Science and Global Security, 15:211–233, 2007

ISSN: 0892-9882 print / 1547-7800 online
DOI: 10.1080/08929880701715100

Routledge
Taylor & Francis Group

The Optimal Spatiotemporal Deployment of Radiation Portal Monitors Can Improve Nuclear Detection at Overseas Ports

Lawrence M. Wein,[1] Yifan Liu,[2] Zheng Cao,[3] and Stephen E. Flynn[4]

[1]Graduate School of Business, Stanford University, Stanford, CA, USA
[2]Department of Systems Engineering and Operations Research, George Mason University, Fairfax, VA, USA
[3]Department of Physics, Stanford University, Stanford, CA, USA
[4]Council on Foreign Relations, New York, NY, USA

Radiation portal monitors are starting to be deployed at overseas ports to prevent nuclear weapons from entering the U.S. in a shipping container. Current designs have containers on trucks passing through a portal monitor at approximately 10 mph, before being routed to one of several lanes at the port's front gate for a driver identification check. For a fixed cost of testing, which consists of the costs of radiation portal monitors plus offsite x-ray and possibly manual testing of containers generating a false radiation alarm that cannot be resolved by gamma-ray imaging, the neutron detection limits of the current design are compared with those of two other designs that do not affect truck congestion at the front gate. For a wide range of budgets, it is optimal to have six monitors in each lane that simultaneously test a truck while it is being processed at the front gate. This design is robust against the location (within the container) of the weapon and reduces the detection limit (relative to the current design) by approximately a factor of three (although the accuracy of this value is limited by the lack of publicly available data) for practical budgets, which is enough to offset some shielding for a plutonium weapon, but insufficient to detect an uranium weapon.

Received 13 September 2005; accepted 25 September 2007.
This research was supported by the Center for Social Innovation, Graduate School of Business, Stanford University. The authors thank Terry Gibson of SAIC for sharing information about their detection equipment and the HKCTOA pilot project, and thank the reviewer for comments that significantly improved the article. From 2003 to 2005, Stephen Flynn served as a consultant to Science Applications International Corporation in support of the development and deployment of the Integrated Container Inspection System (ICIS) pilot project in Hong Kong.
Address correspondence to Lawrence M. Wein, Graduate School of Business, 518 Memorial Way, Stanford University, Stanford, CA 94305-5015. E-mail: lwein@stanford.edu

INTRODUCTION

Preventing nuclear materials and weapons from being smuggled into, or through, U.S. seaports was a key topic in the 2004 U.S. Presidential Debates, and has become one of the highest priorities of the Department of Homeland Security. This task requires that containers be tested at overseas ports for long enough to detect fissile material such as highly enriched uranium, while not increasing congestion at these ports. An important step toward this goal is a pilot demonstration project sponsored by the Hong Kong Container Terminal Operators Association (HKCTOA) that is being performed in two marine container terminals at the world's second-busiest port. The project uses Science Applications International Corporation's (SAIC) Integrated Container Inspection System (ICIS), which consists of three elements:[1] an Exploranium Radiation Portal Monitor (RPM), which measures neutron and gamma-ray emissions as a container truck drives through a portal, the VACIS gamma-ray imaging system, which takes a two-dimensional image of the container with the hope of identifying dense material that can shield the emissions, and the Optical Character Recognition System, which records the unique container identification number that is located on each container. The radiographic profile and gamma image are stored in a computer file for subsequent analysis and review. Because the human analysis of each scan requires approximately two minutes, resources are not currently available to also perform 100% scan analysis. Legal shipments generate numerous gamma-ray alarms at the RPM[2] and consequently the only scans that are analyzed are those corresponding to containers that trigger a neutron alarm; although not part of the HKCTOA project, U.S. Customs and Border Protection could also review the images from containers it targets as suspicious using its Automated Targeting System (ATS). ATS is an expert system that associates data provided in a cargo manifest with intelligence and other potential indicators that suggest that a container has a heightened risk of having been compromised by criminals or terrorists. If scan analysts cannot make a satisfactory determination that the contents of a container pose no risk, they can require that the container be removed from the marine terminal and transported to a customs inspection facility where it is scanned again using a high-energy x-ray imager. In the event that these results still prove inconclusive, the final recourse is for inspectors to open the container and physically remove and examine its contents.

The HKCTOA project is successfully proving that 100% passive radiation detection plus image scanning can be performed at the world's second largest port without disrupting the normal flow of truck traffic entering the marine terminals; in fact, this approach is being implemented on a small scale with 6 overseas ports, which cover 7% of in-bound cargo.[3] However, the HKCTOA project does not directly assess the other vital requirement for a successful system: whether fissile material can be reliably detected by RPMs. It is generally

believed that RPMs are capable of detecting moderately shielded plutonium but incapable of detecting highly enriched uranium.[4] A recent mathematical study[5] suggests that longer RPM testing times could improve the likelihood of detecting highly enriched uranium, but at the expense of increased congestion. In this study, these ideas are taken a step further, by comparing the spatial design used in the HKCTOA pilot project to two other designs that do not affect congestion at the front gate, but make more efficient use of the RPMs. More generally, the goal is to identify a RPM deployment strategy that minimizes the detection limit, defined here as the neutron emissions level (in neutrons/sec) at which detection can be achieved with probability 0.95, subject to a fixed budget for RPM and offsite testing. The optimal deployment involves a complex trade-off of exploiting time and space. The time aspect is straightforward because the signal-to-noise ratio increases as the testing time increases. The spatial aspect stems from the facts that a terrorist can place a weapon anywhere in the 40-ft container and the number of neutrons detected by a RPM is inversely proportional to the square of the distance between the weapon and the RPM. Hence, for a given budget that allows for multiple RPMs, it is not clear whether it is preferable to have several RPMs simultaneously monitoring a single container (in the hope that one of the RPMs is close to the weapon) for a certain amount of time, or to have one RPM monitoring a container but with an increased testing time.

THE MODEL

The mathematical model, which is formulated in the Appendix, includes a queueing model at the front gate, a neutron emissions and detection model, and a cost model.

The Queueing Model

Queueing theory is the mathematical study of waiting lines.[6] In a queueing model, "customers" arrive randomly to a set of "servers," who provide the customers service, typically one at a time (each customer's service time is also random) and in a first-come first-served fashion. The goal of queueing theory is to predict the amount of congestion in the system (e.g., the total number of customers in the system) under steady-state conditions (i.e., after the stochastic system has been operating for a long period of time) in terms of the statistical characteristics of the arrival and service processes.

In the queueing model (Figure 1a), trucks play the role of customers. They arrive to the terminal according to a Poisson process (i.e., the time between consecutive arrivals is an exponential random variable) at a rate of 3.88/min (all parameter values in the model are estimated in the Appendix and appear in Table A1 of the Appendix), and enter a single waiting line that feeds $m = 4$

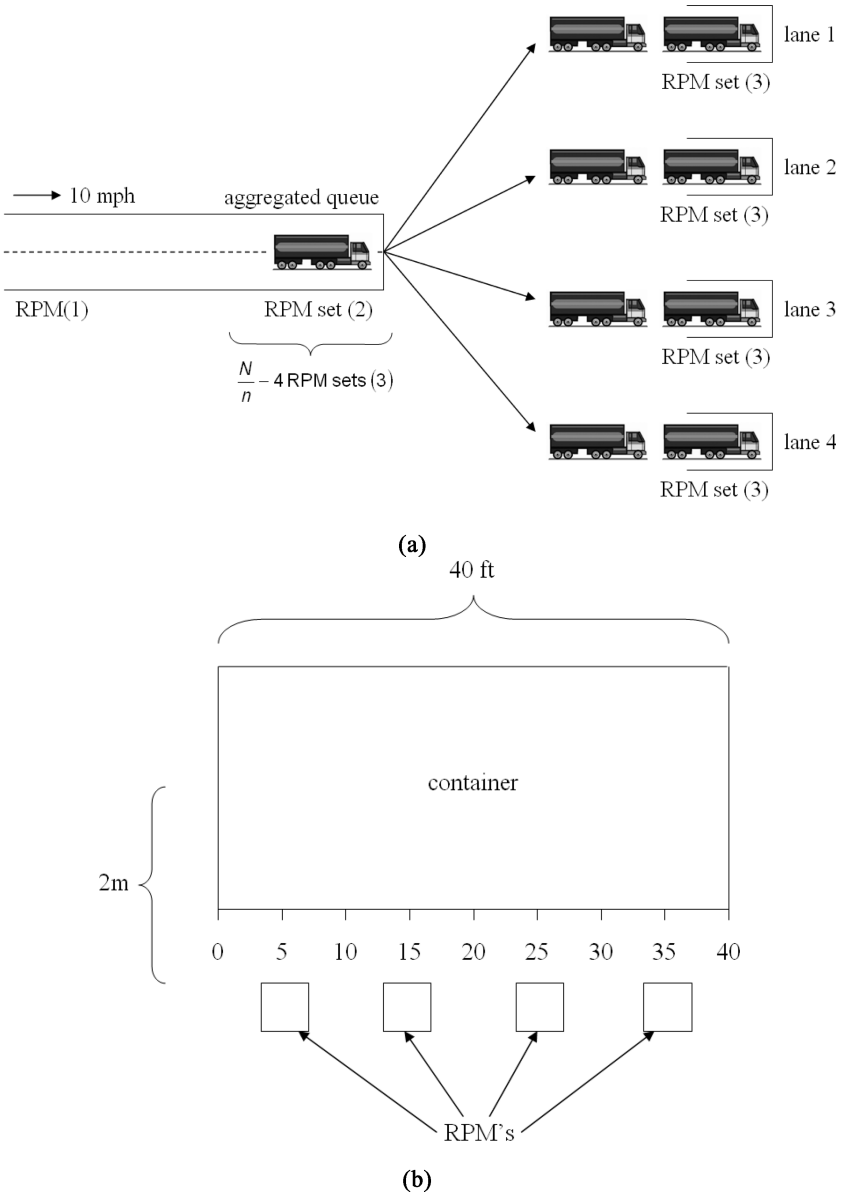

Figure 1: **(a)** The three spatial deployment designs. A single aggregated queue feeds four lanes, and each lane has room for two container trucks. The numbers in parentheses refer to the numbers of the designs that have RPMs (or sets of RPMs) at specific locations: Design 1 has a single RPM far from the front gate, Design 2 has a single set of n RPMs at the front position of the aggregated queue, and Design 3 has N RPMs in total, which consist of a set of n RPMs at the front position in each of the 4 lanes, and a set of n RPMs at each of the front $\frac{N}{n} - 4$ positions of the aggregated queue. **(b)** The n RPMs within each set in Designs 2 and 3 are positioned so that they each monitor an interval (measured along the length of the 40-ft container) of length $\frac{40}{n}$ ft, and each is located at the center of its interval, as shown in the $n = 4$ case.

lanes at the front gate of the terminal. A Poisson arrival process is valid if each customer arrives independently of one another,[8] which is the case here, aside from the batching of traffic that occurs due to traffic lights on the roads leading directly to the terminal; based on empirical data, there is a long history of using queueing models with Poisson arrivals to model vehicular traffic at toll booths (e.g., Ref. [9]). To avoid server idling in these lanes, up to $l = 2$ container trucks are allowed to queue in each lane, and if there are more than $lm = 8$ trucks in total, then the additional trucks wait in a single aggregated queue. Each lane has a human server that verifies the identity of the truck's driver and tells the driver where in the shipyard to deposit the container. This service time is random with mean 60 s and coefficient of variation (i.e., standard deviation divided by the mean) 0.5. Let Q represent the steady-state number of trucks in the system (in queue plus in service). An analysis in §A of the Appendix shows that Q can be approximated by a geometric random variable with a mean of 24 trucks, which leads to an average of 6 trucks for each of the 4 lanes. This level of congestion is consistent with the observations of the HKCTOA project. Indeed, if the demand-to-capacity ratio of $3.88/4 = 0.97$ was to increase above 1.0, then the queue length would dramatically increase to an impractical level,[6] and there is little economic incentive for terminal operators to invest in significant excess capacity (i.e., additional lanes) that would reduce queues to negligible levels.

The RPM Model

Because of the lack of specificity of gamma-ray measurements from RPMs (many legal shipments emit a large number of gamma rays),[2] the emissions and detection model, which specifies a RPM's detection probability and its false positive probability, is restricted to neutrons. Emissions can emanate from three sources: the weapon, the contents of a typical container that contains no weapons, and the background level in the absence of containers. However, because 163,000 roadside field tests for container trucks resulted in no neutron alarms,[7] neutron emissions from the contents of a weaponless container are ignored, and only emissions from the background and a weapon are considered.

First, the neutron emissions detected by a RPM due to testing a weapon-free container are described. In the RPM designs, which are constructed later, a truck sometimes drives through a RPM and at other times it remains in place at the RPM. The number of neutrons detected at a RPM is a Poisson random variable with mean $A\epsilon b(t^d + t^p)$, where $A = 0.3$ m^2 is the area of a neutron detector,[10] $\epsilon = 0.14$ is the efficiency of a neutron detector,[10] b is the background rate, t^d is the drive-through time and t^p is the in-place time. To capture the fact that the background rate varies with solar magnetic activity and magnetospheric storms, time of day, and changes in the local environment, b itself is assumed to be a random variable. More specifically, using data from extensive controlled experiments[7] and the mean background neutron rate of

50 neutrons/m^2·sec in Ref. [10], b is taken to be a lognormal random variable (i.e., $\ln b$ is a normal random variable) with median 43.1 neutrons/m^2·sec and dispersal factor 1.73 (implying that 95% of the time, the background rate is between $\frac{43.1}{1.73^2} = 14.4$ and $1.73^2(43.1) = 129.0$ neutrons/m^2·sec). Hence, the number of neutrons detected at a RPM is actually a mixture of Poissons,[11] and its probability mass function is given in Equation 3 of the Appendix. It is assumed that for a given container, the same realization of the random variable b is used for all RPMs testing it because the background neutron rate is not likely to change substantially over the scale of time that a container is in the system.

For a weaponized container (i.e., a container containing a weapon), the neutron emissions detected by a RPM is a Poisson random variable with mean $A\epsilon b(t^d + t^p) + W$, where W respresents the neutron emissions at the RPM due to the weapon. As with the background term, the weapon term W has an in-place component and a drive-through component. If the neutron source of the weapon is S, which is allowed to vary in the model, then the in-place term is $A\epsilon St^p/4\pi d^2$,[10] where d is the distance between the RPM and the weapon, which (as explained later) depends on both the the spatial deployment and the weapon location. The drive-through term is derived in Equation 6 of the Appendix and incorporates the fact that the distance between the weapon and the RPM changes as the truck drives through the RPM.

The Three Designs

The three designs are illustrated in Figures 1a and 1b. In Design 1, which is employed in the HKCTOA pilot project, the RPM is located approximately 75 m from the front gate, and trucks pass through it at 10 mph. Because trucks do not sit in place at a RPM in Design 1, there is no need to consider deploying a set of RPMs to simultaneously monitor a container. In contrast, in Designs 2 and 3, there is a set of n RPMs that simultaneously monitor a container at a given position in the queue. Because each RPM requires approximately 6 ft of space as measured along the length of the 40-ft container, the decision variable n is restricted to be an integer value between 1 and 6. The n RPMs within a set are positioned so that they each monitor an interval (measured along the length of the 40-ft container) of length $\frac{40}{n}$ ft, and each is located at the center of its interval (Figure 1b).

In Design 2, a single set of n RPMs is located at the front position of the aggregated queue. If an arriving truck finds less than 8 trucks in the system, then he drives through the set of RPMs at 10 mph and immediately enters one of the 4 lanes. If an arriving truck finds 8 or more trucks in the system, then he is in place at the set of RPMs for 15 s (on average, trucks in the aggregated queue move forward one position every 15 s because there are 4 trucks in service, each with a mean service time of 60 s) and his drive-through speed is 5 mph (because he needs to come to a full stop at the set of RPMs).

Design 3 has N RPMs in total, which are deployed as $\frac{N}{n}$ sets of n RPMs. One set is located at the front queue position in each of the 4 lanes, and the remaining $\frac{N}{n} - 4$ sets are located at the front $\frac{N}{n} - 4$ positions in the aggregated queue, where the decision variables N and n are chosen so that $\frac{N}{n} - 4$ is a nonnegative integer. Arriving trucks drive at 10 mph until they reach the end of the queue. As in Design 2, if there are less than 8 trucks in the system, the arriving truck drives through all the RPMs in the aggregated queue at 10 mph and then immediately enters one of the 4 lanes, where it resides at the set of RPMs in this lane for 60 s and drive through it at 5 mph. At the other extreme, if there are more than $\frac{N}{n} + 4$ trucks in queue (which is the 8 trucks in the 4 lanes plus the $\frac{N}{n} - 4$ positions in the aggregated queue), then the truck will initially stop in the part of the aggregated queue where there are no RPMs, and then will subsequently drive through the $\frac{N}{n} - 4$ sets of RPMs at 5 mph and will stay in place there for 15 s, and finally will reside for 60 s (and drive through at 5 mph) at the set of RPMs in one of the 4 lanes. In the intermediate case (which is specified precisely in the Appendix) in which the queue length upon arrival is between 8 and $\frac{N}{n} + 4$ trucks, the arriving truck will drive through some sets of RPMs in the aggregated queue at 10 mph without residing in place, and will drive through other sets of RPMs in the aggregated queue at 5 mph and will stay in place for 15 s, before entering one of the 4 lanes, where it will stay in place for 60 s and drive through a RPM at 5 mph.

The Weapon Location

During the time that a container is in place at a set of RPMs, the distance d between the weapon and a RPM depends upon the locations (as measured along the 40-ft dimension of the container length) of the weapon and the RPM, as well as the minimum distance (taken to be 2 m)[7] between the weapon and the RPM as the truck drives through the RPM (Figure 1b). Two simplifying assumptions are made concerning the distance d between the weapon and the RPM. First, the weapon and the RPM's sensor are assumed to be at the same height so that the model can be constructed in two spatial dimensions. Second, the weapon is placed in the middle of the container widthwise (Figure 1b). Actual RPMs have sensors on both sides of the container. Given that the inverse square function (the emissions measured from a distance d is proportional to d^{-2} [Ref. 10]) is convex, if the readings from the two sensors are summed, then it is indeed optimal (in the sense of avoiding detection) for a terrorist to place the weapon in the middle of the container widthwise. Hence, although Figure 1b shows sensors on only one side of the container, one can view the model as summing the detected emissions from both sides of the RPM.

The weapon location is assumed to be a random variable that is uniformly located along the 40-ft length, and then choose the design variables to minimize the mean detection limit, where the expectation is taken over this uniform

random variable. To assess the robustness of Designs 2 and 3, the detection limits are also computed when the weapon is placed as close to (i.e., the best case), and as far from (i.e., the worst case), the nearest RPM. In the best case, the weapon is located such that when the truck is in place, the weapon and one of the RPMs are perfectly aligned lengthwise and the distance between the weapon and one of the RPMs in the set is 2 m. In the worst case, the weapon is located such that when the truck is in place, the lengthwise distance between the weapon and the closest RPM in the set is at its maximum, which is $\frac{40}{2n}$ ft, and the distance between the weapon and the closest RPM is $\sqrt{(2\text{ m})^2 + (\frac{40}{2n}\text{ ft})^2}$.

The RPM Threshold Levels

For Design 1, a single random emissions reading is obtained—call it X if it comes from a weapon-free container and Y if it comes from a weaponized container—as a truck drives through a single RPM. If $\bar{n}$ denotes the neutron threshold level above which an alarm is generated, then the detection probability for a given source S is $d(S) = P(Y > \bar{n})$. The detection limit is defined to be the source term S such that the detection probability $d(S)$ equals 0.95. The false positive probability is $P(X > \bar{n})$.

It is not as straightforward to decide when to generate an alarm in Designs 2 and 3 for two reasons: multiple emissions readings are obtained and the amount of testing time (both drive-through and in-place) depends on the number of trucks in the system at the time of a truck arrival. To keep the complexity manageable, a single threshold is still used, which is based on the multiple emissions readings. However, because the queue length upon arrival is observable, this threshold is allowed to take on a different value for each possible value of the queue length Q; in particular, the threshold value when $Q = k$ is denoted by $\bar{n}_k$.

For Design 2, emissions readings are obtained from each of n RPMs in a set (let the random variables be X_j and Y_j for $j = 1, \ldots, n$), and one of these RPMs will be closest to the weapon. Hence, a single threshold is again used (but now it can have a different value for each possible value of the queue length Q), but based on $\max_{j=1,\ldots,n} X_j$ and $\max_{j=1,\ldots,n} Y_j$, so that the detection probability and false positive probability when $Q = k$ are $P(\max_{j=1,\ldots,n} Y_j > \bar{n})$ and $P(\max_{j=1,\ldots,n} X_j > \bar{n}_k)$. Note that because the testing time in Design 2 only depends on whether $Q \leq 7$ or $Q \geq 8$, only two different values for the threshold are needed. In each of these two cases (i.e., $Q \leq 7$ and $Q \geq 8$), the detection limit is computed to be the source term that gives a 0.95 detection probability, and then the mean detection limit is found, which is the probability that $Q \leq 7$ times the detection limit when $Q \leq 7$, plus the probability that $Q \geq 8$ times the detection limit when $Q \geq 8$. Similarly, the overall false positive probability is the probability that $Q \leq 7$ times the false positive probability when $Q \leq 7$, plus the probability that $Q \geq 8$ times the false positive probability when $Q \geq 8$.

In Design 3, emissions readings are obtained from the RPM in position j (within the set) in set i for $j = 1, \ldots, n$ and $i = 1, \ldots, \frac{N}{n} - 3$ (i.e., the set in the lane plus the $\frac{N}{n} - 4$ sets in the aggregated queue). Let the emissions readings for the weapon-free container and the weaponized container by these RPMs be the random variables X_{ij} and Y_{ij}, respectively. Because the weapon does not move during its sojourn through the testing system, for all sets the weapon will be closest to one of the n positions within a set; that is, for some value $j = 1, \ldots, n$, let us call it j^*, one would expect (ignoring statistical variability) $Y_{ij^*} > Y_{ij}$ for all $j \neq j^*$ and for each $i = 1, \ldots, \frac{N}{n} - 3$. Hence, the emissions readings are summed over i (i.e., over the various sets) for each specific position in the set, and then the maximum over position j is found, which is likely to be achieved at position j^*. That is, a single threshold (that varies with the value of the queue length Q) is again used, but based on $\max_{j=1,\ldots,n} \sum_{i=1}^{\frac{N}{n}-3} X_j$ and $\max_{j=1,\ldots,n} \sum_{i=1}^{\frac{N}{n}-3} Y_j$, so that the detection probability and false positive probability when $Q = k$ are $P(\max_{j=1,\ldots,n} \sum_{i=1}^{\frac{N}{n}-3} Y_j > \bar{n}_k)$ and $P(\max_{j=1,\ldots,n} \sum_{i=1}^{\frac{N}{n}-3} X_j > \bar{n}_k)$. The mean detection limit and the overall false positive probability are computed by taking the expectations of these quantities with respect to the probability mass function of the queue length Q.

For Design 3, because the testing time takes on one value for $Q \leq 7$ and takes on a second value for $Q \geq \frac{N}{n} + 3$ (in this case, a truck resides in-place at all sets of RPMs in the aggregated queue), the thresholds $\bar{n}_k$ are chosen so that they take on one value for $k \leq 7$ and a second value for $k \geq \frac{N}{n} + 3$. To ease the computational burden for Design 3, for $k = 8, \ldots, \frac{N}{n} + 2$, the threshold $\bar{n}_k$ is forced to be such that the corresponding false positive probability when the queue length equals k (there is a one-to-one correspondence between the threshold $\bar{n}_k$ and the false positive probability when the queue length equals k, for any given values of N and n) is a linear interpolation between the false positive probabilities when $k = 7$ and $k = \frac{N}{n} + 3$.

Costs

To calculate the testing costs, it is assumed that the only VACIS scans that are analyzed are those that fail passive testing, plus an additional fraction $u = 0.05$ of untrusted containers that are flagged by the ATS,[12] which more accurately captures how this testing system might be used by U.S. Customs and Border Protection. It is assumed that the VACIS test correctly diagnoses the false positives or the untrusted containers with probability $p_v = 0.95$, and that a fraction $1 - p_v$ of the containers undergoing scan analysis need to undergo further testing offsite.

The expected total annual cost has two components, the annualized cost of each portal monitor, and the cost of offsite testing. The cost of scan analysis, which takes only two minutes, is ignored because it is dominated by the offsite testing cost, which can take several hours. Each RPM costs \$100,000[1] and it is annualized over a five-year horizon,[5] so that the annual cost per RPM is

$c_p = \$20,000$. It is assumed that the same amount of labor is required to run the RPMs for all three designs: one employee may be needed to direct traffic and one employee can monitor test alarms.

To compute the mean offsite testing costs, it is assumed for lack of data that a high-energy x-ray provides conclusive results for $p_x = 0.5$ of the containers it scans, and the others need to be opened manually. The current charge to importers for offsite testing in the HKCTOA Project is \$200–300 if only a high-energy x-ray is needed, and \$1500–2000 if the container also needs to be unladed. The x-ray cost is taken to be $c_x = \$250$ and the additional cost of a manual test is $c_m = 1500$. Hence, the mean offsite testing cost per container that undergoes scan analysis is $c_o = (1 - p_v)[c_x + (1 - p_x)c_m] = \$50$.

RESULTS

Each design is characterized by its mean detection limit versus cost curve, under the assumption that the weapon location is uniformly distributed along the 40-ft length of the container. The neutron threshold level $\bar{n}$ is varied to generate these curves for Design 1, which is the currently implemented design. For each fixed cost, the values of the set size n and the thresholds $\bar{n}_k$ for Design 2, or the values of N, n, and $\bar{n}_k$ for Design 3, are chosen to minimize the mean detection limit subject to a constraint on the annual cost. By solving these optimization problems for various values of the cost, the detection limit versus cost curves are generated for Designs 2 and 3. For the optimal versions of Designs 2 and 3, the detection limit versus cost curves are also computed when the weapon is in its best-case location and its worst-case location.

Before comparing the designs, the annual budget values on the horizontal axis in Figure 2 are put into perspective. The annual cost with a single RPM is \$5.12 M, \$5.22 M, and \$6.09 M if the false positive rate is 0, 10^{-3}, and 10^{-2}, respectively. That is, the fixed annual cost of testing the untrusted containers flagged by the ATS is approximately \$5.1 M, and the annual cost when a single RPM is deployed is typically in the range of \$5.2 M–6.1 M.

As shown in Figure 2, Design 2 is the best of the three designs for annual budgets below \$5.4 M and Design 3 is the best design for budgets over \$5.4 M; the nonsmooth behavior of the Design 1 curve in Figure 2b is because the neutron threshold level takes on discrete values that are much smaller than those in Designs 2 and 3. For annual budgets between \$5.6 M and \$10 M, the detection limit for Design 3 is approximately a factor of three less than the detection limit achieved by Design 1. There are decreasing marginal returns to investments in all three designs; for example, for Design 3, the detection limit can be reduced from 100 k to 20 k neutrons/sec by increasing the annual budget from \$5.2 M to \$5.4 M, from 20 k to 11 k by increasing the budget from \$5.4 M to \$6 M, and from 11 k to 6 k by increasing the budget from \$6 M to \$10 M, but over \$30 M/yr would be needed to lower the detection limit to 3 k neutrons/sec.

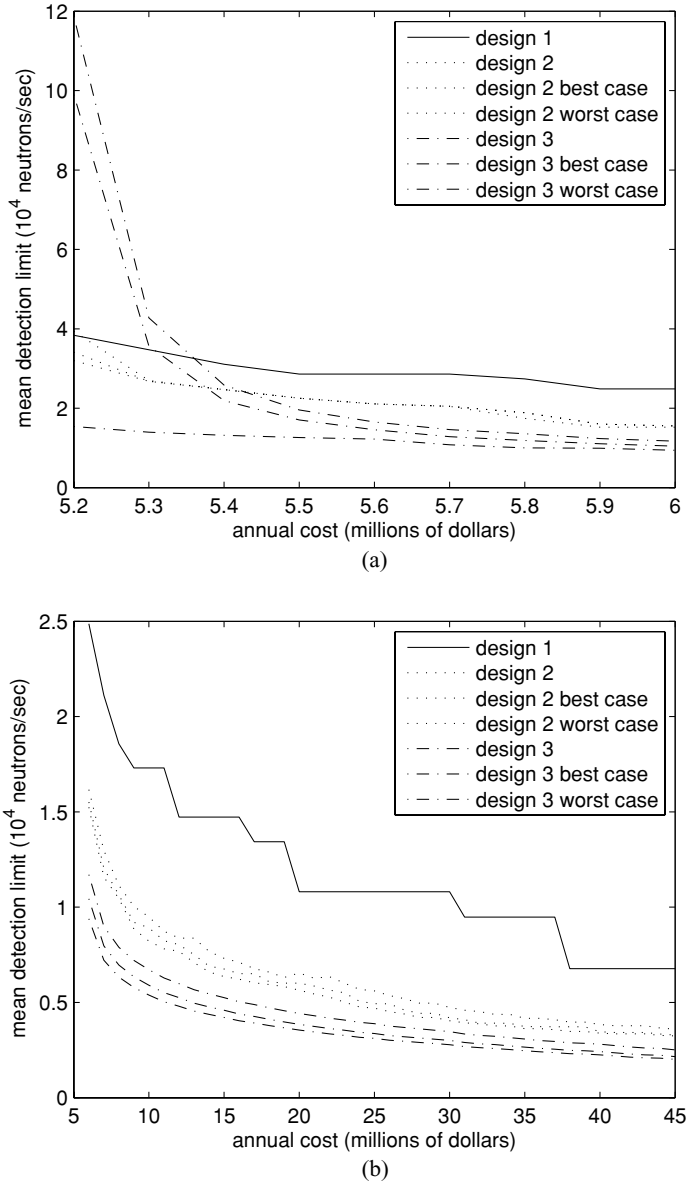

Figure 2: The mean detection limit vs. annual cost for the three designs. The worst-case and best-case detection limits for the optimal versions of Designs 2 and 3 are also given. The detection limit is the source emissions for which the detection probability is 0.95. The annual budget is (a) between \$5.2 M and \$6 M; and (b) above \$6 M.

The emissions rates of a former Soviet nuclear warhead made of plutonium and uranium are 400 k neutrons/sec and 30 neutrons/sec, respectively.[10] Hence. Figure 2 suggests that all three designs can detect moderately shielded plutonium but cannot detect highly enriched uranium, which is in agreement with current

thinking.[4] In other words, using Design 3 rather than Design 1 and/or increasing the annual budget from \$5.5 M to \$10 M would offset a modest amount of shielding for plutonium weapons.

For both Designs 2 and 3, except for very small budgets (e.g., < 5.9 M/yr), it is optimal to use the maximal set size of six RPMs (i.e., $n = 6$). That is, it is more important to have a RPM to be in the proximity of the weapon than it is to increase the testing time. The use of six RPMs in a set creates a detection system that is robust with respect to weapon location, as indicated by the tight clustering of the worst-case and best-case detection limits for each of Designs 2 and 3 when the budget is greater than \$5.6 M (and the optimal set size is at least five); in contrast, for budgets of \$5.2 M, \$5.3 M, \$5.4 M, and \$5.5 M, Design 3 can only afford sets of size $n = 1, 2, 3$, and 4, respectively, and the worst-case (and hence mean) detection limit degrades dramatically.

Moreover, for Design 3, it is optimal to deploy exactly four sets of RPMs (i.e., $\frac{N}{n} = 4$), one in each of the four lanes and none in the aggregated queue. The intuition behind this nonobvious result (note that $n = 6$ and $\frac{N}{n} = 4$ holds only for the particular set of parameter values considered here) is that trucks are tested for 60 s by the RPMs in the lanes with probability one, but are tested for 15 s by the RPMs in the aggregated queue only in the case when the queue length is large (at least eight), and hence it is optimal to spend additional money by lowering the neutron threshold limit than by adding more RPMs. In Design 2, the threshold levels are chosen so that the false positive probability when the queue is large is nearly the same as when the queue is small. For the optimal version of Design 3 (i.e., $\frac{N}{n} = 4$), there is only a single threshold level and the false positive probability is independent of queue size. When the inequality $\frac{N}{n} > 4$ was imposed in the model, the false positive probabilities varied by less than 0.02 with queue size.

DISCUSSION

The existing design has two shortcomings: the container testing time is only 3 s and the RPM is idle most of the time. Offsite testing is the key cost-driver and the signal-to-noise ratio increases with the testing time, suggesting that longer RPM testing times that do not impose additional congestion should improve performance. The signal is also proportional to the inverse square distance between the nuclear weapon and the sensor, suggesting that using multiple sensors for each truck might improve the robustness of the detection system with respect to the location (within the container) of the weapon. The results confirm these hypotheses and identify (via Design 3) the optimal mix of spatial and temporal exploitation for neutron testing by RPMs. The results only pertain to the detection of neutron emissions; spectroscopic gamma portal monitors, which are able to distinguish fissile material from other legal shipments that emit radioactivity, are currently in development,[4] and the methods used

here can be applied to improve the deployment of these portal monitors in the future.

However, there are several caveats. First, the sensitivity and specificity data that are used are from the late 1990s[7] and may not be representative of the equipment that is currently in use, whose detailed performance metrics are proprietary; moreover, the area A taken from Ref. [10] is probably smaller than that used by a RPM that has sensors on both sides. This precludes the direct use of Figure 2 by the U.S. government or a nuclear smuggler, and suggests that readers should focus on the relative performance—rather than the absolute performance—of the designs. The experiments in Ref. [7] were performed over a period of months (May-September 1998 and again in the spring of 1999), which should be enough time to incorporate the range of variation in the mean background rate. Although these data are sufficient to calibrate (e.g., estimate the parameters for) the model, they are insufficient to actually validate the form of the radiation sensor model described in section B of the Appendix. Indeed, some of these designs were also assessed using the radiation sensor model in section A of the Appendix to Ref. [5], which led to significantly larger improvements in performance (i.e., larger reductions in the detection limit relative to Design 1) than seen in Figure 2 (data not shown). Although the detection model used in this article is viewed as being superior to the one used in Ref. [5] (the model in Ref. [5] used a mixture of normals, and so the number of neutrons detected was a continuous random variable that could occasionally take on negative values), it is difficult to place much confidence in the quantitative results reported here until the form of the model is validated. However, due to the proprietary nature of the data in this research area, sufficient data to calibrate the model may never become publicly available.

Nonetheless, the qualitative nature of the results should still hold, and Design 3 should offer at least a modest improvement over Design 1. The relative performance of Design 1 versus Design 3 is very insensitive to the mean service time at the front gate (Figure 3a). Moreover, Design 1 and the optimal (i.e., $\frac{N}{n} = 4$) version of Design 3 are independent of the queue length, and hence of the amount of congestion. Another key factor that dictates the relative amount of improvement possible is the mean offsite testing cost per container that undergoes a scan analysis (c_o) divided by the RPM cost (c_p). There is considerable uncertainty in c_o because there is no documented data on the fraction of containers that are diagnosed by VACIS or by high-energy x-ray, and the x-ray and manual testing costs may vary by terminal and may not equal the prices charged to importers. Nonetheless, Figure 3b shows that the difference in performance between Designs 1 and 3 is nearly independent of c_o over practical ranges (although, as expected, as c_o increases with a fixed budget, the detection limit also increases for both designs), and hence is also insensitive to the RPM cost c_p. Finally, although the focus is on the congestion at the front gate, the more false positives there are, the more disruption to the ship loading plans.

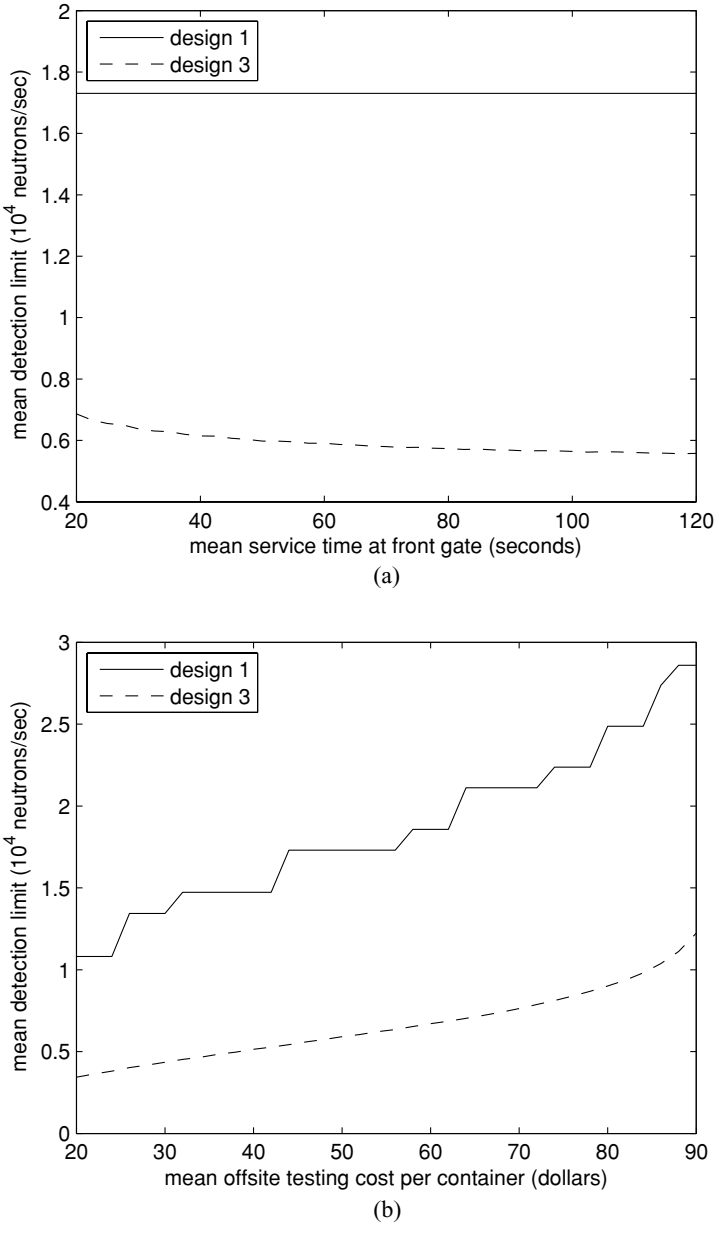

Figure 3: Sensitivity analysis. **(a)** The mean detection limit for Designs 1 and 3 at a fixed cost of \$10 M/yr, as a function of **(a)** the mean processing time at the front gate (μ^{-1}), which has a base-case value of 60 s; and **(b)** the mean offsite testing cost per container undergoing scan analysis (c_O), which has a base-case value of \$50.

Designs 2 and 3 have fewer false positives for a given cost because they use more RPMs, and hence for a given cost Designs 2 and 3 are actually less disruptive, that is, fewer containers miss their scheduled voyage, than Design 1.

CONCLUSION

Detecting highly enriched uranium and plutonium at ports is one of the most important and challenging problems in homeland security. In the model, an aggressive spatiotemporal deployment of RPMs, which uses a set of six RPMs in each truck lane to simultaneously monitor a truck while it is waiting to be processed at the front gate, offers a reduction (by a factor of three, although the accuracy of this value is difficult to assess) in the neutron detection limit. Such a reduction is not sufficient to detect highly enriched uranium, but should offset some shielding around a plutonium weapon. If such an improvement led smugglers to use more shielding, then detection via radiographic scan analysis would increase if the scans were analyzed. As the deployment of RPMs is rolled out, this design should be given serious consideration. The geometry of the problem (i.e., $40 \times 8 \times 8$ ft container, RPMs that are at least 2 m away from the center of the container widthwise, and a maximum of six RPMs that fit lengthwise along container, see Figure 1b) suggests that the best opportunity for further improvements (beyond the ones explored here) in exploiting the spatial aspects of the problem is to design RPMs so that the sensors on both sides can be placed up against the sides of the container during in-place testing. This enhancement would approximately double the detector counts because the inverse square distance would be reduced from $(\frac{1}{2\text{ m}})^2$ to $(\frac{1}{4\text{ ft}})^2$.

NOTES AND REFERENCES

1. SAIC. Integrated Container Inspection System brochure. Scientific Applications International Corporation, San Diego, CA (2004).

2. B. Rooney. Detecting nuclear weapons and radiological materials: How effective is available technology? Testimony before the Subcommittee on Prevention of Nuclear and Biological Attacks and the Subcommittee on Emergency Prepardeness, Science and Technology, The House Commmittee on Homeland Security, June 21, 2005.

3. E. Lipton. "U.S. to Expand Cargo Scans to Detect Nuclear Material." *New York Times*, December 8, 2006, A20.

4. D. Huizenga. Detecting nuclear weapons and radiological materials: How effective is available technology? Testimony before the Subcommittee on Prevention of Nuclear and Biological Attacks and the Subcommittee on Emergency Prepardeness, Science and Technology, The House Commmittee on Homeland Security, June 21, 2005.

5. L. M. Wein, A. H. Wilkins, M. Baveja, and S. E. Flynn. Preventing the importation of illicit nuclear materials in shipping containers. *Risk Analysis* 26 (2006): 1377–1393.

6. D. Gross and C. M. Harris. Fundamentals of Queueing Theory, 2nd Edition. (New York: John Wiley and Sons, 1985).

7. P. Beck. Illicit Trafficking Radiation Detection Assessment Program. Paper OEFZS-G-0005, Austrian Research Centers, Seibersdorf (2000). Accessed at http://www.tsasystems.com/library/reports/itrapfinreport.pdf, October 29, 2007.

8. E. Cinlar. "Superposition of Point Processes." In *Stochastic Point Processes: Statistical Analysis, Theory and Applications*, ed. P. A. W. Lewis (New York: Wiley, 1972), 549–606.

9. L. C. Edie. "Traffic Delays at Toll Booths." *J. Operations Research Soc. America* 2, (1954): 107–138.

10. S. Fetter, V. A. Frolov, M. Miller, R. Mozley, O. F. Prilutsky, S. N. Rodionov, and R. Z. Sagdeev. "Detecting Nuclear Warheads." *Science & Global Security* 1 (1990): 225–302.

11. D. M. Titterington, A. F. M. Smith, and U. E. Makov. *Statistical Analysis of Finite Mixture Distributions* (New York: John Wiley & Sons, 1985).

12. Press Office, U.S. Customs and Border Protection. Fact Sheet: Cargo Container security—U.S. Customs and Border Protection Reality. October 2004. Accessed at http://www.cbp.gov/linkhandler/cgov/newsroom/fact_sheets/port_security/5percent_myth.ctt/5percent_myth.doc, August 16, 2005.

APPENDIX

This appendix describes the mathematical model that generated the results reported in the main text. The queueing model is in section A, the RPM detection model is in section B, the total cost is derived in section C, and the detection limit vs. cost curves are constructed in section D. The model's parameters and their values are given in Table A1.

A Queueing

In the queueing model, trucks arrive at the terminal according to a Poisson process with rate λ, and enter a single line that feeds $m = 4$ lanes at the front gate of the terminal. To avoid server idling in these lanes, up to $l = 2$ container trucks are allowed to queue in each lane, and if there are more than $lm = 8$ trucks in total, then the additional trucks wait in a single aggregated queue. Each lane has a human server with a service time that is random with mean $\mu^{-1} = 60$ s and coefficient of variation (i.e., standard deviation divided by the mean) $c_s = 0.5$. Although there are no available data on the service time variability, the service times are likely to be less variable than an exponential random variable, which has a standard deviation equal to the mean, because of the uniformity and simplicity of the tasks involved. The traffic intensity is given by $\rho = \frac{\lambda}{m\mu}$, which is the fraction of time over the long run that each server is busy. Let Q represent the steady-state number of trucks in the system (in queue plus in service). The heavy-traffic approximation (e.g., Equation 2.13 in Ref. [1]), which is reasonably accurate for moderate m and ρ close to, but smaller than, one, states that Q is exponentially distributed with mean

$$E[Q] = \frac{\rho^2}{1-\rho}\frac{1+c_s^2}{2} + m\rho. \tag{1}$$

Letting $E[Q] = 24$ (there are approximately six trucks per lane in Ref. [2]) and solving Equation 1 for λ, gives $\lambda = 3.883$/min and $\rho = 0.971$, which indeed is consistent with the heavy-traffic regime. It is assumed that Q is geometrically

Table A1: Descriptions and values of model parameters.

Notation	Description	Value
λ	Truck arrival rate	3.883/min
m	Number of lanes	4
l	Trucks/lane after disaggregation	2
μ	Service rate	1/min
c_s	Coefficient of variation of service times	0.5
v_1	Drive-through velocity if not stopping	10 mph
	Drive-through velocity if stopping	5 mph
L	Length of container	40 ft
A	Area of neutron detector	0.3 m^2
ϵ	Efficiency of neutron detector	0.14
e^{θ}	Median neutron background rate	43.1 neutrons/m^2·sec
e^{σ}	Dispersal factor of neutron background rate	1.73
r	Distance of neutron detector	2 m
S	Neutron source	400,000 neutrons/sec (plutonium) 30 neutrons/sec (uranium)
u	Fraction of untrusted containers	0.05
p_v	Probability of diagnosis by VACIS	0.95
c_p	Annual cost per RPM	$20,000
p_x	Probability of diagnosis by high-energy x-ray	0.5
c_x	Cost of offsite high-energy x-ray test	$250
c_m	Cost of offsite manual test	$1500

distributed (with a mean of 24) rather than exponentially distributed because Q is integer-valued. More specifically, it is assumed that

$$Pr(Q = k) = p(1 - p)^k \quad \text{for} \quad k = 0, 1, 2, \ldots, \tag{2}$$

where $p = 0.04$.

The more complex approximation for the probability distribution of Q given in section 5.4 of Ref. [1] was also considered but found to be no more accurate (by comparing it to the distribution computed via Monte Carlo simulation) than the approximation in Equation 1.

B RPM Performance

The performance of a RPM is characterized by its detection probability d and its false positive probability f. Emissions can emanate from the weapon and the background level. The performance for Design 1 is derived in section B.1 and the performance for Designs 2 and 3 is derived in section B.2.

B.1 Performance of Design 1

In Design 1, each container travels through the RPM at $v_1 = 10$ mph. If the authors denote $L = 40$ ft to be the length of a container, then the testing time

is $t^d = \frac{L}{v_1} = 2.73$ s, where the superscript is mnemonic for drive-through. The variable X denotes the number of neutrons detected from testing a weapon-free container. It is assumed that X is a Poisson random variable with parameter (i.e., mean) $A\epsilon bt^d$, where $A = 0.3$ m^2 is the area of a neutron detector,[3] $\epsilon = 0.14$ is the efficiency of a neutron detector,[3] and b is the background rate, which itself is a log-normal random variable with median e^θ and dispersal factor e^σ (i.e., $\ln b$ is a normal random variable with mean θ and standard deviation σ). Hence, the probability mass function of X is

$$P(X = x) = \int_0^\infty \frac{e^{-\frac{(\ln y-\theta)^2}{2\sigma^2}}(A\epsilon y t^d)^x e^{-A\epsilon y t^d}}{\sqrt{2\pi}\sigma y x!}\, dy \quad \text{for } x = 0, 1, \ldots. \tag{3}$$

The mean neutron background rate is assumed to be 50 neutrons/m^2·sec,[3] which implies that $\theta = \ln 50 - \frac{\sigma^2}{2}$. To estimate σ, it is noted that when the detection time is $t = 10$ s, the distance between the sensor and the source is $r = 2$ meters and the source is $S = 20$ k neutrons/sec, the false positive probability and the false negative probability from extensive controlled experiments (with a stationary source) were 10^{-4} and 10^{-3}, respectively.[4] The emissions detected from the stationary source in these experiments is modeled as $\frac{A\epsilon St}{4\pi r^2}$.[3] Solving

$$P(X > \bar{n}) = 10^{-4}, \tag{4}$$

$$P\left(X + \frac{A\epsilon St}{4\pi r^2} < \bar{n}\right) = 10^{-3}, \tag{5}$$

for the threshold level $\bar{n}$ and σ yields $\sigma = 0.546$, and hence the median background rate is $e^\theta = 43.1$ neutrons/m^2·sec and the dispersal factor is $e^\sigma = 1.73$.

The variable Y is defined to be the number of neutrons detected from testing a weaponized container. So that only two spatial dimensions need to be considered, it is assumed that the RPM and the weapon are at the same height, and that $r = 2$ m (Ref. [4]) is the minimum distance between the weapon (which is located in the middle of the vehicle widthwise) and the RPM as the container passes through the RPM (see Figure 1b in the main text). The cumulative emissions W at the RPM are calculated by integrating over time, from time 0, which corresponds to when the front of the 40-ft container passes by the RPM, to time $t^d = \frac{L}{v_1}$, which is when the back of the container passes by the RPM. Without loss of generality for the drive-through calculation, the weapon is assumed to be in the middle of the vehicle lengthwise (this is not assumed for the in-place calculations in Designs 2 and 3), so that the weapon is at a distance $r = 2$ m from the RPM at time $t = \frac{L}{2v_1}$, and more generally the distance between the weapon and the RPM at time t is $\sqrt{r^2 + (v_1 t - \frac{L}{2})^2}$. Hence, if the neutron source of the weapon is S, which is allowed to vary in the computational study, then the cumulative emissions W at the detector due to the weapon passing through

the RPM is

$$W = \frac{A\epsilon S}{4\pi} \int_0^{t^d} \frac{dt}{r^2 + \left(v_1 t - \frac{L}{2}\right)^2} = \frac{\tan^{-1}\left(\frac{L}{2r}\right) A\epsilon S}{2\pi v_1 r}. \tag{6}$$

The authors assume Y is a Poisson random variable with parameter $A\epsilon b t^d + W$, where again b is a log-normal random variable.

If $\bar{n}$ denotes the neutron threshold level above which an alarm is generated, then the detection probability and the false positive probability are given by, respectively,

$$d(S) = P(Y > \bar{n}), \tag{7}$$

$$f = P(X > \bar{n}), \tag{8}$$

where the dependence of the detection probability on the source term is made explicit. The detection limit is defined by

$$D(\bar{n}) = \{S | d(S) = 0.95\}. \tag{9}$$

B.2 Designs 2 and 3

RPM placement. Designs 2 and 3 are described using the same framework. More notation is needed than in describing Design 1 because now there is a set of RPMs that simultaneously monitor a container, and containers may stop at a set of RPMs. There are N RPMs in total, which are collected into sets of n RPMs. The n RPMs within a set simultaneously monitor a container, the RPMs within a set are positioned so that they each monitor an interval (measured along the length of the container) of length $\frac{L}{n}$, and each is located at the center of its interval. For example, RPMs are placed at 10 and 30 ft if $n = 2$, and at 5, 15, 25, and 35 ft if $n = 4$ (see Figure 1b in the main text). More generally, if $x = 0$ correspond to the front of the container and $x = L$ corresponds to the back of the container, then RPM j in a set is located at $\frac{(j-\frac{1}{2})L}{n}$ for $j = 1, \ldots, n$. Because each RPM takes up approximately 6 ft along the length of the container, the value of n is assumed to be no more than 6. Each set of n RPMs is located at a particular position in queue, where $i = 1$ denotes the set of RPMs closest to the front gate (there are $m = 4$ sets for $i = 1$, one for each of the 4 lanes), and $i = 2, \ldots, \frac{N}{n} - m + 1$ are the sets in the aggregated queue; note that there is a position in queue behind $i = 1$ that is not indexed because there are no RPMs there. The two decision variables, N and n, are chosen so that the number of sets, $\frac{N}{n} - m$, is a nonnegative integer.

RPM testing times. Let t_{ij}^d and t_{ij}^p be the drive-through time and the in-place time for RPM j in set i, which are random because they depend on the realization of the steady-state queue length Q, which represents the number

of trucks in the system when a new truck arrives, not including the new one. These testing times are now specified for Designs 2 and 3. Design 2 has only a single set of RPMs (i.e., $N = n$) that is located at position $i = 2$ at the head of the aggregated queue. Design 2 is fit into this notational framework by setting $t_{ij}^d = t_{ij}^p = 0$ for $i \neq 2$. If $Q \leq lm - 1 = 7$, which occurs with probability $\sum_{i=0}^{7} p(1-p)^i$, then an arriving truck passes through the RPM at 10 mph, as in Design 1, in which case $t_{2j}^d = 2.73$ s and $t_{2j}^p = 0$. In the present model, if a truck stays in place at a set of RPMs for any positive amount of time, then it is assumed that the drive-through velocity (i.e., the velocity before and after he stops) is only $v_2 = 5$ mph, which leads to a drive-through time of 5.46 s. Hence, if $Q \geq lm = 8$ in Design 2, then an arriving truck has drive-through time $t_{2j}^d = 5.46$ s and in-place time $t_{2j}^d = \frac{\mu^{-1}}{m} = 15$ because a truck at the head of the aggregated queue leaves this position after 1 of the 4 trucks completes its service at the front gate (while this amount of time is actually random, t_{2j} is assumed to be a constant that equals the mean of this random variable).

Turning to Design 3, there are three cases to consider. If $Q \leq lm - 1$ then the arriving truck does not stop in the aggregated queue, and for $j = 1, \ldots, n$, the testing times are $t_{1j}^d = 5.46$ s, $t_{1j}^p = \mu^{-1} = 60$ s, and $t_{ij}^d = 2.73$ s and $t_{ij}^p = 0$ for $i = 2, \ldots, \frac{N}{n} - m + 1$. If $Q \in [lm, (l-1)m + \frac{N}{n})$, then the arriving truck stops at sets $i = 2, \ldots, Q - lm + 2$ but does not stop at sets $i = Q - lm + 3, \ldots, \frac{N}{n} - m + 1$. Hence, if $Q \in [lm, (l-1)m + \frac{N}{n})$, then for $j = 1, \ldots, n$, the testing times are $t_{1j}^d = 5.46$ s and $t_{1j}^p = \mu^{-1} = 60$ s, $t_{ij}^d = 5.46$ s and $t_{ij}^p = \frac{\mu^{-1}}{m} = 15$ s for $i = 2, \ldots, Q - lm + 2$, and $t_{ij}^d = 2.73$ s and $t_{ij}^p = 0$ for $i = Q - lm + 3, \ldots, \frac{N}{n} - m + 1$. Finally, if $Q \geq (l-1)m + \frac{N}{n}$, then the arriving truck will initially stop in the part of the aggregated queue where there are no RPMs, and will subsequently slowly move through each set of RPMs, so that for $j = 1, \ldots, n$, the testing times are $t_{1j}^d = 5.46$ s and $t_{1j}^p = \mu^{-1} = 60$ s, and $t_{ij}^d = 5.46$ sec and $t_{ij}^p = \frac{\mu^{-1}}{m} = 15$ s for $i = 2, \ldots, \frac{N}{n} - m + 1$.

Neutrons detected by a RPM. In Designs 2 and 3, RPMs now detect neutrons not only while a truck drives through a RPM, but also while a truck stops at a RPM. Let X_{ij} denote the number of neutrons detected by RPM j in set i from testing a weapon-free container; note that although there are 4 sets of RPMs at position 1, any given container only passes through one of these sets. It is assumed that X_{ij} is a Poisson random variable with parameter $A\epsilon b(t_{ij}^d + t_{ij}^p)$. While b is still a log-normal random variable, it is assumed that the same realization of b is used for any given container, because the background rate varies on a longer time scale than the amount of time that a container spends in the queueing system (i.e., the background rate is random but fixed during a container's residence in the system).

The amount of emissions detected from the weapon by RPM j in set i includes not only the drive-through term, which is now $\frac{\tan^{-1}(\frac{L}{2r})A\epsilon S}{2\pi v_{ij} r}$, where v_{ij} is

either 5 mph or 10 mph, according to all the cases above (i.e., $v_{ij} = 10$ mph when $t_{ij}^d = 2.73$ s and $v_{ij} = 5$ mph when $t_{ij}^d = 5.46$ s), but also an in-place term. The in-place term is $\frac{A\epsilon St_{ij}^p}{4\pi d_{ij}^2}$,[3] where d_{ij}, which is the distance between the weapon and RPM j in set i, is specified in the next paragraph. Hence, the cumulative neutron emissions detected by RPM j in set i due to the weapon is $W_{ij} = \frac{\tan^{-1} l(\frac{L}{2r})A\epsilon S}{2\pi v_{ij} r} + \frac{A\epsilon St_{ij}^p}{4\pi d_{ij}^2}$, and the number of neutrons detected by RPM j in set i from testing a weaponized container is a Poisson random variable Y_{ij} with parameter $A\epsilon b(t_{ij}^d + t_{ij}^p) + W_{ij}$.

Weapon location. If the weapon is placed at position x along the length of the container (where $x = 0$ and $x = L$ represent the front and back of the container), then recalling the placement of RPMs with a set, the authors have that the distance between the weapon and RPM j in set i is $d_{ij} = \sqrt{r^2 + (x - \frac{(j-\frac{1}{2})L}{n})^2}$. Designs 2 and 3 are optimized under the assumption that the weapon location x is uniformly distributed over $[0, L]$. To assess robustness, the detection limits for the optimal versions of these designs are also computed for the base-case and worst-case weapon locations. In the best case, the weapon is located at $x = \frac{L}{2n}$, which minimizes the distance d_{i1} between the weapon and the first RPM in the set (by symmetry, any RPM in the set could be chosen), with this distance being r. Similarly, in the worst case, the weapon is located at $x = 0$, which maximizes the distance between the weapon and the first RPM in the set, leading to $d_{i1} = \sqrt{r^2 + (\frac{L}{2n})^2}$.

RPM threshold levels. The queue length upon arrival, and hence the testing time, is random but observable during the testing process in Designs 2 and 3, and so a different threshold level $\bar{n}$ can be chosen for each possible value of the queue length Q. Let $\bar{n}_k$ be the threshold value when $Q = k$.

Although there are a variety of ways to decide whether the realizations of X_{ij} and Y_{ij} set off an alarm, because the weapon is stationary throughout the testing process, for each position $j = 1, \ldots, n$ the emissions are summed over all sets i, and then the maximum emissions from any position is considered. That is, when $Q = k$, the detection probability and false positive probability are equal to

$$d_k(S) = P\left(\max_{j=1,\ldots,n} \sum_{i=1}^{\frac{N}{n}-m+1} Y_{ij} > \bar{n}_k \right), \tag{10}$$

$$f_k = P\left(\max_{j=1,\ldots,n} \sum_{i=1}^{\frac{N}{n}-m+1} X_{ij} > \bar{n}_k \right). \tag{11}$$

The detection limit when $Q = k$ is

$$D_k = \{S | d_k(S) = 0.95\}. \tag{12}$$

The mean detection limit and the overall false positive probability are computed by taking expectations with respect to the queue length Q defined in Equation 2. That is, the mean detection limit is

$$D = \sum_{k=0}^{\infty} p(1-p)^k D_k, \tag{13}$$

and the overall false positive probability is

$$f = \sum_{k=0}^{\infty} p(1-p)^k f_k. \tag{14}$$

C Costs

To calculate the testing costs, it is assumed that the only VACIS scans that are analyzed are those that fail passive testing, plus an additional fraction $u = 0.05$ of untrusted containers that are flagged by the ATS,[5] which more accurately captures how this testing system might be used by U.S. Customs and Border Protection. By Equation 33 in the Appendix of Ref. [6], the arrival rate to the VACIS scan analysts is

$$\lambda_o = \lambda[1 - (1-u)(1-f)], \tag{15}$$

where the false positive probability f is given by Equation 8 for Design 1 and Equation 14 for Designs 2 and 3. It is assumed that the VACIS test correctly diagnoses the false positives or the untrusted containers with probability $p_v = 0.95$, and that a fraction $1 - p_v$ of the containers undergoing scan analysis need to undergo further testing offsite.

The expected total annual cost has two components, the annualized cost of each portal monitor, and the cost of offsite testing. As described in the main text, the annual cost per RPM is $c_p = \$20,000$[2,6] and the mean offsite testing cost per container that undergoes scan analysis is

$$c_o = (1 - p_v)[c_x + (1 - p_x)c_m], \tag{16}$$

which is \$50 in the base case, where $p_x = 0.5$ is the probability of diagnosis by high-energy x-ray, $c_x = \$250$ is the cost of an offsite high-energy x-ray test, and $c_m = \$1,500$ is the cost of an offsite manual test. Let n_k be the number of RPMs used in Design k = 1,...,3, so that $n_1 = 1$, $n_2 = n$ and $n_3 = N$. Because λ_o in Equation 15 is in terms of tests per minute, the expected total annual cost is

$$C = c_p n_k + 525,600 c_o \lambda_o. \tag{17}$$

D Detection Limit vs. Cost Curves

Each design is characterized by its mean detection limit versus cost curve. For Design 1, the testing time is a constant and a detection limit versus cost curve can be generated by varying $\bar{n}$. The curves for Designs 2 and 3 require the solution of some optimization problems. For each value of the cost, the values of n and $\bar{n}_k$ for Design 2, or the values of N, n, and $\bar{n}_k$ for Design 3 are chosen, to minimize the mean detection limit D in Equation 13 subject to a cost constraint using Equation 17.

However, additional restrictions on $\bar{n}_k$ in Design 3 are imposed to ease the computational burden. Note that for given values of N and n, there is a one-to-one correspondence between f_k and $\bar{n}_k$ via Equation 11 (i.e., each value of $\bar{n}_k$ generates a corresponding unique value for f_k) for each value of k. Hence, constraints on f_k are imposed that lead to corresponding constraints on $\bar{n}_k$. In Design 3, a truck is in-place at one set of RPMs if $Q \leq lm - 1$ and is in-place at all sets of RPMs if $Q \geq (l-1)m + \frac{N}{n} - 1$. Hence, it is assumed without loss of optimality that $f_k = \underline{f}$ if $k \leq lm - 1$ and $f_k = \bar{f}$ if $k \geq (l-1)m + \frac{N}{n} - 1$, where $\underline{f}$ and $\bar{f}$ are parameters to be optimized. The additional constraints require that f_k be a linear interpolation of $\underline{f}$ and $\bar{f}$ for intermediate values of k:

$$f_k = \underline{f} + \frac{k - lm + 1}{(l-1)m + \frac{N}{n} - lm}(\bar{f} - \underline{f}) \quad if \quad k \in \left[lm, (l-1)m + \frac{N}{n} - 2 \right]. \tag{18}$$

For each value of N and n in Design 3, the constraints in Equation 18 allow for optimization over two parameters, $\underline{f}$ and $\bar{f}$, rather than over the thresholds $\bar{n}_k, k = 0, 1, \ldots$.

Solving these optimization problems for various cost values gives the detection limit versus cost curves in Figure 2 of the main text.

APPENDIX NOTES AND REFERENCES

1. W. Whitt. "Approximations for the $GI/G/m$ Queue," *Production Operations Managment* 2 (1993): 114–161.

2. SAIC. *Integrated Container Inspection System brochure*. Scientific Applications International Corporation, San Diego, CA (2004).

3. S. Fetter, V. A. Frolov, M. Miller, R. Mozley, O. F. Prilutsky, S. N. Rodionov, R. Z. Sagdeev. "Detecting nuclear warheads." *Science & Global Security*, 1 (1990): 225–302.

4. P. Beck. Illicit Trafficking Radiation Detection Assessment Program. Paper OEFZS-G-0005, Austrian Research Centers, Seibersdorf (2000).

5. Press Office, U.S. Customs and Border Protection. Fact Sheet: Cargo Container Security—U.S. Customs and Border Protection Reality. October 2004. Accessed at http://www.cbp.gov/linkhandler/cgov/newsroom/fact_sheets/port_security/5percent_myth.ctt/ 5percent_myth.doc, August 16, 2005.

6. L. M. Wein, A. H. Wilkins, M. Baveja, and S. E. Flynn. "Preventing the Importation of Illicit Nuclear Materials in Shipping Containers." *Risk Analysis* 26 (2006): 1377–1393.