A Decision Theoretic Model for Determining Verification Requirements[1]

RICHARD L. PATTERSON
Bendix Systems Division

WYMAN RICHARDSON
Institute of Science and Technology, University of Michigan

As national security, even national existence, has come to depend on rational self-restraint on the part of an ever increasing number of sovereign governments, a willingness to consider accepting the risks of an arms control or disarmament treaty becomes manifest. Careful thought is necessary to make the risk as small as possible. If the system for verifying compliance with the treaty is too lax, national security is at the mercy of a potential violator. If too strict a system is insisted upon, the treaty may never be negotiated and the danger of the present arms race will continue and increase.

United States diplomats can negotiate a disarmament treaty most effectively if they keep in mind a final offer, a requirement for strictness of verification below which they will not go. The final offer must be one that strikes the best balance between the two risks R_T, the risk under a treaty, and R_N, the risk under no treaty. It appears that α, the United States negotiators' perception of the probability that a given final offer will be accepted by the Soviet Union, increases as strictness decreases. Hence R_T, considered as a function of α, increases with α. The total risk R is

$$R(\alpha) = (1-\alpha)R_N + \alpha R_T(\alpha).$$

The best final offer is a requirement for strictness of verification minimizing $R(\alpha)$.

The evaluation of the best final offer is necessarily subjective since α, R_N, and R_T depend on many factors which are hard to measure. It will be assumed in this note that the tolerable risk R_T, corresponding to the best final offer, has been decided upon. A model will be described specifying what quantity and quality of inspection information is consistent with this bound.[2]

As an initial approach to a complicated problem, a very simple inspection system is assumed. It is a system that from time to time generates an opinion (called a "detection") that a certain type of event (e.g.,

[1] The authors, in part, developed the ideas contained herein for the Bendix Systems Division of the Bendix Corporation under a contract with the United States Arms Control and Disarmament Agency. The ideas and conclusions presented herein, however, are solely those of the authors. They do not necessarily coincide with those of the United States Arms Control and Disarmament Agency or with those of any other representative of the United States.

[2] A detailed description of the model is given in Bendix Systems Division (1963).

nuclear testing) has occurred. It is assumed that each event is detected with constant probability p, that some opinions may be false alarms according to a Poisson distribution with mean β, and that opinions are arrived at independently. The detection probability p is considered a measure of the quantity of information and the false alarm rate β, a measure of the quality of information.

Values of p and β consistent with the bound on R_T can be considered as verification requirements. To evaluate R_T, the risk under the treaty, it is assumed that, after a time interval has elapsed, the inspecting nation chooses among alternative responses and that the appropriateness of each response is measured by a loss function depending on the response and the number of events that have occurred. For any response rule, the expected loss conditional upon a given number of events having occurred is obtained by a straightforward probability calculation. These conditional expected losses are combined into a total expected loss (risk) by assigning weights equal to the responding country's prior opinion of the number of events likely to occur in the time interval. It is assumed that the response rule minimizing the risk is in force. R_T is considered to be this minimal risk. The same response rule is obtained by choosing the response minimizing the expected loss with respect to the Bayesian posterior distribution of the number of events.[3]

A simple case of the decision procedure occurs when only two responses are assumed possible, one appropriate when the number of events exceeds the legal limit (a violation) and the other when the legal limit is respected. Four possible response losses are assumed: those of responding correctly and incorrectly in each case. The best decision rule can be shown to be to respond "violation" if and only if the number of detections exceeds a critical level.

Since the formulas for the response rule depend on the prior distribution of the number of events, a foolishly conceived distribution is likely to result in foolish decisions. Guidelines are needed for the wise specification of a prior distribution. The following is one possible construction.

Each event that a nation perpetrates in violation of the treaty is assumed to yield a military gain that is greatest for the first violation and decreases for succeeding violations. A loss attendant to discovery is assumed to be taken following a reported detection.

If the military gain and the discovery loss were known exactly by the responding nation, then that nation would know exactly how many violations were planned, namely, the number optimizing the expected net gain. Since the gains and losses are not known exactly, their values may be represented by a probability distribution. This distribution generates a probability distribution of the optimal number of violations, which serves as the prior distribution of the response model. The advantage of this procedure is that one single prior distribution of losses and gains generates a family of prior distributions varying with the detection probability and hence with the incentive to violate.

A computer program was written to carry out the otherwise tedious calculations required to implement the model. When the program was run using the prior distribution just described, a discrepancy came to light. Certain low values of p were sufficient to supposedly deter a potential violator but useless for recognizing violations. The fault lay with the supposition that a violator is deterred by the detection report of a violation

[3] An explanation of the logical principles underlying Bayesian decision theory is given in Savage, 1954.

rather than by the recognition of, and response to, a violation. A reported violation masked by numerous false alarms would be very likely to go unpunished.

Continued research on the model in a number of directions is indicated. A prior distribution depending on fear of recognition of, and response to, a violation might be devised, possibly by an iterative procedure starting with a guessed distribution, running the two-response model to calculate the probability of response to a violation, obtaining a better prior distribution from these probabilities and a distribution of gains and losses, and so on until the procedure hopefully converges. Modifications of the prior distribution to allow for irrational decision-making; compromise between pressure groups; or other factors, e.g., good faith independent of the probability of discovery, might be considered.

The distribution of the number of events required to produce d detections when $\beta = 0$ is the well-known negative binomial distribution. This fact suggests the tempting possibility of obtaining a posterior distribution of the number of events without having to postulate a prior distribution. The authors are examining a generalization of the negative binomial distribution that is appropriate when $\beta > 0$.

The model assumes inspection inputs in the form of independent opinions that an event has occurred, some of which are false alarms and others refer to events detected with a constant probability p. It is most unlikely that any actual inspection system would obey this assumption. A system to detect underground nuclear tests, for example, reports several detection criteria depending on components of the seismic waves, each with its own false alarm rate and detection probability. Missile flight tests are detected by infrared, radar, and acoustic sensors. These examples suggest that to be realistic, the decision model should be modified to accept multichannel inspection inputs.

It is obviously unrealistic to assume that each nation waits until January first to respond to a violation. Rather, a nation would be expected to respond after a critical number of detections accumulates. Although the probability distribution of the number of events is useful information, a decision to respond must weigh the expected cost of a mistake against the expected cost of postponing a decision.

The decision model has operated from the standpoint of the unilateral decision-maker. If a decision model is to be used by an International Disarmament Organization to convict a violator, the requirements are somewhat different. It would not be feasible to convict, for example, if the number of detections were no greater than the legal limit of the number of events, even if the parameter values were small enough to render almost certain the posterior probability of violation.

If there are n possible types of violation and an inspection system to detect each type, what is the best way to decide whether to respond as to a violation? It is not clear that the best way is to replicate the single-type model n times. The chance of a false accusation would be greatly increased. A new balance between false accusation and failure to respond might have to be struck.

Bayesian optimality is not the only reasonable criterion for choosing a decision rule. Another criterion which is highly regarded by contemporary scholars is the minimax principle. Instead of averaging the expected losses, each conditional upon a given number of events having occurred, and minimizing the average, the minimax principle declares as best that response rule which minimizes the largest of the conditional expected losses. This criterion emphasizes safety; day-in, day-out profit may not be as great but protection is given against the worst possible eventuality. If one does not accept the minimax principle entirely, one might still keep in mind the possibility of finding a response rule that greatly increases safety with only a slight sacrifice of Bayesian optimality.[4]

Mathematical Description of the Model

Given that there were k events, the random variable d, the number of detections, is the sum of m and f, the number of detected events and false alarms respectively. It is assumed that m is binomial (k, p) and f is assumed Poisson (β). The distribution $P(d|k)$ of d given k is the convolution of the distributions of m and f:

$$\sum_{m=0}^{d}\left[\binom{k}{m}p^m(1-p)^{k-m}\right]\left[\frac{e^{-\beta}\beta^{(d-m)}}{(d-m)!}\right].$$

d has a mean of $pk+\beta$ and a variance of $p(1-p)k+\beta$.

The inspecting nation's prior opinion of the number of events likely to occur in a unit time interval is summarized by a probability distribution $S(k)$. After evaluating the inspection information d, the inspecting nation has a modified opinion represented by the probability distribution $Q(k|d)$ which is calculated by Bayes' formula:

$$Q(k|d) = \frac{S(k)P(d|k)}{T(d)},$$

where $T(d) = \sum_k S(k)P(d|k)$.

Let L_{rk} be the loss incurred by making the response r when there were k events. The loss to be expected from using the response rule $r(d)$ when there were k events is

$$\rho(k) = \sum_d L_{r(d),k}P(d|k).$$

Since the response rule is determined prior to the occurrence of the events, the best rule is the one minimizing ρ, the expected value of $\rho(k)$ with respect to the prior distribution of k.

$$\begin{aligned}\rho &= \sum_k\sum_d L_{r(d),k}P(d|k)S(k)\\ &= \sum_d T(d)\sum_k L_{r(d),k}\frac{P(d|k)S(k)}{T(d)}\\ &= \sum_d T(d)\sum_k L_{r(d),k}Q(k|d).\end{aligned}$$

Thus the response rule minimizing ρ is to make the response minimizing the expected loss with respect to the posterior distribution of k. The risk R_T under the treaty is considered to be this minimal value of ρ.

For a given inspection apparatus, it may be possible to increase p at the expense of increasing β by being less fussy about what is termed a detection. R_T may be used as a figure of merit for determining the best adjustment of p and β. R_T is also a useful criterion for comparing alternative treaty provisions. In each of these applications, changes in p and β result in a change in the incentive to violate, which should be reflected in changed specifications of the prior distribution of the number of events.

In the two-response model, the set of possible responses is restricted to two: response r_1 appropriate when k exceeds the legal limit N and r_0 appropriate when $k \leqq N$. Four possible losses are assumed: L_{11} and L_{01} incurred by responding correctly and incorrectly, respectively, to a violation, and L_{00} and L_{10} incurred by responding correctly and incorrectly, respectively, to no violation.

The best response rule can be shown to be to respond "violation" if and only if

$$\sum_{k>N} Q(k|d) > \frac{L_{10}-L_{00}}{L_{10}-L_{00}+L_{01}-L_{11}}.$$

[4] A comparison of the minimax criterion with other criteria is given in Milnor (1954).

The right side of the inequality is a constant determined by the response losses. A proof is given in Bendix Systems Division (1963) that the left side, which is the posterior probability of violation given d, is an increasing function of d.[5] The decision rule is therefore to respond "violation" when and only when the number of detections exceeds a certain critical number d_c.

The calculation of the posterior distribution $Q(k|d)$, the critical level d_c of detections, the probability of responding incorrectly, and the expected response loss R_T requires many sums of products and would be tedious on a hand machine. A program to carry out the calculations was therefore written in Fortran and compiled on the Bendix G-20 computer. The inputs required are the values of p, β, N, the response losses (for the two-response case), and the prior distribution of the number of events (which may be positive only for $k \leqq 20$). The execution time is less than a minute.

A prior distribution varying with the incentive to violate was constructed by assuming that a military gain $g(i)$ of each violation i is realized until a detection is reported, which occurs after the first violation with probability p, after the second with probability $(1-p)p$, after the third with probability $(1-p)^2p$, and so on.[6] Thereupon a loss L is taken. The expected net gain from a plan of v violations is

$$\sum_{i=1}^{v}[G(i)-L]p(1-p)^{i-1}+ G(v)\sum_{v+1}^{\infty}p(1-p)^{i-1},$$

where $G(i)=\sum_{i=1}^{i}g(j)$. The marginal expected net gain of the vth violation is, after algebra,

$$[g(v)-Lp](1-p)^{v-1}.$$

If $g(v)$ decreases monotonically, then the optimal number of planned violations is the largest integer v for which

$$g(v)-Lp>0.$$

Mindful of the pertinence of the model to the control of nuclear tests and missile flight tests, a maximum M was postulated for the cumulative military gain $G(i)$. Each $g(i)$ was assumed to advance $G(i)$ a fixed ratio c of the remaining distance to M. As a result,

$$g(i)=Mc(1-c)^{i-1}.$$

The optimal number of violations is the greatest integer preceding x, the solution of

$$Mc(1-c)^{x-1}-Lp=0,$$

which is

$$x=1+\frac{\log(Lp/Mc)}{\log(1-c)}.$$

If a prior distribution is specified for Lp/Mc, which is the ratio of the expected discovery loss on the first violation to the military gain of the first violation, then a distribution of x, a monotonic function of Lp/Mc, is determined. From the distribution of x is obtained the distribution of the greatest integer less than x, which serves as the prior distribution of the number of events.

[5] The authors wish to thank Robert M. Thrall for help in constructing this proof.

[6] This result assumes that every event is a violation. If N unannounced events are legal and events beyond the Nth are violations, the probability that detection will occur after the ith violation is

$$\binom{i+N-1}{N}p^{N+1}(1-p)^{i-1}.$$

The disadvantage of writing a treaty with $N>0$ is shown by the following table of the expected number of violations before detection as a function of N and p:

	$p=.2$	$p=.5$	$p=.8$
$N=0$	5	2	1.2
$N=5$	25	7	2.5
$N=25$	105	27	7.5