we can assume that the screening cost is substantially smaller than the ultimate inspection involving the manual unpacking of a container. Thus, in our cost analysis (1) we separate the "test *T*" from the unpacking, i.e., we consider an inspection strategy as a two-component procedure, involving a test *T* to "flag" some of the containers, followed by unpacking of the "flagged" ones.

It is important that both the detection rate *d* and rate of false alarm $f$ as we argue below, can be viewed as functions of our real budget *b*. Our subsequent focus is on the realistic scenario when we have a powerful testing policy *T*, the cost $b_T^*$ of which exceeds our available budget b. In this situation we can try to stay below our budget b, by applying some components of our inspection policy to only some of the containers, producing various operating characteristics of achievable detection rate as function of our budget b. Some corresponding curves are shown in Figure 2, for the specific choices of $C_T = 1, C_U = 15, C_I = 5, \pi = .1$, and the performance *(f,d)=(0.1,0.8)* as given in Figure 1(a).

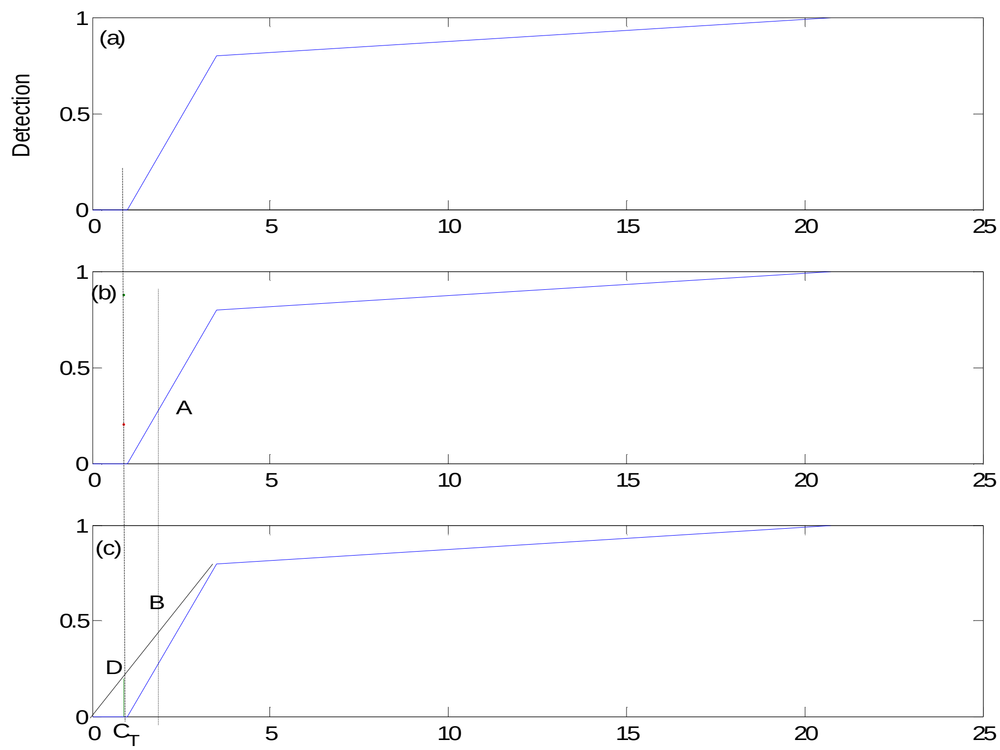


*Figure 2. Plots of the relation between the fraction of all dangerous cargoes detected, and the budget. Parameters are set at (f,d)=(0.1, 0.8), using the detection scheme of Figure 1(a). (a) For budgets less than $C_T$ there is no detection at all. (b) For budgets between $C_T$ and b* random selection of containers flagged by the test will produce the detection level shown by [A]. However, randomly selecting containers to be tested, and then opening all of the ones that are flagged, will produce a higher detection level shown by [B]. (c) This principle is even more effective when the budget [C] is less than $C_T$, and the mixed strategy produces substantial non-zero detection as shown by [D].*

One of the possibilities is to manually unpack only a fraction of the flagged containers. The corresponding achievable detection rate as a function of the real budget b is shown in Figure 2(a). The striking feature of this figure is that until the amount $C_T$ has been spent, there is *no detection at all* (i.e. for $b < C_T$ we cannot guarantee any detection). We have, however, better possibilities, using ideas from game theory. To provide a realistic scale, we draw on reports that current port operations in the US are said to open or examine between 6% and 9% of all incoming containers (Etheridge, 2006; Martonosi et al., 2006). With the assumed performance numbers, this corresponds to a real budget b somewhere between $C_T$ and $b_T*$. In this range, one may randomly select from among the flagged containers until the funds are exhausted. This brings performance to the point labeled A in Figure 2(b). For example, if we randomly examine half of the flagged containers, we will find, on average, half of the contraband contained in all the flagged containers. Note that randomization (rather than systematic rules) provides an essential element of deception and is key to the success of this strategy. If, for example, we only opened flagged containers at a particular terminal "on Mondays, Wednesdays and Fridays" the rule would eventually be discovered by our opponents, who would then concentrate their efforts on arriving on the other days.

It is possible to do even better, by adopting a mixed or randomized strategy from the very beginning. If we randomly select containers to be *screened*, we can, at the same budget, achieve a higher detection rate, corresponding to the point labeled B. This principle clearly applies with even more force to budgets less than $C_T$, where the naïve alternative would be to detect *nothing*. The principle is shown by the straight line in Figure 2(c), which connects the origin to the *operating point* $d^*(b_T^*)$. Points on this straight line represent the effectiveness and budget of *mixed screening strategies.* For example, the point marked A is 25% of the way from the origin to the optimal point. It represents a strategy in which only 25% of the containers are tested at all, while the others are untested. If the reading for a tested container falls above the threshold, it is unpacked. As is apparent from the geometry of Figure 2(c), the cost of this strategy is *25%*$b_T^*$, and the corresponding detection rate is *25%d**. Game theory also assures us that we cannot do better, if our budget *b* is only 25% of $b_T^*$.

**Tests themselves may be ranked using the screening power index**. Before estimating how much this approach might improve detection performance, let us see how *different* screening tests can be easily compared. Suppose there is an entirely different screening strategy (S), with its own ROC curve, as shown in Figure 1(b), and its own cost $C_S$. We can perform exactly the same calculation, and obtain the corresponding point $(d^*_S, b^*_S)$. The two lines connecting those points to the origin are strictly ordered (i.e. one is lower than the other). This ordering is determined by the ratios $\left\{d^*_X / b^*_X\right\}_{X=S,T}$, where *X* represents the particular screening method (see Figure 3(a)). This ratio, as the construction makes clear, depends on several factors: the cost of screening, the cost of unpacking and interruption to commerce, the performance of the test, and the prior estimate of the probability of a harmful container.

Without revealing any sensitive information to researchers, a terminal operator may readily calculate this critical ratio, which we call the "Screening Power Index" (SPI). This can be done for any method of screening whose cost and performance are known. Comparison of the SPI shows which method provides the highest level of protection, *for any given, but "inadequate" budget*. The choice among screening protocols is then guided in a rational way, by a calculation that is readily performed, and readily explained. This way of approaching the calculation makes the decision among screening methods, when budgets are inadequate, "as simple as possible, but not simpler".

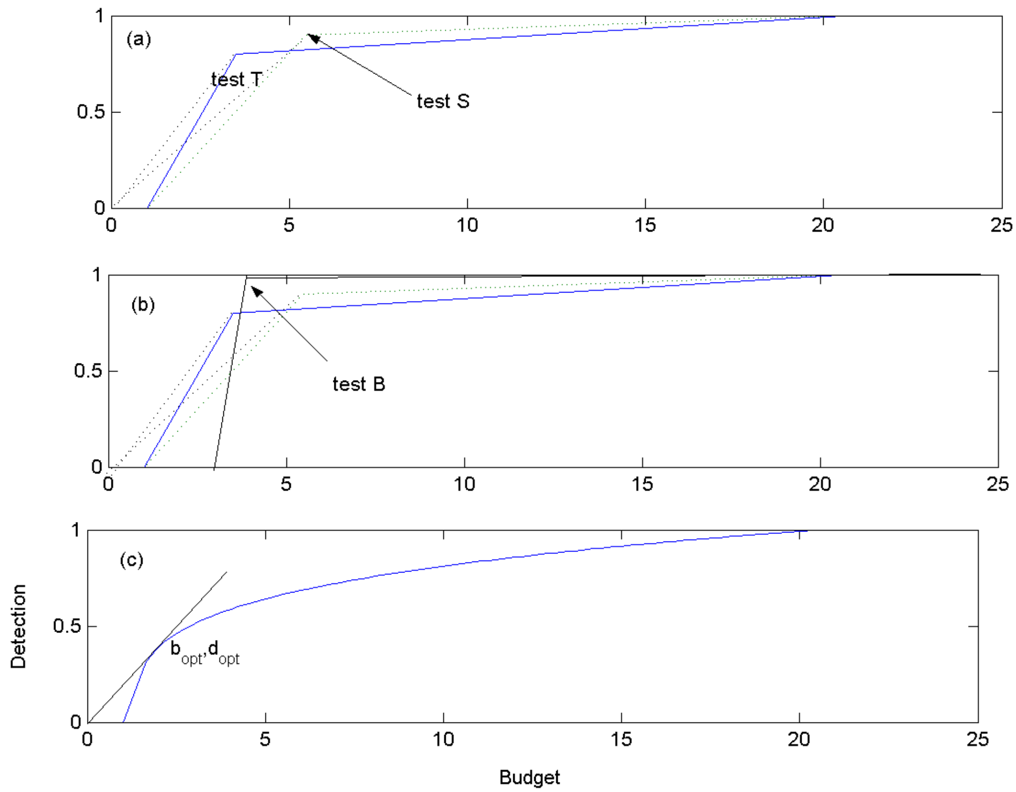


*Figure 3. Systems can be compared by examining the slope of the line from the origin to the "top" of the curve. The slope of the highest line represents the screening power index (SPI). (a) Test T is therefore better, for inadequate budgets, than is Test S. (b) But test B, which is much more expensive, is in fact better than either of them. (c) The same principle applies to operating characteristics such as the power law form shown here, but the optimal point is found either analytically, or by numerical search. The one dimension search is highly efficient in every case, as the SCI is a concave function of the threshold.*

As an example, with the values as shown in Figure 3(b), the screening method *B* is clearly preferred to the other two methods, even though its "natural" operating expense would be the highest of the three, and well beyond the available funds.

**For more complex tests, the threshold can be selected optimally**. For more complex screening methods (called "scanning" in some of the literature), such as those involving radiation detectors, both *d* and *f* can be decreased (or increased) by raising (or lowering) a parameter called the *threshold.* In the limits, one either unpacks none of the containers, or all of them. A typical continuous *f-d* relation is shown in Figure 1(c). It might be determined from physical principles (e.g. by calculation of how much radiation leaks from a container, at various frequencies, with various kinds of shielding).

An ROC with several linear portions may, on the other hand, result from empirical data (e.g. by learning what fraction of the containers coming from country X, and having documents Y, contain a contraband of interest). We are not able to work with real data in this report, but we illustrate the principles using model forms for the ROC. Figure 1(c) shows an ROC produced by assuming that signals follow an exponential distribution. Since the detection and false alarm rates may now depend on a threshold *t* we write them as *(f(t), d(t))*. The choices for *t* may vary smoothly (for example, click rates on a radiation detector) or may fall at a few discrete points (for example, combinations of document check information). Figure 1(d) shows a model example in which there are only a few possible threshold settings (the "discrete" case).

At any setting, the cost of operation (per budget period, or year) is calculated, as before, as the cost of screening plus the cost of unpacking, plus the cost of interruption to commerce.

$$b(t) = C_T + C_U \pi d(t) + (C_U + C_I)(1-\pi) f(t) \qquad (2)$$

Since the same parameter *t* determines both the detection rate *d(t)* and the budget required, *b(t ),* and each of them is monotonically dependent on the threshold, we can combine the calculations and draw a single curve relating *b* and *d.* Because the points *(d(t),f(t))* form a monotone increasing curve as a function of the threshold *t*, each possible cost value *b* is attained for a unique point *(d,f)* on this ROC curve). An example is shown in Figure 3(c). This curve gives precisely the information that a policy maker needs. The goal is to *minimize* the risk of missing something dangerous (that is to *maximize d*) and the constraint is the amount of money that can be spent. Note that there could be other constraints as well, in which case the same principles used here apply, but the calculation, while straight-forward (see, e.g. Boros et al. (Boros et al., 2006)), becomes multi-dimensional and cannot be easily represented by graphs.

To select a threshold, we choose the operating point so that the *Screening Power Index* is a maximum, defining the *optimal screening power point* on the *d(b)* curve, with coordinates $(b^{opt,} d^{opt})$. When the budget is less than the budget required to reach that operating point, we connect the origin that point by a straight line.
As before, points on the straight line represent the effectiveness of *mixed strategies.* The line is chosen so that no other line, from the origin to a point on the *d(b)*curve, has a greater rate of increase. We can express this mathematically by saying that $b^{opt}$ is the argument at the maximum of the Screening Power Index *d(b)/b*, where *b* is given by (2). Expressing both *d* and *f* as function *s* of *b* we get:

$$b^{opt} = \arg\max\{d(b) / \left(C_T + C_U \pi d(b) + (C_U + C_I)(1-\pi) f(b)\right)\} \qquad (3)$$

Note that the selection of this point requires consideration of all of the cost parameters, and the presumed prior probability of contraband. Note also that in all cases of interest, the operating point for a test with single fixed threshold is also the optimal operating point. (Otherwise, it would simply better to open containers at random!).

*In practice, an operating point is sometimes selected by convention (for example, that the rate of missing dangerous cargoes (false negatives) be equal to the rate of false positives). In other studies it is selected by hypothesizing the relative cost of false negatives and false positives. The* *points selected in these ways are unlikely to be the correct point* $(b^{opt}, d^{opt})$ *needed to optimize the effect of an inadequate budget.*

**Complex tests are compared using the screening power index.**. As before, an entirely different screening strategy (S) will have its own ROC curve, and its own cost $C_S$. Repeating the calculations yields the corresponding curve $d_S(b)$. When they are shown together we may see that the curves may cross more than once, so that neither is "absolutely better than the other". However, each has its own unique optimal point $b^*$ and the two lines connecting those points to the origin are strictly ordered as determined by the ratio $d_x(b^*)/b_x^*$ (where $X$ labels the particular screening strategy or test). Thus tests with an adjustable threshold can be separately tuned to each give its own best performance, and they can then be ranked to select the highest level of protection, for a given "inadequate" budget.

# 2 Impact of Effectiveness of Screening.

**Implications for actual port security are quite significant.** . For example, suppose that a screening method has a *d=.80 f=.20* performance, which is generally considered a good level. This means that the single operating point *(b*,d*)* will have coordinates $(b^*, d^*) = (C_T + 0.8C_U\pi + 0.2(C_U + C_I)(1-\pi), 0.8)$. Since $\pi$ is quite small for many serious threats we can neglect it for purposes of discussion. Then $(b^*, d^*) = (C_T + 0.2(C_U + C_I), 0.8)$. Now, with standard approaches, one screens every single container, and then unpacks until the budget is exhausted. If, for example, we are currently unpacking say 7% of all containers, two things may be happening. First, we may be adjusting our definition of "high threat" to conform to our budgetary capabilities. This may have been done by adjusting the thresholds applied, so that the operating point is somewhere around $(f, d) = (0.07, 0.28)$. In other words, 28% of the dangerous containers are being unpacked. However, with a mixed strategy we could achieve a higher level of detection. Let $C_T$=1 and the sum of unpacking and interruption costs, $C_I$+$C_U$=20. Then the mixed strategy achieves a performance given by

$$d_{optimal}(b_{now}) = 0.8/(1+20\times.20)\times(1+20\times.07) = 0.8/5\times2.4 = 0.384 \qquad (4)$$

This is an increase in detection, from 28% to 38% (that is, a nearly 35% greater detection rate) at absolutely no increase in cost!

**Multiple screens may be reduced to an optimal subset: k-out-of-N tests.**. Suppose that there are $N$ different document checks available, which each have their own performance characteristics $(f_i, d_i)\ i = 1,\cdots,N$. Each of the tests may flag a container for unpacking. The terminal operator should use all of this information to decide which ones to unpack. For simplicity, we examine the case in which all the tests are *stochastically independent*, and all have *the same performance* characteristic. In this case the probability that *k* of the *N* tests will flag the container depends only on the number *k* and not on *which particular tests* flagged the container. Since for independent tests the conditional probabilities can be computed as products, we have, for the overall test *k-out-of-N:*

$$f^{k-out-of-N} = \binom{N}{k} f^k (1-f)^{N-k}$$
$$d^{k-out-of-N} = \binom{N}{k} d^k (1-d)^{N-k} \quad (5)$$

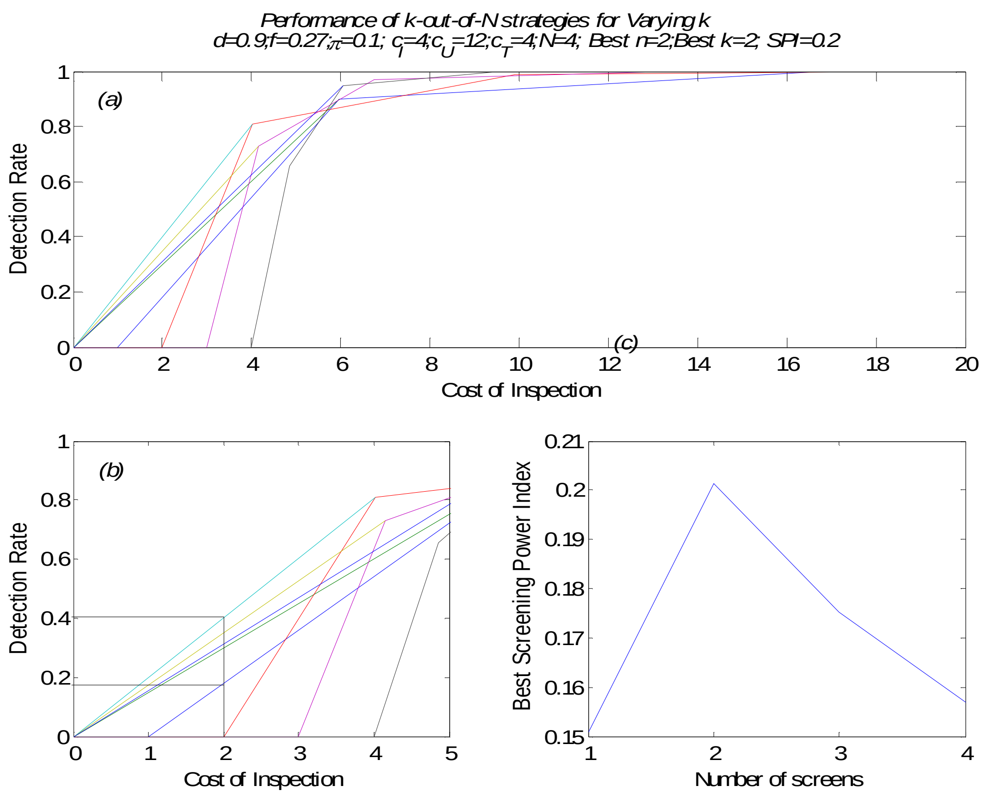


*Figure 4. We consider the situation where 4 independent screens, with the same operating characteristic, are available. The operator may choose to use 1,2, 3 or all 4 of the screens. For each of these choices (which corresponds to a curve in (a)), there is a specific optimal value of k, the number of positives needed to trigger an unpacking operation. (b) This is selected by comparing the tangent lines from the origin, shown in an expanded view. (c) The dependence of the best achievable SPI is shown as a function of the number of tests deployed. It has a clear maximum at N=2, reflecting the fact that using fewer tests would decrease detection by almost 30% and increasing it would provide no gain, but approximately a 15% decrease in performance.*

This is computed for the example values $C_T=1$; $c_U=12$; $c_I=4$; $f=0.27$; $d=0.9$; $\pi=0.1$; and is shown in Figure 4. What does this have to do with practical port security? It is known that at present some 6% to 9% of containers are “screened”, and this is usually described by saying “all the highly suspicious containers are checked". We conjecture that what happens, in practice, is that the threshold (in this case, the value of *k*) is set so that the resulting operations are possible with the available budget. We explore, by example, how much improvement may be possible with a mixed strategy.

We find that in fact, even though a number N of inexpensive screens may be available, it may be optimal to pay for only some smaller number of them, in order to achieve the optimal SPI. In this example (see Figure 4 caption for details) it is best to use just 2 of the 4 available screens.

**Parameter uncertainties will affect the decision process.** The technical exposition of the Screening Power Index is complete. However, the real world may be less cooperative than is assumed in our model. To begin with, there will be some uncertainty in the underlying performance characteristic, *(d(t),f(t))*. This can be explored using a variety of methods (some of which may simply be expert judgments) to come up with a band of *plausible ROC curves.* In addition, there is sure to be uncertainty about the value of the *a priori* probability $\pi$. This can also be factored into the analysis, and the result will finally be some *family* of *cost-performance* curves corresponding to each method of screening. Each such curve has an optimal point, and the resulting SPI values span some range.

Now, consider any two competing tests. In some cases the range of SPI values for one will always be higher than the range of SPI values for the other. In such cases, the choice is clear. In some cases the ranges will overlap. There are no universally accepted rules for making the decision in this case. There is some preference for the range that reaches higher, but that choice would require careful scrutiny of the specific assumptions about performance, costs, and the *a priori* probability, which make it reach these highest values. In a competitive economic or political environment, advocates can be expected to argue vigorously about these "edge cases". Because there is a rigorous theory underlying the computation of the SPI, however, it will be difficult to "game the decision". The underlying ROC arises from detailed technical or empirical data, and is hard to manipulate. Similarly, although *proposed* costs for military and security systems exhibit a degree of fantasy, insofar as there are any *real* performance data *(d,f)*, there must have been real systems built and operated, so that the costs should be known.

Finally, of course, it may be that the ranges of SPI factors for the two systems essentially coincide. In this case, the choice of systems may be determined by other principles. The first is the *deception principle*, used above. Insofar as one can conceal which of the two equivalent tests is in use, it becomes more difficult for an opponent to use countermeasures. The second is the *robustness principle*, which favors using more than one approach, in anticipation of the possibility that the opponent may discover ways reduce the effectiveness of any given system. From the political point of view one also notes the *distributional principle,* which favors having some part of the overall effort "*made in each congressional district".*

# 3 Conclusions

**Optimal strategies balance testing and unpacking to maximize detection with inadequate budgets**. In summary, for any screening test, or combinations of tests, there is a corresponding operating characteristic. This may be combined with known cost information, and estimated probabilities, to determine the curve of detection as a function of budget. The entire cost-detection performance curve may be summarized by a single number, the Screening Power Index, which corresponds to the optimal operating point on the curve. Finally, the available budget is balanced according to the optimal random strategy, by either testing or ignoring containers. The result can be substantial increases in the detection of contraband, with essentially no increase in costs for screening or inspection. It may be helpful to think of these mixed strategies as shifting a certain portion of the budget from screening to unpacking, resulting in a higher overall detection rate.

**To protect multiple ports extensions may be needed.** . To protect the entire nation, one also needs to know the *a priori* probability that containers entering via different ports contain contraband. This prior probability, $\pi_P$, will affect the best detection strategy. If our assessment of these prior probabilities could be kept confidential, and could not reasonably be worked out by an opponent, then the optimal strategy concentrates our resources first on the port with the highest *a priori* risk. This is analogous to the problem of assigning risk reduction funds to cities or to states.

On the other hand, if this ranking could reasonably be estimated by an opponent, then the opponent would concentrate efforts precisely where we have estimated them to be least likely. The resulting situation can be treated as "zero-sum" game. That is, the opponent seeks to maximize the harm to us, and we seek to minimize the very same harm. These extensions of the problem will be discussed elsewhere.

# 4 References


Boros, E., Fedzhora, L., Kantor, P., Saeger, K., & Stroud, P. (2006). *Large scale LP model for finding optimal container inspection strategies* No. RRR 26-2006 (Submitted to NRLQ))

Egan, J. P. (1975). *Signal detection theory and ROC analysis*. New York: Academic Press.

Etheridge, B. (2006). Etheridge calls for 100 percent container inspection at U.S. ports. , May 4, 2006

Jacobson, S. H., Karnani, T., & Kobza, J. E. (2005). Assessing the impact of deterrence on aviation checked baggage screening strategies. *International Journal of Risk Assessment and Management, 5*(1), 1-15.

Madigan, D., Mittal, S., & Roberts, F. (2007). Sequential decisions-making algorithms for port-of-entry inspection: Overcoming computational challenges. *Proceedings IEEE Intelligence and Security Informatics,* New Brunswick, NJ. *, ISI2007* 1-7.

Martonosi, S. E., Ortiz, D. S., & Willis, H. H. (2006). *Evaluating the viability of 100 per cent container inspection at america's ports* No. RP 1220)

McLay, L., Jacobson, S., & Kobza, J. (2005). Making skies safer. *OR MS TODAY, 32*(5), 24.

McLay, L. A., Jacobson, S. H., & Kobza, J. E. (2006). Multilevel passenger screening strategies for aviation security systems. *Naval Research Logistics,*

Stroud, P. D., & Saeger, K. J. (2003). Enumeration of increasing boolean expressions and alternative digraph implementations for diagnostic applications. *Proceedings of Computer, Communication and Control Technologies,* Orlando, FL. *, IV* 328-333.

Swets, J. (1973). The relative operating characteristic in psychology. *Science, 182*(December), 990-1000.

Virta, J., Jacobson, S. H., & Kobza, J. E. (2003). Analyzing the cost of screening selectee and non-selectee baggage. *Risk Analysis, 23*(5), 897-908.

Zhou, K. *Http://splweb.bwh.harvard.edu:8000/pages/ppl/zou/roc.html*