Player Aggregation in the Traveling Inspector Model

JERZY A. FILAR

Abstract—**We consider a model of dynamic inspection/surveillance of a number of facilities in different geographical locations. The inspector in this process travels from one facility to another and performs an inspection at each facility he visits. His aim is to devise an inspection/travel schedule which minimizes the losses to society (or to his employer) resulting both from undetected violations of the regulations and from the costs of the policing operation. This model is formulated as a noncooperative, single-controller, stochastic game. The existence of stationary Nash equilibria is established as a consequence of aggregating all the inspectees into a single "aggregated inspectee." It is shown that such player aggregation causes no loss of generality under very mild assumptions. A notion of an "optimal Nash equilibrium" for the inspector is introduced and proven to be well-defined in this context. The issue of the inspector's power to "enforce" such an equilibrium is also discussed.**

I. Introduction

The traveling inspector model (TIM) is a mathematical model of dynamic inspection/surveillance of a number of facilities (these will sometimes be called sites, plants, or inspectees) in different geographical locations. The present author originally proposed this model and a number of alternative methods of analysis in [14], and in [12] an implementation of one of these methods is discussed. Conceptually, TIM is an inspection process with the following structure.[1]

1) There are S *inspectees* (or facilities or sites) in different locations.

2) There is one *inspector* who can perform only one inspection during the current inspection period (e.g., day, week, etc.).

3) The inspector *travels* from site to site and performs an inspection at the new site at which he "just arrived."

4) The inspectees know *the last* inspection site but not the next.

5) The inspector wishes to minimize the overall *cost to society* (or to his employer): this may include costs due to violation of regulations/cheating, travel costs, and inspection costs.

6) The *duration* of the process (i.e., number of inspection periods, or stages) can be either finite and known, or infinite.

In this paper we formulate the above process as a noncooperative, single-controller, stochastic game with either infinite or finite horizon. We show that the game possesses stationary Nash equilibria which can be found by recently developed techniques. In stochastic games, stationary strategies are the easiest strategies to implement, and hence it is important that the existence of solutions in this class of strategies enables us to restrict consideration to this class only (there are examples: Blackwell and Ferguson's [4] "big match" is one, where this simplification is impossible). The above results are obtained by showing that under weak assumptions (in the inspector/inspectee context), there is no loss of generality in assuming that the inspectees have "amalgamated" (i.e., are jointly coordinating their behavior/actions), so that some known results for two-person, single controller stochastic games can be applied to our problem. The key part of this argument lies in the interesting general conditions for "player aggregation" developed recently by Goldman and Shier [16] and Goldman [17]. Further, we introduce the notion of the inspector's "optimal Nash equilibrium," show that it exists, and that it can, in principle at least, be computed by a finite algorithm (see Theorem 3.5). However, the critical issue of the inspector's power to "enforce" such an equilibrium point (or any other) is only partially resolved (see Lemmas 3.7 and 3.9).

The paper has two objectives. First, we believe that the class of motivating situations is important enough and the structure of the model rich enough to provide stimulation for other researchers. This view seems confirmed by the fact that the model easily lends itself to many modifications and generalizations (for instance, see Section V), some of which lead naturally to unanswered theoretical questions. Of course, at present, TIM is only a "prototype" model which would have to be specialized and adapted appropriately to yield an accurate representation of a specific actual inspector/inspectee conflict situation.

Second, we wish to draw attention to the fact that due to recent algorithmic developments in the theory of stochastic games, it has become feasible to attack a relatively complex problem such as the TIM by modeling and solving it as a stochastic game with an appropriate special structure.

In the past 15 years or so, there has been a marked increase in research activity aimed at identifying particular classes of stochastic games which admit "simple" solutions and at developing algorithms for constructing these solutions (see, for instance, [20], [37], [8]–[10], [11], [13], [38], [39], [21], [22], [34], [6], and [2], just to name a few). At the same time applications of stochastic games are also beginning to emerge (see, for instance, [5], [35], [1], and [40]).

To conclude this Introduction we mention that most of the earlier game-theoretic models of the inspector-inspectee conflict has been static (i.e., single-stage) games, and thus essentially different from the model proposed here. See for instance, [28] and [18]. However, one consequence of the existence of stationary equilibria established here, is that despite its underlying dynamic nature the traveling inspector model can also be solved as a static game. The latter observation may be particularly important in these potential applications in which the inspector is a mechanical/electronic device programmed to follow a particular strategy, in an environment that does not readily permit the use of an adaptive strategy.

II. The TIM as a Stochastic Game

We shall consider a game with $(S + 1)$ players. Players 1, 2, $\cdots$, S will be the inspectees and player $(S + 1) = II$ will be the inspector. The sth inspectee can be thought of as the manager of the sth plant/site. The symbol S will have dual meaning, denoting both the Sth inspectee and the set $\{1, 2, \cdots, S\}$ of sites to be inspected. Now, during a typical inspection period[2] $[t, t + 1)$ the

Manuscript received March 9, 1984; revised December 5, 1984. Paper recommended by Associate Editor, J. Walrand. This work was supported in part by the National Science Foundation under Grant ECS-8204677.

The author is with the Department of Mathematical Sciences, The Johns Hopkins University, Baltimore, MD 21218.

[1] Depending on the particular method of analysis chosen, some of the following assumptions can be easily relaxed or modified.

[2] The precise timing of events during an "inspection period" is left open as it would almost certainly depend on the context of the model. For instance, a violation could be continuous or instantaneous in such a period. Similarly, the inspection in some contexts could be an "audit" of a preceding period.

0018-9286/85/0800-0723$01.00

pth inspectee ($p \in S$) is implementing an action $v_p \in V(p)$, where $V(p)$ is a finite set of actions which can be thought of as violation levels that p can commit and includes an action which corresponds to no violation at all. The inspector, on the other hand, chooses an action $i = (i_1, i_2)$ at time t, where $i_1 \in S$ is the site which will be inspected during the period $[t, t+1)$ and i_2 is the level of inspection to be performed at i_1. We assume that $\mathcal{I}$ is the finite set of inspection levels, and hence that the number of actions i open to the inspector is w, the cardinality of $S \times \mathcal{I}$. The *state* of the game at time t is the site at which the inspector has just completed an inspection. The consequences of inspectees' choices of $v_p \in V(p)$; $p = 1, \cdots, S$ and of the inspector's choice of $i \in S \times \mathcal{I}$ are the following.

i) A transition from the current state, say s, to the state i_1 (determined by i) at the time $t+1$.

ii) Rewards $r_p(v_p, i, s)$; $p = 1, \cdots, S$ earned by each of the inspectees for the period $[t, t+1)$. Note that these depend only on the actions of the particular inspectee and the inspector, and could easily be negative if the inspectee is caught committing a violation.

iii) Reward $r_{II}(v_1, \cdots, v_S, i, s)$, earned by the inspector for the period $[t, t+1)$ (again interpreted as a loss whenever it is negative).

Remark 2.1: Since travel costs (or costs of moving an "inspection team") are likely to be of importance to the inspector, it is natural that his reward function depends on the argument s, the site of the last inspection. The situations where s affects the rewards of the inspectees are less obvious but certainly conceivable: for instance, if the inspector is obliged to make a long trip from s to i_1, the time "wasted" on travel may enable the inspectees to accomplish a "complete" violation of a certain type.

The game we have just described belongs to a special class of games which are sometimes called "single-controller stochastic games." This class was first studied by Stern [36] and has received some attention in recent years (see, for instance, [30], [38], [21], [22], [13], and [10]). Of course, in the TIM, the inspector is the "single" controller since it is only his actions which determine the transitions from state to state (i.e., from site to site). In general, the law governing the possible transitions is stochastic (not deterministic as in TIM), and is often called "the law of motion" of the stochastic game.

A *stationary strategy* f_p of the pth inspectee is a probability vector on the actions in $V(p)$, whose hth component is $f_p(s, h) = Pr\,\{p$ chooses action h whenever the inspection just completed occurred at site $s\}$.

Similarly, we can define a set of probability vectors $g(s)$ whose ith component $g(s, i)$ denotes the probability that the inspector chooses the action $i = (i_1, i_2)$ whenever he completes an inspection at s. Thus, a stationary strategy for the inspector can be regarded as the composite vector $g = (g(1), g(2), \cdots, g(S))$. By contrast, *general* (i.e., not necessarily stationary) *strategies* in a stochastic game can depend not only on the current state but also on the complete history of the game up to the current state. The class of *Markov strategies* lies in between stationary and general strategies, in that the players' randomized strategies at time t, can depend on t as well as on the current state s. Let F_p, FM_p, and FS_p denote the sets of all general, Markov, and stationary strategies, respectively, of the pth inspectee. Clearly, $F_p \supset FM_p \supset FS_p$. Similarly, define G, GM, and GS as the corresponding sets of strategies for the inspector. Once an $(S+1)$-tuple of strategies $(f_1, \cdots, f_S, g)$ is given, the expected gain/loss $\Pi_t^\nu(f_1, \cdots, f_S, g, s)$ to the νth player for the tth stage (i.e., $[t, t+1)$), given that the initial state was s, is well defined. The two types of stochastic games which we shall consider here, are distinguished by the manner in which the players evaluate a stream of expected gains $(\Pi_1^\nu, \Pi_2^\nu, \cdots)$. They are the following. a) The *T-stage* or *finite stochastic games*, if the payoff to the νth player resulting from the use of strategies $(f_1, \cdots, f_s, g)$ is given by

$$\Phi^\nu(f_1, \cdots, f_S, g, s) = \sum_{t=1}^{T} \Pi_t^\nu(f_1, \cdots, f_S, g, s)$$

where $\nu = 1, 2, \cdots, S+1$, and $s \in S$ is the initial state. Here T is, of course, the number of stages after which the process stops.

b) The *undiscounted* or *limiting average reward stochastic games*, if the payoff to the νth player is given by

$$\Phi^\nu(f_1, \cdots, f_S, g, s) = \liminf_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} \Pi_t^\nu(f_1, \cdots, f_S, g, s)$$

where $\nu = 1, \cdots, S+1$ and $s \in S$.

In the sequel it will be convenient to use the following more compact notation. Let $F = X_{p=1}^S F_p$ and $\Gamma = F \times G$, so that $\gamma \in \Gamma$ represents an $(S+1)$-tuple of general strategies. The symbols FM, FS, ΓM, and ΓS will be given analogous meanings in terms of Markov and stationary strategies, respectively. For any $\gamma \in \Gamma$ we shall denote by (γ, p, f_p) the member of Γ obtained from γ by changing the coordinate corresponding to the pth inspectee to f_p, with $(\gamma, S+1, g)$ defined similarly.

We shall say that $\gamma^o \in \Gamma$ is a *(Nash) equilibrium point* for the TIM game if for every $f_p \in F_p$; $p = 1, \cdots, S$ and $g \in G$, the relations

$$\Phi^p(\gamma^o, p, f_p, s) > \Phi^p(\gamma^o, s); \qquad p = 1, \cdots, S, \; s \in S$$

$$\Phi^{S+1}(\gamma^o, S+1, g, s) > \Phi^{S+1}(\gamma^o, s); \qquad s \in S$$

are all *false*.[3] That is, if we think of γ^o as the players' "present" choice of strategies, then by the above definition no player has an incentive to deviate unilaterally from his present choice.

We conclude this section by remarking that the issue of existence of Nash equilibria in noncooperative stochastic games has now been studied quite extensively (see, for instance, [15], [32], [33], and [7]). Nonetheless, for the class of undiscounted stochastic games there is still no general existence theorem even when the state and action spaces are finite. However, under some additional assumptions (see, for instance, [19] or [7]) Nash equilibria are known to exist. We do not discuss these conditions here since TIM does not satisfy them. Instead, we shall show that under mild conditions there is no loss of generality in replacing the set of S inspectees by a single "aggregated" inspectee, thus creating a two-person, noncooperative, undiscounted game which is known to possess *stationary Nash equilibria* (see [30]) which are also equilibria for the original game. Of course, the existence of *Markov Nash equilibria* in the T-stage stochastic games is well known (see, for instance, [19, ch. 9]).

III. Inspectee Aggregation and its Consequences

In this section we shall consider the results of assuming that the S inspectees in the TIM form an aggregated player I in a two-person game, with the inspector acting as player II. We shall assume that the strategy space of this aggregated player I is $F = X_{p=1}^S F_p$ (as in Section II), and that G is the strategy space for player II. Let $\Gamma = F \times G$ and $R_p(s) = \{\Phi^p(\gamma, s) | \gamma \in \Gamma\}$; that is, $R_p(s)$ denotes the set of possible rewards to the pth inspectee if the initial state is s. Further, let $R_I(s) = X_{p=1}^S R_p(s)$ and let the triple (r_I, p, r_p') denote the "outcome" of the aggregated inspectee I, when the pth component of r_I (i.e., the reward of the pth inspectee) is changed from r_p to r_p' (note that the argument s is suppressed here to simplify the notation, that is, $r_I = r_I(s) \in R_I(s)$, etc.).

Now, for any $\gamma = (f, g) \in \Gamma$ and initial state s, the payoff function of the inspector will still be

$$\Phi^{II}(\gamma, s) = \Phi^{S+1}(\gamma, s) \tag{3.1}$$

while the payoff of the aggregated inspectee will be taken to be

$$\Phi^I(\gamma, s) = \psi[\{\Phi^p(\gamma, s) : p \in S\}] \tag{3.2}$$

[3] This definition applies to both types of games defined in a) and b).

where $\psi:\mathbb{R}^S \to \mathbb{R}^1$ and is *stricly monotone* in each of its arguments; that is, for any $p \in S$ and $s \in S$

$$r_p'(s) > r_p(s) \text{ implies } \psi(r_I,\ p,\ r_p') > \psi(r_I) \tag{3.3}$$

for all $r_I \in R_I(s)$, where $r_p'(s)$, $r_p(s) \in R_p(s)$.

The last condition reflects the simple notion that if outcome $r_p'(s)$ is preferred to $r_p(s)$ individually, then the unilateral deviation from $r_p(s)$ to $r_p'(s)$ in the "aggregated outcome" r_I will be advantageous for the aggregated inspectee as well. It should be noted that in TIM "natural" aggregation functions such as

$$\psi(r_I) = \sum_{p=1}^{S} \lambda_p r_p(s) \tag{3.4}$$

(with λ_p's positive), satisfy (3.3).

Once the inspectees are aggregated as above to form player I, we have the "aggregated traveling inspector model" or ATIM. Note that it is only the inspectees and their payoffs which are aggregated and not their strategy spaces. Thus, F, FM, and FS are still the sets of general, Markov, and stationary strategies for player I in ATIM.

The above method of aggregating payoffs is "global" in the sense that it is the payoffs for the whole game which are aggregated via the function ψ, however, it says nothing about the manner in which the payoffs at each stage are combined. Hence, from now on, we shall impose the *consistency assumption of ATIM*. There exists a stage-by-stage aggregation of the rewards of the inspectees which induces a two-person game whose set of Nash equilibria in the smallest class of strategies (general, Markov, or stationary) for which it is nonempty coincides with the corresponding set of Nash equilibria of ATIM.

Note that in most likely applications we would want the above "local" aggregation to be performed by the same aggregation function as that in (3.2) which was used to aggregate at the "global" level. For instance, for the T-stage payoff criterion a), local aggregation via a function of the form (3.4) is equivalent to global aggregation via the same function. For the undiscounted payoff criterion b), the same is true as long as all the players use stationary strategies.

Theorem 3.1:

i) The sets of Nash equilibria of TIM and ATIM coincide.

ii) If the undiscounted payoff criterion b) is used, then both TIM and ATIM possess Nash equilibria in stationary strategies.

iii) If the T-stage payofff criterion a) is used, then both TIM and ATIM possess Nash equilibria in Markov strategies.

Proof: i) The proof of this will closely follow the line of argument used to prove [16, Theorems 1 and 2] (the even more general results on player aggregation due to Goldman [17] could also be invoked). The essential observation is this: due to our definition of the rewards $r_p(v_p,\ i,\ s)$; $p = 1, \cdots, S$, of the inspectees, and due to the fact that only the inspector determines the transitions from state to state, the expected rewards for the tth stage simplify to $\Pi_t^p(f_1, \cdots, f_S, g, s) \equiv \Pi_t^p(f_p, g, s)$; $p = 1, \cdots, S$, for every initial state s. Consequently, each inspectee's payoff function also depends only on his own strategy, the inspector's strategy, and the initial state; that is, for $p = 1, \cdots, S$

$$\Phi^p(\gamma,\ s) \equiv \Phi^p(f_p,\ g,\ s). \tag{3.5}$$

Now, let $\gamma^0 = (f_1, \cdots, f_S^0, g^0)$ be an equilibrium point of ATIM and suppose that it is not an equilibrium point of TIM. Then either for some $s \in S$, $p \in S$ and $f_p \in F_p$

$$\Phi^p(f_p,\ g^0,\ s) = \Phi^p(\gamma^0,\ p,\ f_p,\ s) > \Phi^p(\gamma^0,\ s) = \Phi^p(f_p^0,\ g^0,\ s) \tag{3.6}$$

or for some $s \in S$ and $g \in G$

$$\Phi^{S+1}(\gamma^0,\ S+1,\ g,\ s) > \Phi^{S+1}(\gamma^0,\ s). \tag{3.7}$$

If (3.6) held, then the fact that (for all $k \neq p$) $\Phi^k(\gamma^0, p, f_p, s) \equiv \Phi^k(\gamma^0, s)$ [see (3.5)], together with monotonicity of ψ imply that $\Phi^I(\gamma^0, p, f_p, s) > \Phi^I(\gamma^0, s)$, thus, contradicting the hypothesis that γ^0 is an equilibrium point of ATIM. If (3.7) held, then (3.1) immediately yields the same contradiction.

Conversely, suppose that γ^0 is an equilbrium of TIM but not of ATIM. Again (3.1) immediately shows that the inequality $\Phi^{II}(\gamma^0, II, g, s) > \Phi^{II}(\gamma^0, s)$ is impossible for any $g \in G$ and $s \in S$. Suppose then that for some $s \in S$ and $f = (f_1, \cdots, f_S) \in F$ the inequality

$$\Phi^I(\gamma^0,\ I,\ f,\ s) > \Phi^I(\gamma^0,\ s) \tag{3.8}$$

holds, where in the above we regard γ^0 as the pair (f^0, g^0). Now, let us define the set $I^+(s) = \{k \in S | \Phi^k(\gamma^0, k, f_k, s) > \Phi^k(\gamma^0, s)\}$. The sets $I^-(s)$ and $I^0(s)$ are defined similarly by replacing $>$ with $<$ and $=$, respectively. We shall show that $I^+(s)$ is nonempty, thus contradicting the hypothesis that γ^0 is an equilibrium point of TIM. Suppose then that $I^+(s) = \{\ \}$. If $I^-(s)$ also empty, then by (3.5) for every $k \in S$, $\Phi^k(f, g^0, s) \equiv \Phi^k(f_k, g^0, s) \equiv \Phi^k(\gamma^0, k, f_k, s) \equiv \Phi^k(\gamma^0, s)$ which contradicts (3.8) [see (3.2)]. If $I^-(s)$ is nonempty, then repeated application of (3.3) to change $\{f_k^0 | k \in I^-(s)\}$ "one by one" to $\{f_k | k \in I^-(s)\}$, together with (3.5) imply that $\Phi^I(\gamma^o, s) > \Phi^I(\gamma^o, I, f, s)$ contradicting (3.8).

ii) Of course, part i) is useful only if the sets of Nash equilibria of these games are nonempty. However, we know from [30, Section 5] that in every two-person, single-controller undiscounted stochastic game there exists $\gamma^o = (f^o, g^o) \in FS \times GS$ which is a Nash equilibrium in ATIM, and hence also in TIM by i), and the consistency assumption.

iii) This follows in an analogous way from [19, Theorem 9.5]. □

Corollary 3.2: Let $\gamma^o = (f^o, g^o) \in \Gamma$ be any Nash equilibrium of the undiscounted TIM game, and let $\beta(s) = \Phi^{II}(\gamma^o, s)$ for $s = 1, \cdots, S$. Then $\beta(s) = \beta$ for every s; that is, the equilibrium rewards of the inspector are independent of the initial state.

Proof: Suppose to the contrary that $\beta(s) > \beta(s')$ for some pair of initial states s and s'. Then define a strategy $\hat{g} \in G$ for the inspector as follows. If the initial state is s' go to s next and perform any inspection there, and thereafter use g^o. If, on the other hand, the initial state is not s' then use g^o throughout the game. Since we are using limiting average rewards it should be clear that

$$\Phi^{II}(f^o,\ \hat{g},\ s') = \Phi^{II}(f^o,\ g^o,\ s) = \beta(s).$$

But since γ^0 is an equilibrium point

$$\beta(s') = \Phi^{II}(f^o,\ g^o,\ s') \geq \Phi^{II}(f^o,\ \hat{g},\ s') = \beta(s)$$

which yields the desired contradiction. □

Since in TIM the duration of the process would typically be large but unknown, we shall concentrate on the undiscounted version defined by b) of Section II. Further, we shall be particularly interested in the set of stationary Nash equilibria in TIM (or equivalently in ATIM). However, by the consistency assumption of ATIM it is sufficient to consider the set of stationary Nash equilibria of a two-person, single-controller stochastic game whose "local" rewards of player I are the appropriate aggregations of the "local" rewards of the S inspectees (again, aggregations of the form (3.4) can be used). Hence, from now on, we shall not differentiate between ATIM and its "locally" aggregated counterpart.

More precisely, if we define EPS to be the set of all stationary (Nash) equilibria of ATIM (which is nonempty by Theorem 3.1), then it is possible to give a finite characterization of this set as follows. There exists a finite set $\bar{\chi} = \{f^i \in FS\}$ of *extreme equilibrium strategies* for player I such that with every subset χ of $\bar{\chi}$ we can associate (a possibly empty) set $E(\chi) = \{g \in GS | (f,$

$g) \in EPS$ for all $f \in \chi\}$, it then turns out that

$$EPS = \bigcup_{\chi \subset \bar{\chi}} [C(\chi) \times E(\chi)] \tag{3.9}$$

where $C(\chi)$ is the convex hull of χ. This result is proved in [10, Section 3] and is an analog of a classical result for bimatrix games due to Kuhn [24]. Furthermore, it is shown in [10] that the members of $\bar{\chi}$ form a subset of the set of extreme points of a polyhedral set defined by a system of linear constraints whose coefficients are given by the data of the game. Hence, if these data are rational the set $\bar{\chi}$ can be constructed by one of a number of well-known finite algorithms (for a review of some of these, see [29]). Further, each of the sets $E(\chi)$ in (3.9) is itself the image of a certain polyhedral set under a finitely executable transformation (see [10, Section 4]).

These results show that the set of *all* stationary equilibria of ATIM (and, hence, of TIM) can be fully characterized by finite algorithms provided only that the data are rational. However, these algorithms depend on algorithms for enumerating all vertices of polyhedral sets (e.g., see [29]), and hence could prove impractical in many applications. If we are interested in computing just one equilibrium point of ATIM, the method which promises to be most efficient is that of solving a certain quadratic program and then converting the solution to a member of EPS. This approach, which was proposed by Filar [11], generalizes a classical result about bimatrix games which is due to Mangasarian and Stone [27].

Remark 3.3: One major factor affecting the potential solvability of ATIM is the exponential growth of the action space of the aggregated inspectee. Note that a typical action of such player I is an S-tuple $v = (v_1, \cdots, v_S)$; $v_p \in V(p)$. Hence, I possesses $\Pi_{p=1}^{S} |V(p)| = u_I$ such actions! However, this growth of u_I with S exhibits what might be called a "natural" increase of the difficulty of the problem: a single inspector could not be expected to be very effective against many inspectees. There are various approaches which might prove useful in alleviating this "curse of dimensionality"; however, these lead to open theoretical problems, some of which deserve deeper investigation (see also Section V).

One problem which arises whenever Nash equilibria are used as a solution concept is that of choosing between alternative equilibrium points, since they typically result in different payoffs. In an inspector/inspectee context it is not unreasonable to consider what we shall call the *inspector's optimal equilibria*, namely, those equilibria which maximize his payoff function. More precisely, with each $\gamma = (f, g) \in EPS$ we can associate $\beta(\gamma) = \Phi^{II}(\gamma, s)$ (recall that $\beta(\gamma)$ is independent of s by Corollary 3.2). Then the inspector's *optimal equilibria* are the solutions of the maximization problem

$$\sup \beta(\gamma)$$

$$\text{subject to: } \gamma \in EPS. \tag{3.10}$$

The next result shows that optimal equilibria of undiscounted ATIM do exist.

Theorem 3.4: Let $\widehat{EP} = \{\hat{\gamma} \in EPS | \beta(\hat{\gamma}) = \max \beta(\gamma)$ over $EPS\}$. Then $\widehat{EP}$ is nonempty, and a member of $\widehat{EP}$ can be computed by a finite method provided only that the data of the process are rational.

Remark 3.5: The proof of the above theorem is outlined in the Appendix. It must be emphasized, however, that the finite method for finding the inspector's optimal equilibrium, is again based on enumerating all vertices of polyhedral sets.

Of course, Theorem 3.4 merely establishes an upper bound in the inspector's reward resulting from the use of any stationary equilibrium point. However, if we choose some $\hat{\gamma} = (\hat{f}, \hat{g}) \in \widehat{EP}$ and consider the set $E(\hat{g}) = \{\bar{f} \in FS | (\bar{f}, \hat{g}) \in EPS\}$ [note that $\hat{f} \in E(\hat{g})$], then we can only assert that $\Phi^{II}(\bar{f}, \hat{g}, s) \leq \beta(\hat{\gamma})$. Since it is natural to expect that the use of strategy $\hat{g}$ by the inspector will induce a truly noncooperative aggregated inspectee to play some $\bar{f} \in E(\hat{g})$, the fact that in general the strict inequality $\Phi^{II}(\bar{f}, \hat{g}, s) < \beta(\hat{\gamma})$ is possible shows that our optimal Nash equilibrium point $\hat{\gamma}$ may not be "enforceable" by the inspector.

We are thus led to the following definition. An equilibrium point $\gamma^o = (f^o, g^o) \in EPS$ is *enforceable* for the inspector if $\Phi^{II}(\bar{f}, g^o, s) = \bar{\beta}$, a constant, for all $\bar{f} \in E(g^o)$, and $s \in S$.

Remark 3.6: It might be worth mentioning that the assumption that the game is noncooperative (i.e., that each player tries to maximize his own reward function independently) may not hold in many situations, for instance, if the possibility of a "corrupt" inspector is allowed. Nonetheless, it is a rather natural assumption to make initially at least and, if accepted, it leads us to the (Nash) equilibrium point solutions with their inherent difficulties. Since in the ATIM model above it is not unreasonable to assume that the inspector's equilibrium strategy g^0 (as above) would be either known, or could be estimated by the aggregated inspectee, the set $E(g^0)$ constitutes his "rational" choices of stationary strategies. The fact that different members of this set can, in general, result in different payoffs to the inspector (who is using g^0) suggests the need for determining whether a given equilibrium point is "enforceable" in the sense defined above.

Lemma 3.7: Let $\gamma^o = (f^o, g^o) \in EPS$, then there exists a finite method for checking whether γ^o is enforceable for the inspector, provided only that the data of the process are rational.

Remark 3.8: One consequence of the above result (the proof of which is outlined in the Appendix) is that if, in particular, an optimal equilibrium point $\hat{\gamma} = (\hat{f}, \hat{g})$ is found to be enforceable for the inspector, then this provides strong argument for the inspector to actually adopt the strategy $\hat{g}$ because now he can guarantee himself the reward of max $\beta(\gamma)$ over EPS, against "a rationally behaving" noncooperative aggregated inspectee.

We now mention a result which shows that in an important special case all stationary equilibria are enforceable for the inspector, and with an identical payoff.

Lemma 3.9: Assume that the reward functions of the inspector and the aggregated inspectee are such that for every $g \in GS$

$$X(g) = \{f_g | \Phi^I(f_g, g, s) = \max_{FS} \Phi^I(f, g, s) \text{ for all } s \in S\}$$

$$= \{f_g | \Phi^{II}(f_g, g, s) = \min_{FS} \Phi^{II}(f, g, s) \text{ for all } s \in S\}.$$

Then for any $(\hat{f}, \hat{g}) \in EPS$, the corresponding payoff to the inspector is the constant

$$\alpha^o = \max_{GS} \min_{FS} \Phi^{II}(f, g, s),$$

for all $s \in S$.

Remark 3.10: The above lemma is proved in the Appendix. It should be noted that the conditions of this lemma are satisfied in the case where $\Phi^{II}(f, g, s) = -\Phi^I(f, g, s) - C(g, s)$; that is, when the inspector's reward is the inspectee's loss (e.g., fines) except for the term that depends only on the inspector's strategy (e.g., travel/inspection costs).

Remark 3.11: For the T-stage payoffs defined by a) in Section II, TIM can still be solved by aggregating the inspectees and solving ATIM, which is now a T-stage, two-person stochastic game. Thus, at each stage the two players play one of S possible bimatrix games. In particular, if at time t the state of the game is s, and the inspectees and the inspector choose actions: $v = (v_1 \cdots v_S)$ and $i = (i_1, i_2)$, respectively, then their corresponding current rewards are $r_I(v, i, s)$ and $r_{II}(v, i, s)$, where $r_I(v, i, s)$ is some aggregation of $r_p(v_p, i, s)$'s for $p \in S$ consistent with (3.2). The set of all such pairs of rewards as v and i range over all possible actions of players I and II constitute the sth bimatrix game. Of course, the next bimatrix game to be played is determined by the first component of i. Since an equilibrium point of a bimatrix game can be found either by the method of Lemke and Howson

[25], or by that of Mangasarian and Stone [27], an equilibrium point of ATIM can now be computed using the usual backward recursion of dynamic programming. Of course, the equilibrium strategies will now depend on the stage of the game as well as on the current state; that is, they will belong to the class of Markov strategies. The details of this method can be easily reconstructed by following the proof of Theorem 9.5 in [19].

IV. ATIM as a Zero-Sum Game

When the aggregated traveling inspector model is considered from the point of view of the inspector, one method of analysis which ought not to be overlooked is the *minimax approach*. That is, for each potential outcome $(\boldsymbol{v}, \boldsymbol{i}, s)$ the inspector will now assume that his own loss (negative reward): $l(\boldsymbol{v}, \boldsymbol{i}, s) \equiv -r_{II}(\boldsymbol{v}, \boldsymbol{i}, s) \equiv r_I(\boldsymbol{v}, \boldsymbol{i}, s)$, also represents the reward to the aggregated inspectees. This "zero-sum" assumption of directly opposing interests, while conservative, may be one that the inspector cannot afford not to make! In addition, it has the advantage of eliminating what is perhaps the greatest single obstacle in the modeling of such a process from the viewpoint of the "inspection agency": the uncertainty as to what exactly the inspectee's reward function is. The resulting game now becomes a two-person, zero-sum, single-controller stochastic game. Such a game with the undiscounted payoff criterion is now solvable in stationary strategies by efficient linear-programming algorithms (see, for instance, [38] and [22]). In the context of ATIM these algorithms have been further simplified and implemented by Filar and Schultz [12]. Practical numerical solution has been obtained in over 30 examples with up to eight inspectees and three violation levels per inspectee (this corresponds to 3^8 pure actions $\boldsymbol{v}$ available to the aggregated inspectee at every stage of the game).[4]

When the T-stage payoff criterion is used, the model is still easily solvable by backward recursion such as that commonly used in dynamic programming. A technique almost identical with that of Charnes and Schroeder [5, p. 309] will yield the value vector and optimal strategies for both players. Of course, at each iteration of this method a set of S matrix games has to be solved. As in the T-stage noncooperative case, the solution strategies will typically be Markov.

We shall now illustrate the traveling inspector model by a simple numerical example taken from Filar and Schultz [12] which we call the "gun smuggling problem."

We consider the situation where the inspector's task is the prevention/capture of contraband, say guns, from entering the region to which he is assigned. We assume that shipments of guns can enter the region at only S sites (e.g., bridges, roads, ports, etc.). Of course, the inspector can be present at only one site at any given time. One stage will be a 24 h period beginning at midnight. We suppose that the inspectee knows at which site the inspector begins the stage. The inspectee then sends up to S shipments of guns (at most one for each site) which were stored at some central cache and which will arrive at their assigned site (entry point) at a random time (due to local conditions) which is uniformly distributed [5] between noon and the following midnight.

If the inspector is present at a site when a shipment of k guns arrives, they will be seized and his gain will be $2k$. This reflects the notion that such a capture not only deprives "the enemy" of k guns but also delivers k guns to the inspector's side. A failure to capture this shipment will amount to a loss of k to the inspector. The inspector's travel costs from site s to site s' are given as fractions of the interval from noon to midnight during which all sites are left unguarded, and will be denoted by $\gamma(s, s')$; $s, s' \in \{1, \cdots, S\}$. Let $g(v_p)$ denote the number of guns in a shipment destined to enter at site p which corresponds to a violation at level v_p being committed at that site. The inspector's typical decision (see Section II) is now $\boldsymbol{i} = (i_1, i_2) = (s', 1)$, since we are assuming only a single level of inspection (i.e., interception of the shipment). Now, with each site p we can associate a (ficticious) inspectee whose reward function on a given day is defined by

$$
\begin{aligned}
&r_p(v_p, i, s)\\
&\equiv r_p(v_p, (s', 1), s)\\
&= \begin{cases} g(v_p) & \text{if } p \neq s' \\ \gamma(s, p)g(v_p) - 2(1-\gamma(s, p))g(v_p) & \text{if } p = s'. \end{cases}
\end{aligned}
$$

The above reward represents the net expected gain to the "inspectee's side" (measured in number of guns) associated with site p for that day. Hence, if on a particular day a vector of shipments $\boldsymbol{v} = (v_1, \cdots, v_S)$ was dispatched to enter at sites $1, 2, \cdots, S$, respectively, and if the inspector decided to travel to site s' (assuming that he was at s last), then the inspector's loss for that day will be defined by

$$
l(v, (s', 1), s) = \sum_{p=1}^{S} r_p(v_p, (s', 1), s) \tag{4.1}
$$

that is, the net expected loss (in guns) by the "inspector's side" for that day.

Next, we summarize the optimal solution to a 3-state example of this problem. The data are as given in Tables I and II.

Table I indicates that only one type of shipment (of 200 guns) can enter through site 1 each day, while two different types of shipments can enter through sites 2 and 3 ($v_p = 0$ corresponds to no shipment dispatched to site p). The value $\gamma(1, 3) = 0.6$ in Table II indicates that the inspector uses 60 percent of the available inspection time traveling from site 1 to site 3. The underlying, zero-sum, single-controller stochastic game now consists of three 18×3 payoff matrices L^p; $p \in \{1, 2, 3\}$ whose entries are given by (4.1).

This example was solved using an algorithm described in [12]. The minimax optimal stationary strategies are given in Figs. 1 and 2.

The optimal strategy given in Fig. 1 for the inspectee should be interpreted as follows. If the inspector was observed at, say, site 3 at the end of the last stage, then the inspectee should choose composite actions (1, 0, 2), (1, 1, 0), and (1, 1, 1) with probabilities 0.214, 0.729, and 0.057, respectively. Of course, the action (1, 0, 2) means that shipments of 200 and 250 guns will be directed to enter through sites 1 and 3, respectively, and no shipment is sent to site 2 (see Table I). Similarly, the optimal stationary strategy for the inspector should be interpreted as follows. Whenever he just completed an inspection at, say, site 2, he should then go to one of the sites 1, 2, or 3 with probabilities 0.190, 0.333, and 0.476, respectively. The value of the game $\theta = 148.57$ and it represents the long-run average net gain of guns per day by the inspectee's side when the optimal strategies are used. This can be contrasted with the net gain of (200 + 150 + 250) = 600 which the inspectee could achieve if there were no inspector.

V. Possible Extensions and Some Open Problems

The traveling inspector model presented in the preceding sections easily lends itself to many modifications and generalizations. Some of these can be handled by already existing techniques while for others no adequate treatment appears to be known.

The problem of what to do when equilibrium/optimal strategies for the inspector dictate that certain plants should never be visited can perhaps be modeled with the help of additional constraints. In the infinite horizon model, for instance, these constraints could

[4] Problems with eight inspectees took approximately 9 min of CPU time on the VAX 11/780 to yield optimal strategies for both the inspector and the inspectees.

[5] Actually any distribution, even one which depends on the size of the shipment and its destined site could be easily incorporated in this model. Similarly, the length of a stage is flexible.

TABLE I
VIOLATION COSTS

$g(v_p)$	p=1	p=2	p=3
$v_p=0$	0	0	0
$v_p=1$	200	100	200
$v_p=2$		150	250

TABLE II
TRAVEL COSTS

$\gamma(s,s')$	s'=1	s'=2	s'=3
s=1	0	.3	.6
s=2	.3	0	.3
s=3	.6	.3	0

Model: Gun Smuggling I Policy: Average Optimal for violator

INSPECTOR AT SITE	VIOLATION LEVELS AT SITES 1 2 3	PROBABILITY OF USE
1	0 0 2	0.143
1	0 2 2	0.282
1	1 2 2	0.575
2	1 0 0	0.233
2	1 2 0	0.052
2	1 2 2	0.714
3	1 0 2	0.214
3	1 1 0	0.729
3	1 1 1	0.057

Fig. 1. Optimal stationary strategy for the inspectee.

Model: Gun Smuggling I Policy: Average Optimal for inspector.

INSPECTOR AT SITE	NEW SITE	INSPECTION LEVEL	PROBABILITY OF USE
1	1	1	0.333
1	2	1	0.476
1	3	1	0.190
2	1	1	0.190
2	2	1	0.333
2	3	1	0.476
3	1	1	0.190
3	2	1	0.476
3	3	1	0.333

Fig. 2. Optimal stationary strategy for the inspector.

take the form of lower bounds on linear functions of the "long-run average state-action frequencies." Under relatively mild assumptions it has been shown by Hordijk and Kallenberg [22] that the so constrained zero-sum undiscounted stochastic game has a value (if the problem is feasible, of course) and both the inspector and the inspectee possess optimal strategies. Further, an algorithm for computing these strategies is also given in [22]. For the noncooperative constrained game no such algorithm is yet available. If the losses at subsequent periods are discounted by some discount factor $\beta \in [0, 1)$, then all of the stochastic game formulations of Section III remain solvable by algorithms that are even simpler than those for the average-reward models. If the duration of the period between successive inspections is not a constant but a random variable, the distribution of which can depend on the last site visited and on the inspector's last decision, the zero-sum stochastic game formulations can be replaced by the corresponding zero-sum semi-Markov game formulations, which under some assumptions are solvable by algorithms similar to those mentioned previously (see [23, ch. 7]).

In Section III we briefly mentioned algorithms, based on the results of [10] and [11], for finding the equilibria of two-person, single-controller, stochastic games. The development of truly efficient implementations of these (or similar) methods is a problem whose solution would permit numerical experimentation with larger scale examples of the traveling inspector model. While in the problem of generating *all* "extreme" equilibria we are unlikely to avoid the need to generate all vertices of certain polyhedral sets, the conventional wisdom that algorithms involving an enumeration of such vertices are bad may not apply in our case if full advantage is taken of the special structure of the problem. The latter statement is based on encouraging numerical results reported by Mangasarian [26] in the special (simpler) case of bimatrix games.

Finally, it is tempting to impose the "leader-follower" structure on the aggregated traveling inspector model, with the inspector as the leader since his decisions are presumably far more detectable by his opponent than vice versa (otherwise the inspecting would be unnecessary). The "Stackelberg equilibrium" solutions of such games could then be investigated. While a detailed analysis of such "Stackelberg TIM" models is a subject for future research, it follows that under the conditions of Lemma 3.9 (proved in the Appendix) the inspector's gain resulting from any stationary equilibrium point coincides with his Stackelberg equilibrium reward. More precisely, following the conventions of Basar and Olsder [3] (see ch. 4) we see that for each $g \in GS$ player I's *optimal response set* is precisely the set $X(g)$ of Lemma 3.9. Then the inspector's *Stackelberg reward* is, for each $s \in S$, given by

$$\max_{GS} \min_{X(g)} \Phi^{II}(f, g, s) = \max_{GS} \min_{FS} \Phi^{II}(f, g, s) = \alpha^o$$

where the last equality follows from Lemma 3.9.

APPENDIX

In this section we shall sketch the proofs of those results of Section III which were not justified there. The first of these depends heavily on the results derived in [10].

OUTLINE OF THE PROOF OF THEOREM 3.4

Let $\bar{\chi} \subset FS$ be as in (3.9) (it is defined precisely in [10, Section III]). By [10, Theorem 3.1], the elements of $\bar{\chi}$ can be enumerated by generating all vertices of a certain polyhedron. However, some members of $\bar{\chi}$ can be "dummies" in the sense that if $\chi^i = \{f^i\}$ then $E(\chi^i) = \phi$; that is, f^i is not part of any equilibrium strategy. These dummy members of $\bar{\chi}$ can be easily identified using techniques developed in [10].

Suppose now that $\chi^* = \{f^1, \cdots, f^m\}$ is the set of "nondummy" members of $\bar{\chi}$ and define

$$\beta^i = \max_g \Phi^{II}(f^i, g, s) \tag{6.1}$$

for every $i = 1, \cdots, m$ and every initial state s. Note that the above definition is meaningful since with $f^i \in FS$ fixed in (6.1), the above maximization problem is equivalent to solving an undiscounted Markovian decision process with respect to $g \in G$. It is well known (for instance, see [23]) that this maximum is achieved at a nonrandomized stationary strategy which can be computed by linear programming. In addition, an argument similar to that in Corollary 3.2 will show that the maximum in (6.1) is independent of the initial state s. Now define

$$\beta^* = \max_{1 \le i \le m} \beta^i. \tag{6.2}$$

It is now possible to show with the help of [10, Lemma 2.1 and Proposition 4.1] that

$$\beta^* = \max \{\beta(\gamma) | \gamma \in EPS\}$$

and in the process the maximizing stationary equilibrium point is constructed by a finite algorithm. □

Proof of Lemma 3.7: Recall that with $\gamma^o = (f^o, g^o) \in$

EPS we associated the set of stationary policies $E(g^o) = \{\bar{f} \in FS | (\bar{f}, g^o) \in EPS\}$ for the aggregated inspectee. By an argument such as that in [10, Section III] it can be seen that $E(g^o)$ is a compact convex set. Further, by [10, Theorem 3.1] if f^i is an extreme point of $E(g^o)$, then $f^i \in \bar{\chi}$. Thus, to check the enforceability of (f^o, g^o) by the inspector, it is sufficient to check whether $\Phi^{II}(f^i, g^o, s)$ is constant for those members of $\bar{\chi}$ which are also in $E(g^o)$. The latter statement follows from the fact that due to the single controller assumption for every $s \in S$

$$\Phi^{II}(f^o, g^o, s) = \sum_i \lambda^i \Phi^{II}(f^i, g, s)$$

where the summation is over the extreme points of $E(g^o)$, and λ^i's are nonnegative numbers summing to 1. □

Proof of Lemma 3.9: Let $\gamma^* = (f^*, g^*) \in EPS$ and $\beta(\gamma^*) = \Phi^{II}(f^*, g^*, s)$ (independent of s by Corollary 3.2). Also consider a two-person, zero-sum, single-controller stochastic game with $\Phi^{II}(f, g, s)$ as the payoff kernel and player II as the maximizer. It follows from the results in [30], [22], or [38] that such a game possesses an optimal stationary strategy pair $\gamma^o = (f^o, g^o) \in FS \times GS$ computable by linear programming. By an argument similar to that in Corollary 3.2 it can be shown that the value of this game is independent of the initial state, that is,

$$\Phi^{II}(f^o, g^o, s) = \alpha^o = \max_{GS} \min_{FS} \Phi^{II}(f^o, g^o, s). \tag{6.3}$$

It now follows from the facts that $\gamma^* \in EPS$ and γ^o is an optimal pair in the "$\Phi^{II}(\gamma, s)$, 0-sum game" that for every $s \in S$

$$\beta(\gamma^*) = \Phi^{II}(f^*, g^*, s) \geq \Phi^{II}(f^*, g^o, s) \geq \Phi^{II}(f^o, g^o, s) = \alpha^o. \tag{6.4}$$

In order to prove the opposite inequality note that $f^* \in X(g^*)$, and hence by the hypothesis and (6.3), we have

$$\alpha^o \geq \min_{FS} \Phi^{II}(f, g^*, s) = \Phi^{II}(f^*, g^*, s) = \beta(\gamma^*) \tag{6.5}$$

for every $s \in S$. In view of (6.4) and (6.5) $\beta(\gamma^*) = \alpha^o$, and the Lemma holds. □

Acknowledgment

The author is indebted to A. J. Goldman for his comments and suggestions of improvements, and to the referees for their comments.

References

[1] S. C. Albright and W. Winston, "Optimizing advertising expenditures in a dynamic duopoly," *Oper. Res.*, vol. 27, pp. 682–693, 1979.

[2] A. Alj and A. Haurie, "Dynamic equilibria in multigeneration stochastic games," *IEEE Trans. Automat. Contr.*, vol. AC-28, pp. 193–203, 1983.

[3] T. Basar and G. J. Olsder, *Dynamic Noncooperative Game Theory.* New York: Academic, 1982.

[4] D. Blackwell and T. S. Ferguson, "The big match," *Ann. Math. Stat.*, vol. 39, pp. 159–163, 1968.

[5] A. Charnes and R. G. Schroeder, "On some tactical antisubmarine games," *Naval Res. Logist. Quart.*, vol. 14, pp. 291–311, 1967.

[6] A. Federgruen, "Successive approximation methods in undiscounted stochastic games," *Oper. Res.*, vol. 28, pp. 794–810, 1980.

[7] A. Federgruen, "On *N*-person stochastic games with denumerable state space," *Adv. Appl. Probability*, vol. 10, pp. 452–471, 1978.

[8] J. A. Filar, "A single loop stochastic game which one player can terminate," *OPSEARCH* 18, 1981, pp. 185–203.

[9] ——, "Ordered field property for stochastic games when the player who controls transitions changes from state to state," *J. Optimiz. Theory Appl.*, pp. 505–517, 1981.

[10] ——, "On stationary equilibria of a single-controller, stochastic game," *Math. Programming*, to be published.

[11] ——, "Quadratic programming and the single-controller stochastic game," *J. Math. Anal. Appl.*, to be published.

[12] J. A. Filar and T. A. Schultz, "Interactive solutions for the travelling inspector model and related problems," The Johns Hopkins Univ., Baltimore, MD, Oper. Res. Group Rep. Series 83-06, 1983.

[13] J. A. Filar and T. E. S. Raghavan, "A matrix game solution of the single-controller stochastic game," *Math. Oper. Res.*, vol. 9, pp. 356–362, 1984.

[14] J. A. Filar, "The travelling inspector model," Dep. Math. Sci., The Johns Hopkins Univ., Baltimore, MD, Tech. Rep. 374, 1983.

[15] A. M. Fink, "Equilibrium in a stochastic *n*-person game," *J. Sci. Hiroshima Univ. Ser. A-I*, vol. 28, pp. 89–93, 1964.

[16] A. J. Goldman and D. R. Shier, "Player aggregation in noncooperative games," *J. Res. Nat. Bur. Stand.*, vol. 85, pp. 391–396, 1980.

[17] A. J. Goldman, "Player aggregation in non-cooperative games, II," *J. Res. Nat. Bur. Stand.*, vol. 86, pp. 383–391.

[18] A. J. Goldman and M. H. Pearl, "The dependence of inspection-system performance on levels of penalties and inspection resources," *J. Res Nat. Bur. Stand.*, vol. 80B, pp. 189–236, 1976.

[19] D. P. Heyman and M. J. Sobel, *Stochastic Models in Operations Research Vol. II.* New York: McGraw-Hill, 1984.

[20] A. J. Hoffman and R. M. Karp, "On nonterminating stochastic games," *Management Sci.*, vol. 12, pp. 359–370, 1966.

[21] A. Hordijk and L. C. M. Kallenberg, "Linear programming and Markov games I. In game theory and mathematical economics," O. Moeschlin, D. Pallaschke, Eds. Amsterdam, The Netherlands: North-Holland, 1981, pp. 291–305.

[22] ——, "Linear programming and Markov games II. In game theory and mathematical economics," O. Moeschlin, D. Pallaschke, Eds. Amsterdam, The Netherlands: North-Holland, 1981, pp. 307–319.

[23] L. C. M. Kallenberg, "Linear programming and finite Markovian control problems," Amsterdam, The Netherlands, Math. Centre Tracts 148, 1983.

[24] H. W. Kuhn, "An algorithm for equilibrium points in bimatrix games," *Proc. Nat. Acad. Sci.*, vol. 47, pp. 1657–1662, 1961.

[25] C. E. Lemke and J. F. Howson, "Equilibrium points of bimatrix games," *SIAM J. Appl. Math.*, vol. 12, pp. 413–423, 1964.

[26] O. L. Mangasarian, "Equilibrium points of bimatrix games," *SIAM J. Appl. Math.*, vol. 12, pp. 778–780, 1964.

[27] O. L. Mangasarian and H. Stone, "Two-person nonzero-sum games and quadratic programming," *J. Math. Anal. Appl.*, vol. 9, pp. 348–355, 1964.

[28] M. Maschler, "A price leadership method for solving the inspector's non-constant sum game," *Naval Res. Logist. Quarter.*, vol. 13, pp. 11–33, 1966.

[29] T. H. Matheiss and D. S. Rubin, "A survey and comparison of methods for finding all vertices of convex polyhedral sets," *Math. Oper. Res.*, vol. 5, pp. 167–185, 1980.

[30] T. Parthasarathy and T. E. S. Raghavan, "An orderfield property for stochastic games, when one player controls the transition probabilities," *J. Optimiz. Theory Appl.*, vol. 33, pp. 375–392, 1981.

[31] M. Pollatschek and B. Avi-Itzhak, "Algorithms for stochastic games with geometrical interpretation," *Management Sci.*, vol. 15, pp. 399–415, 1969.

[32] P. D. Rogers, "Nonzero sum stochastic games," Ph.D. dissertation, Univ. California, Berkeley, 1969.

[33] M. J. Sobel, "Noncooperative stochastic games," *Ann. Math. Stat.*, vol. 42, pp. 1930–1935, 1971.

[34] M. J. Sobel, "Myopic solutions of Markov decision processes and stochastic games," *Oper. Res.*, vol. 29, pp. 995–1009, 1981.

[35] M. J. Sobel, "Stochastic fishery games with myopic equilibria," in *The Economics of Renewable Resources,* L. J. Mirman and D. Spulber, Eds. New York: Elsevier-North-Holland, 1981.

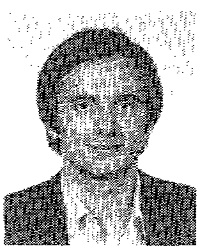
[36] M. Stern, "On stochastic games with limiting average payoff," Ph.D. dissertation, Univ. Illinois at Chicago Circle, IL, 1975.

[37] J. Van der Wal, "Discounted Markov games: Successive approximations and stopping times," *Int. J. Games Theory*, vol. 6, pp. 11–22, 1977.

[38] O. J. Vrieze, "Linear programming and undiscounted stochastic games in which one player controls transitions," *OR Spektrum*, vol. 3, pp. 29–35, 1981.

[39] O. J. Vrieze, "Stochastic games with finite state and action spaces," Ph.D. dissertation, Dep. Math., Catholic Univ., Nijmegen, The Netherlands, 1983.

[40] W. Winston, "A stochastic game model of a weapons development competition," *SIAM J. Contr. Optimiz.*, vol. 16, pp. 411–419, 1978.



Jerzy A. Filar was born in Warsaw, Poland, on August 30, 1949. He received the B.Sc. degree, with honours, from the University of Melbourne, Melbourne, Australia, in 1972, the M.Sc. degree from Monash University in 1975, and the M.A. and Ph.D. degrees from the University of Illinois at Chicago, in 1977 and 1980, respectively.

Since 1980 he has been an Assistant Professor in the Department of Mathematical Sciences at the Johns Hopkins University, Baltimore, MD. His current research interests are in game theory and Markov decision processes.

Dr. Filar received the Dr. Gurdas Chatterjee Award for the best paper published in Opsearch in 1981.