Calhoun: The NPS Institutional Archive

DSpace Repository

Faculty and Researchers | Faculty and Researchers' Publications

2008

Efficient Employment of Non-Reactive Sensors

Kress, Moshe; Jones, Jason S.; Szechtman, Roberto

2008 Efficient Employment of Non-Reactive Sensors, (with R. Szechtman and J. S. Jones), Military Operations Research, V. 13, No. 4, pp 45-57.
https://hdl.handle.net/10945/38159

defined in Title 17, United States Code, Section 101. Copyright protection is not available for this work in the United States.

Downloaded from NPS Archive: Calhoun

http://www.nps.edu/library

Calhoun is the Naval Postgraduate School's public access digital repository for research materials and institutional publications created by the NPS community. Calhoun is named for Professor of Mathematics Guy K. Calhoun, NPS's first appointed -- and published -- scholarly author.

Dudley Knox Library / Naval Postgraduate School
411 Dyer Road / 1 University Circle
Monterey, California USA 93943

Table of Contents

Volume 13, Number 4

2008

Military Applications Society

Institute for Operations Research and the Management Sciences

The Military Applications Society (MAS) is a democratically constituted professional society of open membership dedicated to the free and open pursuit of the science, engineering and art of military operations. It is the first society of the Institute for Operations Research and the Management Sciences (INFORMS). MAS advances research in military operations, fosters higher standards of practice of military operations research, promotes the exchange of information among developers and users of military operations research; and encourages the study of military operations research.

The society sponsors two open meetings annually, publishes a book series "Topics in Operations Research," and, among other activities, jointly publishes *PHALANX* and *MILITARY OPERATIONS RESEARCH* with the Military Operations Research Society.

President
Prof Patrick J. Driscoll
Department of Systems Engineering
United States Military Academy
patrick.driscoll@usma.edu

Vice President/President Elect
Dr. Greg Parlier
SAIC
4901-D Corporate Dr
Huntsville, AL 35805
greg.h.parlier@saic.com

Secretary/Treasurer
Alan Johnson

Awards Committee Chair and Former President
Dr. Keith N. Womer
Hearin Center for Enterprise Science
University of Mississippi
kwomer@bus.olemiss.edu

Council Member
COL Timothy E. Trainor, USA, Ph.D.
Department of Systems Engineering
United States Military Academy
tim.trainor@usma.edu

COL William K. Klimack
Department of Systems Engineering
United States Military Academy
william.klimack@usma.edu

Council Member
Lt Lee Stenson

George Mayernik
Raytheon Integrated Defense System

Dr. Niki Goerger
ERDC

Former President
Edward A. Pohl, Ph.D.
University of Arkansas
epohl@engr.uark.edu

Dr. Richard F. Deckro
Air Force Institute of Technology
richard.deckro@afit.af.mil

Former President
Philipp A. Djang
US Army Research Labs
djang@arl.army.mil

Dr. John P. Ballenger
Raytheon, Inc.
jp_ballenger@raytheon.com

Former President
Dr. Bruce W. Fowler
US Army Aviation Missile Command
fowler-bw@redstone.army.mil

Dr. Tom R. Gulledge
George Mason University
gulledge@gmu.edu

Executive Summaries

DEFINING EFFECTS FOR PROBABILISTIC MODELING

by Dr. Mark A. Gallagher, Gregory J. Ehlers, Wesley D. True, and Marc R. Warburton

With Effects Based Operations (EBO), the Department of Defense is emphasizing achieving effects, rather than accomplishing tasks. For the operations research community to fully engage in this EBO paradigm, the effects have to be quantifiable. As an initial step, the authors propose a rigorous, but versatile, definition of effects. Their definition provides a probability space for effects. As a sample application, they demonstrate a reliability approach to modeling probability of achieving an effect. Given a network of parallel or series independent events, the authors show the overall probability of success and its variability may be determined from the mean and variance of the individual activities. Their effects definition provides a measurable foundation to improve EBO predictions and assessments.

A COMPARISON OF MULTIVARIATE OUTLIER DETECTION METHODS FOR FINDING HYPERSPECTRAL ANOMALIES

by Timothy E. Smetek and Kenneth W. Bauer Jr.

Hyperspectral imagery is an emerging technology useful in locating unusual objects dispersed within some natural background such as a desert scene. Current anomaly detection methods commonly use non-robust statistical methods that may lead to inaccurate detection results. This research explores the use of different multivariate outlier detection methods for the anomaly detection problem. These methods encompass a large literature, which for unknown reasons, have been principally overlooked. In theory, these methods may be better suited than existing anomaly detection methods for finding anomalous objects in a hyperspectral image. This hypothesis is tested by applying a range of outlier detection methods to both simulated and real-world image data. Test results indicate that multivariate outlier detection can achieve superior detector performance relative to benchmark anomaly detection methods.

EFFICIENT EMPLOYMENT OF NON-REACTIVE SENSORS

by Moshe Kress, Roberto Szechtman, and Jason S. Jones

Advances in sensor technologies induce new operational scenarios that necessitate new algorithms for efficient sensor employment. The objective is to find a search pattern that maximizes the informational gain from a constrained set of sensing assets. In this paper we consider two types of non-reactive sensors, whose employment cannot be changed during the search mission, and develop optimal search patterns that maximize the information obtained from the sensor.

REACH-BASED ASSESSMENT OF POSITION

by Lt Col J. Todd Hamill, USAF, Dr. Richard F. Deckro, Dr. Robert F. Mills, and Dr. James W. Chrissis

As the interest in applying social network analysis (SNA) techniques to military problems has increased, so too has the realization that numerous theoretical and practical challenges must be overcome when dealing with social networks that cannot be easily observed. When analyzing clandestine networks, efficiently calculated measures that perform well despite limited information are of increasing interest, particularly to counterterrorism efforts. A new SNA measure, the reach-based assessment of position (RBAP), was specifically designed to serve as a 'screening' tool to identify individuals within such a network who may potentially serve important roles in achieving organizational objectives.

MODELLING THE RURAL INFANTRY BATTLE: THE EFFECTS OF LIVE COMBAT ON MILITARY SKILLS AND BEHAVIOUR DURING THE APPROACH PHASE

by L. R. Speight and D. Rowland

Many combat models are based on estimates of military performance and behaviour that have been obtained in peacetime conditions. However, historical analysis has shown that performance in actual combat typically strays far below the peacetime norm. A survey of the recorded accounts of junior officers and NCOs who have actively participated in dismounted infantry attacks shows an extremely wide range of behaviours. The heroic few have appeared to be essentially unaffected by the combat environment, exercising their military skills to all intents and purposes at their peacetime level. The remainder all seemed to suffer noticeable degradation, stretching at best from impaired use of their small arms to, at worst, zero effective participation in the battle. The present paper seeks to incorporate aspects of this live combat behaviour in a stylised account of an infantry assault. The approach started with a combat model that was carefully aligned to the outcomes of realistic instrumented trials with trained troops. Rules were then superimposed in order to link the behaviour of each individual attacker and defender, as revealed by historical analysis, to the prospect of battle and the characteristics of the opposing fire. The results of a modelling exercise using this enhanced model were encouraging, in that the outcomes were very much in accord with the outcomes of historical battles. This exercise also highlighted a number of topics that warranted further effort if our models are to be made more realistic in terms of genuine combat. The two most important of these were, firstly, the effect of each soldier's behaviour, especially that of heroes, on the military contributions of his immediate comrades in arms; and, secondly, the representation of the close quarter battle and its effect on the final outcome of any engagement.

Defining Effects for Probabilistic Modeling

Dr. Mark A. Gallagher

Headquarters Air Force Studies & Analyses, Assessments, and Lessons Learned
Mark.Gallagher@pentagon.af.mil

Gregory J. Ehlers

United States Strategic Command
Gregory.Ehlers@stratcom.mil

Wesley D. True

United States Strategic Command
Wesley.True@stratcom.mil

Marc R. Warburton

Science Applications International Corporation
Marc.Warburton@stratcom.mil

APPLICATION AREAS

Revolution in Military Affairs, Measures of Effectiveness, Plan Development, Effects-Based Operations

OR METHODOLOGIES

Reliability, Probability, Mathematical Modeling

ABSTRACT

The Department of Defense is implementing Effects-Based Operations (EBO). Critical EBO aspects include planning for effects and assessing achievement of effects. Both of these steps are enhanced with an analytical model that quantifies the probability of effect. Joint Publication 5-0 defines effect as "the physical or behavioral state of a system that results from an action, a set of actions, or another effect." Furthermore, that publication also defines Measure of Effectiveness (MOE) as "a criteria used to assess changes in system behavior, capability, or operational environment that is tied to measuring the attainment of an end state, achievement of an objective, or creation of an effect." In this article, we develop the MOE of Probability of Effect to assess the likelihood of achieving a specified effect given a proposed course of action.

Probability Theory requires a defined sample space, where each possible outcome is enumerated. In this article, we further refine the Joint Publication effect definition as an impact on a **single functional capability or behavior** with four specifications: 1) **scope** (affected entities or individuals), 2) **magnitude** (specifies the desired extent of the capability or behavior, which may constitute a decrease, maintenance, or increase), 3) **start time,** and 4) **minimum duration.** We contend this definition is universally applicable to the functions of any system and behaviors of any group. Furthermore, this definition lays a framework for both operational implementation and rigorous analysis.

For the limited situations where the time-dynamics of actions are not critical and system component reliabilities are independent, we present a reliability construct to predict (before the action) or assess (after the action based on limited observables) effect achievement. Gallagher and Whiteman (2004) calculate the overall mean and variance for serial systems, such as a kill chain, based on the means and variances of the individual components. Similarly, this article presents calculations for the overall mean and variance for components in parallel or combinations of series and parallel. We show that the beta distribution adequately represents the uncertainty of achieving an effect unless the mean is very close to zero or one. In addition, the equivalence of modeling reliability (system functioning) and unreliability (system not functioning) effects is shown.

EFFECTS DEFINED

The Department of Defense is implementing Effects-Based Operations (EBO). A common EBO approach is to consider Systems-of-Systems Analysis. These effects may occur to friendly, adversary, and neutral systems, such as the United States, Coalition Nations, Neutral Nations, Adversary Nations, Terrorists, and other entities. The latter category may include international government bodies (such as the United Nations and the World Trade Organization), Non-government Organizations (such as Amnesty International and the Red Cross), corporations, crime syndicates, religions, or other groups. EBO considers the dynamic impacts and interactions among these entities with the goal of achieving a desired end state for the entities. For example, military attacks have effects on the target, the attacker's resources, and may modify behaviors of third-party observers. Commander's intent should specify the desired end-state for friendly, adversary, and neutral systems and their behaviors. Explicitly defining effects allows for continual assessment of the current status and selection of actions that are integrated across the instruments of national power to progress toward the desired end-state. EBO planning and assessment across this vast breath requires a systematic approach.

Joint Publication 5-0 (2006) describes the EBO process. Planning begins with determining the desired end state with supporting objectives, which specify "clearly defined, decisive, attainable goals". The effects describe measurable system behavior in the operational environment that are the conditions for achieving objectives. Furthermore, an effect is a "physical and/or behavioral state of a system . . ." The system is composed of interacting or interdependent elements or nodes. These nodes may be individuals, places, or things. Links represent the behavioral, physical, or functional relationship between nodes. Tasks direct friendly actions against nodes to achieve desired effects. Our goal is to eval-

uate likelihood of potential tasks, or equivalently friendly actions, achieving desired effects and causing undesired effects.

The complexity of interactions between systems has resulted in stressing collaboration among experts, where subject matter experts become the primary means of planning and integration. Given the vast scope of aspects to consider, collaboration improves the options, however using subjective judgment alone is unlikely to develop an integrated and holistic set of actions. A systematic approach to a problem of this complexity needs to incorporate subjective assessments, but should not completely rely on subjective evaluations.

We used seven attributes in developing our effects definition. First, the terms should aid in communication among senior leaders and managers of the instruments of power. In the military, the system should contribute to clear understanding among the commander, planners, operators, and intelligence officers. Second, consistent with Joint Publication 5-0 direction, the impacts should be measurable. Even if not directly observable, planning predictions and outcome assessments should be stated in quantifiable terms with meaningful units. While behavior can be measured, psychological states or mental attitudes cannot and should not be included as effects. Third, the system needs to account for the uncertainty of prior estimates and after-action assessments. Estimating the extent of uncertainty for effects regarding human behavior is important. Fourth, effects should be developed in a manner that they may be combined toward higher effects, objectives and eventually the desired end state. Fifth, the effects should be defined in a manner that enables consideration of employing various means to achieve them. In other words, effects are described independent of actions that induce them so that various actions may be evaluated in terms of their likelihood in achieving a set of effects. Sixth, effects need to consider the behavior of individuals, not simply mechanical functions. Seventh, the dynamic nature of functions and behaviors should be addressed. We use these seven attributes in developing our proposed system.

The heart of an EBO planning is the definition of effects. Three major definitions have been published. The official DoD (2007) definition is, "A change to a condition, behavior, or degree of freedom." United States Joint Forces Command (2005) defines effects as "the physical, functional, or psychological outcome, event, or consequence that results from specific military or non-military actions." The *Commander's Handbook for Effects-Base Approach to Joint Operations* (Joint Warfare Center, 2006) and Joint Publication 5-0 use the effect definition of "the physical and/or behavioral state of a system that results from an action, a set of actions, or another effect." These definitions, without further specification, are not measurable; hence, ascertaining the probability of an effect, as a prediction before an action or as an assessment after an action, remains subjective and arbitrary. In addition, the ability to combine effects, as defined, in a systematic approach is unclear.

Rather than trying to define an effect, others have tried to *describe* effects. This approach leads to a lexicon with a myriad of verbs, like "degrade" and "deny.' Unfortunately, this approach is not comprehensive, which leads to an ever expanding dictionary of effect verbs. In addition, many of these verbs specify how the effect is achieved, rather than the functional impact. For example, "destroy" implies a kinetic attack, even if a computer network attack or electronic jamming could achieve the necessary functional effect. Historically, the analytical community has worked extensively to build a standardized taxonomy for kinetic damage levels, which provides planners with consistent, reproducible, and comparable evaluations that enables selection of the optimal means. Furthermore, these effect verbs lack the precision required for clear communication or rigorous evaluations. We contend effects should not be defined through the use of descriptor verbs.

In light of the seven desired attributes for an EBO system, we define an effect as an impact on a **single functional capability or behavior** with four specifications: 1) **scope** (the affected entities, which may be geographical such as a facility, city, region; organizational such as individual, group, or nation; or networks), 2) **magnitude** (specifies the desired extent of the capability or behavior whether it is decreased, maintained, or increased), 3) **start**

time, and 4) **minimum duration.** This refinement of the general definitions establishes an exhaustive and mutually exclusive outcome or sample space. While only one outcome is possible, either the effect will/did or will not/did not occur, our estimates need to account for the uncertainty of random influences prior to taking actions and limited observability after actions.

We contend this definition meets all seven of our proposed attributes. First, since we naturally speak of functions and behaviors, this definition fits within our normal vocabulary, which improves communication between commanders, planners, operators, and intelligence officers. Second, the single functional capability or behavior with a specified magnitude ensures quantifiable metric. Joint Publication 5-0 defines these metrics as measures of effectiveness (MOEs), "a criterion used to assess changes in system behavior, capability, or operational environment that is tied to measuring the attainment of an end state, achievement of an objective, or creation of an effect." Third, the uncertainty may be accounted in several ways including probabilities or range of probabilities. Fourth, these effects may be combined in at least three manners. The functional capabilities may be nested, such as the ability of a bridge to support traffic would include the ability to support heavy traffic. One function degrade may lead to a subsequent effect, such as inability to communicate may reduce ability to be cued to incoming threats. The effects may be related in time; for example, we may desire electric power interrupted during one time period and restored at a later time. Fifth, this definition imposes no constraint on the means to achieve the effect. The effect may even be achieved without action based on the adversary's incompetence or reliability failure. Sixth, the definition may include measurable human behavior. Seventh, the definition incorporates the time aspects so the dynamic nature of systems is addressed.

We contend this effect definition offers several advantages. First, it applies to any function or behavior that we may measure in any domain, so it is universally applicable. Second, the effect is not linked to any action so the definition poses no limitations on evaluating alternative actions to achieve it. Third, this definition clearly maps outcomes (events) to a sample space, which enables the application of probability theory. The Joint Technical Coordinating Group for Munitions Effectiveness (JTCG/ME), which publishes the Joint Munitions Effective Manual (JMEM) to support the weaponeering and combat assessment phases of the Joint Targeting Cycle, has adopted this definition. (JTCG/ME, 2006).

This effects approach enables wide consideration of potential approaches to achieve it. For example, rather than specifying "destroying a bridge," the effect statement should be eliminate ability of traffic to cross the bridge for a specified time period. In this example, this effect may be accomplished by a hazard sign, blocking action, computer opening a draw bridge, military occupation, or destruction. These alternative courses of action should be evaluated in terms of their likelihood of success and their impact on other desired effects. Do we want the bridge operating at a later time? Do we prefer a clandestine operation so the population is not enraged against our forces? Is their infrastructure, such as power or communication lines, along the bridge that we do or do not want to affect? Each of these other considerations may be stated as additional effects. The course of action that collectively best achieves the desired effects while minimizing the undesired effects should be selected. Consistent with Joint Publication 5-0, we could call this collection of related effects an objective.

Effect, as defined, is limited to a single, measurable, functional capability or behavior. To construct a comprehensive assessment, multiple effect evaluations are required. A single action may precipitate multiple effects, and a single effect may be precipitated by multiple actions. Additionally, it may be desired that an action have both a particular resultant effect and not have negative consequences or undesired effects. For example, concerns about detection and attribution may be represented as avoiding associated effects of behavior responses. Another example is selecting the best option to achieve two related effects, such as electrical power outage for a time and having it restored in a later time period. In these cases, the associated single dimensional measurements or MOEs for each effect may be combined to form a state vector that varies over

time. Actions should be considered such that the state transitions toward the desired end-state. This concept of an outcome state is consistent with EBO literature. In addition, the concept of a time-varying state vector facilitates direct application of standard engineering approaches including Box-Jenkins forecasting, stochastic estimation and control theory, or simulation.

A variety of analytical techniques may be applied to estimate the probability of effect in a manner that enables a systematic construction to determine the likelihood of achieving the desired end-state. The remainder of this article describes one simple approach that applies when the target functions and vulnerabilities are known and time-dynamics responses are not critical. We contend this is applicable, at a minimum, to many fixed targets. We envision future analysts and planners will employ a variety of techniques, beyond just this one that we present, to estimate the probability of effect.

The description of this one approach is divided into three main sections. First, we discuss system analysis approaches to depict relationships between potential actions and effect achievements. Second, we present how, based upon a reliability structure, a probability distribution function for an effect may be constructed. This stochastic approach enables accounting for intelligence uncertainties in addition to operational variations. Third, we demonstrate our approach. This approach enables us to estimate a probability of effect for either potential future actions with uncertain results or for assessing mission success based on limited observable indicators.

SYSTEM ANALYSIS

The initial step is specification of the desired effect. In our approach, we begin at the impact to facilities or individuals, which can be incorporated into evaluation of higher states. As a preliminary step, we advocate for constructing a list of general facility functions for each of the installation category codes (DIA, 2001) used in the intelligence database. For example, the installations categorized as a command bunker, 86400, would have two measurable primary functions: plan actions and communicate. The planned actions function may be measured by the MOE of equivalent number of simple directions per time period. The MOE for the communication function may be the rate of data received or transmitted. Even this simple specification significantly assists in the implementation of EBO. Rather than describing the damage required, warfighters may begin to plan for consistent functional impacts on installations.

The next step is to link actions on nodes or components to the function specified in the effect. Kapur and Lamberson (1977) state "The reliability of a system is the probability that, when operating under stated environmental conditions, the system will perform its intended function adequately for a specified interval of time." This definition has only the minor distinction from our effects definition in that rather than the "perform its intended function adequately" the effect definition specifies a "the desired extent of the functional capacity." As a result of this similarity, much of the wealth of reliability literature and techniques may be applied to our task. When aging failures and repair times are greater than the duration of interest, the reliability problem may simplify from a dynamics to a static problem. Similarly, a planner's estimate of a kill chain success for military strikes on fixed installations is usually determined without considering time aspects of the attacking forces or the target. The timing for the needed effect clearly drives when the target is attacked in the campaign.

When performing this functional analysis, the planner must identify vulnerable nodes that support the functional effect. These vulnerable points may be entire nodes identified in center-of-gravity analysis or crucial elements of a particular installation. For example, a command bunker's power supply may be critical for its function to create, distribute, and communicate commands. For effects that may be achieved through military strikes, the actions including both weapons and tactics being considered determine the functional vulnerabilities; hence, the level of analytical detail required. Clearly, assessing the impact of a nuclear detonation against a facility rarely requires subsystem descriptions. The application of conventional explosives requires detailing the major components at a minimum. Assessing a computer

network attack necessitates very detailed system analysis. Therefore, while the functional effect may be the same, different actions may require varied levels of analysis to determine the vulnerable objects and their impacts.

We may evaluate the function of a system based on the component reliabilities. We must be careful to consistently model the effect as either the reliability or the unreliability of the system. Actions against a system have success rates (which may be referred to as reliabilities) and affect the adversary's system component reliabilities. Friendly military actions, however, are not part of that adversary system's normal construct. In fact, friendly military actions are often counter to the adversary's system functioning. For example, an attack with a high probability of arrival and probability of damage should result in low adversary system reliability. Modeling in terms of system reliability, R, requires evaluating the failure of military action to disable the adversary system. Conversely, modeling a failure or unreliability effect of a system, E, entails evaluating the success of military action to disable the system. Both of these approaches may incorporate the system inherent reliability/unreliability if it is significant. Let the effect of the ith component be the complement of its reliability; $E_i = 1 - R_i = \bar{R}_i$. For a system of n component, in series,

$$R = \bigcap_{i=1}^{i=n} R_i = 1 - \bigcup_{i=1}^{i=n} \bar{R}_i = 1 - \bigcup_{i=1}^{i=n} E_i = 1 - E$$

and, in parallel,

$$R = \bigcup_{i=1}^{i=n} R_i = 1 - \bigcap_{i=1}^{i=n} E_i = 1 - E.$$

The complementary nature of the reliability and failure effect may be verified also for systems that contain combinations of series and parallel structures. Ross (2000) states that any system can be represented as a series arrange of parallel minimal path structures or a parallel arrangement of serial minimal cuts sets. This dualism is similar to modeling system reliability or mission effectiveness (a failure effect). Realize that the effects definition may encompass functional and behavioral impacts much broader than system failures.

The DoD Architecture Framework (2004) lays out a systematic approach for developing system relationships that may also be applied for EBO planning. Their framework states "system functions are executed by automated systems, while operational activities describe business operations that may be conducted by humans, automated systems, or both." Effects, as we define them, are synonymous with the operational activities. The Operational Activity to Systems Function Traceability Matrix, System View (SV)-5, describes the necessary relationships. We contend that the necessary relationships between potential targeted components and primary functions may be developed through architecture or reliability techniques.

Admittedly, limited or incomplete intelligence data on an adversary's systems may be addressed in this approach in two ways. First, the node may be characterized as not having a particular vulnerability due to the uncertainty. For example, if we know a facility has a pump but are not certain of its location within a building, we may assess the pump to have perfect reliability. Second, we may address the uncertainty through reduced reliability. In the pump example, if the building size limits our chance of successfully terminating its functioning, we may reduce the probability of mission success.

As an example, we applied systems engineering techniques to determine a functional decomposition to vulnerable objects of a hypothetical underground command and control facility. The results are shown in Figure 1. Many systems, potentially even actions that encourage specific behaviors and their responses, may be modeled as construction of series (intersections) and parallel (unions) of components.

PROBABILITY DISTRIBUTION FUNCTION (PDF) FOR PROBABILITY OF EFFECT (PE)

We account for the uncertainty based on limited intelligence or operational variability by incorporating the mean and variance of each step in the action and system component. The

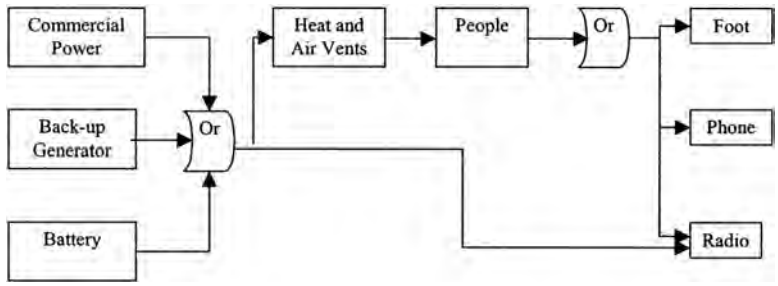

Figure 1. Functional Diagram of Vulnerable Objectives of a Command and Control Bunker.

variances may be based on testing, a subjective assessment, or a conservative upper bound. With this additional information, our approach provides decision makers with an estimate of success and an indication of the certainty of that estimate.

We applied a Bayesian approach to the component reliabilities, action successes, and overall PE for three main reasons. First, treating the estimated PE as a random variable has more intuitive appeal; decision-makers seem to grasp the Bayesian direct representation of the uncertainty of the estimate easier than the classical approach to relate the variation to the fixed, but unknown, "true" value. Second, the Bayesian approach readily incorporates subjective assessments. Third, the Bayesian approach incorporates the uncertainty of individual values into the overall reliability.

We model the system functioning in a reliability construct, which consist of series or parallel connections to accomplish the function of the specified effect. Equivalently an unreliability effect may be modeled by interchanging the series ("and") and the parallel ("or") constructs and replacing component reliability distributions with their complements. We estimate the distribution parameters from the means and variances of the individual component reliabilities and their variances. Finally, we demonstrate the overall distribution provides an adequate model for probability of effect (PE).

We assume the overall system reliability or effect follows a beta distribution. We selected the beta distribution for five reasons: 1) the beta is the distribution of success rate for binomial outcomes; 2) the beta is the distribution for Bernoulli outcomes with partial failures; 3) Gallagher and Whiteman (2004) show theoretically the asymptotic product of a series of probabilities follows a lognormal, but with limited number of terms in the product, a beta distribution fits better; 4) the beta distribution is one of the distributions with no tails that exceed the probability limits of zero and one; 5) the beta distribution is a conjugate prior; hence beta distributions beget beta distributions. (See Gallagher, Weir, and True (1997) for one development.)

We applied our proposed Bayesian approach to determine a beta distribution for PE, as depicted in Figure 2. A narrower probability density function (PDF) or, equivalently, a steeper cumulative distribution function (CDF) indicates more certainty in the assessment. In addition, we may calculate probability intervals for the estimate based on the statistical distribution. For example, the symmetric 90% probability interval for the PE assessment in Figure 2 is (0.78, 0.91). This article presents and tests our Bayesian approach for developing a statistical distribution for probability of effect, either in planning or assessing an action.

A system functioning may be modeled as a reliability network. Let Y_i be defined as either the reliability, R_i, or failure effect (unreliability), E_i, of the ith subsystem or component. We treat Y_i as a random variable in this Bayesian approach.

Gallagher and Whiteman (2004) present the mean and variance of the intersection of n independent random variables, $\cap_{i=1}^{i=n} Y_i$, based on the individual random variables' means and variances.

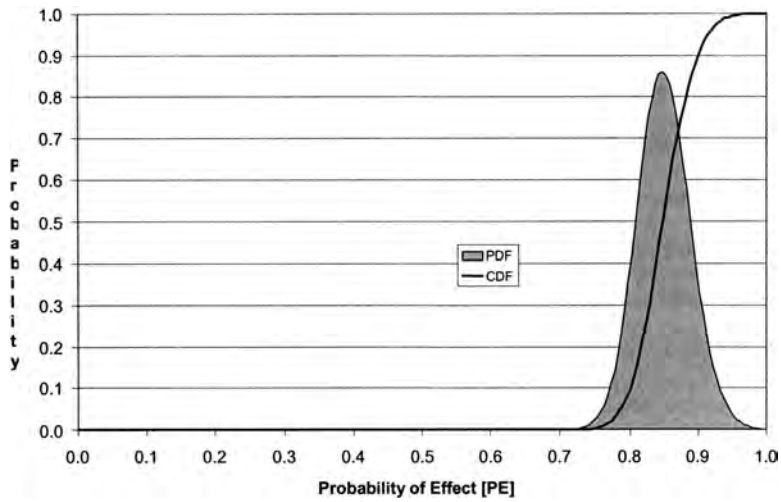

Figure 2. Example Statistical Distribution of Probability of Effect

$$\mu^{i=n}_{\bigcap_{i=1} Y_i} = E\left[\prod_{i=1}^{i=n} Y_i\right] = \prod_{i=1}^{i=n} E[Y_i]. \quad (1)$$

$$\sigma^2_{\bigcap_{i=1}^{i=n} Y_i} = \mathrm{Var}\left[\prod_{i=1}^{i=n} (Y_i)\right] = \prod_{i=1}^{i=n} (\mathrm{Var}[Y_i] + E[Y_i]^2) - \mu^2_{\bigcap_{i=1}^{i=n} Y_i}. \quad (2)$$

Equations (1) and (2) are appropriate for reliabilities of independent components arranged in series or the complement of failure effects of independent components arranged in parallel. Similarly, the mean and variance of the union of n independent random variables, $\bigcup_{i=n}^{i=1} Y_i$, based on the individual random variables' means and variances are

$$\mu^{i=n}_{\bigcup_{i=1} Y_i} = E\left[1 - \prod_{i=1}^{i=n} (1 - Y_i)\right] = 1 - E\left[\prod_{i=1}^{i=n} (1 - Y_i)\right], \text{ and} \quad (3)$$

$$\sigma^2_{\bigcup_{i=1}^{i=n} Y_i} = \mathrm{Var}\left[1 - \prod_{i=1}^{i=n} (1 - Y_i)\right] = \mathrm{Var}\left[\prod_{i=1}^{i=n} (1 - Y_i)\right]$$

$$= E\left[\left(\prod_{i=1}^{i=n} (1 - Y_i)\right)^2\right] - E\left[\prod_{i=1}^{i=n} (1 - Y_i)\right]^2$$

$$= E\left[\prod_{i=1}^{i=n} (1 - Y_i)^2\right] - \left(1 - \mu^{i=n}_{\bigcup_{i=1} Y_i}\right)^2$$

$$= \prod_{i=1}^{i=n} E[(1 - Y_i)^2] - \left(1 - \mu^{i=n}_{\bigcup_{i=1} Y_i}\right)^2$$

$$= \prod_{i=1}^{i=n} (\mathrm{Var}[1 - Y_i] + E[1 - Y_i]^2) - \left(1 - \mu^{i=n}_{\bigcup_{i=1} Y_i}\right)^2$$

$$= \prod_{i=1}^{i=n} (\mathrm{Var}[Y_i] + (1 - E[Y_i])^2) - \left(1 - \mu^{i=n}_{\bigcup_{i=1} Y_i}\right)^2. \quad (4)$$

Equations (3) and (4) model independent component reliabilities arranged in a parallel structure or the complement of independent failure effects arranged in series.

Complex structures with combinations of series and parallel arrangements may also be

assessed. The distribution for each node is determined based on the prior nodes' distributions with repeated applications, as appropriate, of (1) and (2) for series relationships and (3) and (4) for parallel relationships.

If only an estimate $\hat{Y}_i$ is known, with no basis for determining the associate variance, Gallagher and Whiteman (2004) suggest assuming that variance is based on the minimum number of binomial testing in which exactly one failure or one success occurs. This assumption yields a geometric series and an upper bound for the variance of

$$\sigma^2_{Y_i} = \begin{cases} \hat{Y}_i(1-\hat{Y}_i)^2 & \text{for } \hat{Y}_i \geq 0.5 \\ \hat{Y}_i^2(1-\hat{Y}_i) & \text{for } \hat{Y}_i < 0.5 \end{cases}. \quad (5)$$

We approximate the overall reliability with a beta distribution,

$$f(x) = \begin{cases} \dfrac{x^{\alpha-1}(1-x)^{\beta-1}}{\displaystyle\int_{=0}^{=1} t^{\alpha-1}(1-t)^{\beta-1}dt} & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}.$$

The mean and variance of the beta distribution are given by

$$\mu = \frac{\alpha}{\alpha+\beta} \text{ and} \quad (6)$$

$$\sigma^2 = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}. \quad (7)$$

One may view the formulation as either a probability of reliability or a probability of unreliability, equivalently a failure effect. Since unreliability is simply one minus the reliability, the parameters of the beta distribution simply reverse roles. The reliability and unreliability distributions means are the probabilistic complements while the distribution variances are equal.

Algebraic manipulation of the beta's mean and variance formulas, given in (6) and (7), yields parameters for the beta distribution when mean and variance are known as

$$\alpha = \mu\left[\frac{\mu(1-\mu)}{\sigma^2} - 1\right], \text{ and} \quad (8)$$

$$\beta = (1-\mu)\left[\frac{\mu(1-\mu)}{\sigma^2} - 1\right]. \quad (9)$$

Analysts may determine subjectively the beta parameters by specifying the mean and setting the equivalent number of tests equal to the sum of α and β. We can determine α based on (6) and then β as the number of tests minus α. Intelligence uncertainty may be accounted for by a high variance or modeling some component as invulnerable.

DEMONSTRATION

We present some numerical examples to clarify this approach. Consider a system with two components operating in series to accomplish a function that we desire to affect. The reliability of the first component is 0.99 and the reliability of the second component is 0.98 to accomplish their contribution to the system's function. Based on (5), the respective variances are 0.00010 and 0.000392. From (1) and (2), we calculate the system reliability mean as 0.97 and variance as 0.0022. The beta distribution that matches these moments using (8) and (9) has parameters $\alpha = 1.82$ and $\beta = 59.35$. Equivalently, we could model the failure effect (unreliability) with the complements of each term. The result is a beta distribution with parameters $\alpha = 59.35$ and $\beta = 1.82$ with a mean of 0.03 and variance of 0.0022. Hence, the unreliability distribution is the complement of the reliability distribution in that the beta parameters switch, which results in the mean equal to the complement of the system reliability and the same variance. Thus, the probability of effect is 3% for not taking any action due to the system's own inherent reliability.

Let us consider actions that impact the function of this system. We modeled probability distributions for three components. These components may represent phases of the kill chain and components of the targeted system, but all distributions must consistently represent either reliabilities of system operation or system failures. For conditional series arrangement, we could have a mission with a probability of arrival, a probability of damage given

arrival, and unreliability of the target given damage. The equivalent parallel construct would have unreliability of mission arrival, unreliability of mission strike given arrival, and target reliability given strike.

We tested this approach with a Monte Carlo simulation. For each system case, we generated random variables in two nested loops. In an outer loop of 2,000 replications, we generate success rates from the beta distribution for each component in the combined action and affected systems. In the inner loop of 2,000 replications, we generate Bernoulli values for each of the components and determine whether success is achieved. We compare the empirical averages over the inner loop iterations with the theoretical distribution statistics from applying our approach. We used three tests to compare the empirical distribution of the 2,000 resulting means with theoretical parameters based on (1) through (8), as appropriate. We evaluated the difference in the modeled and generated means and variances. Further, we calculated the Kolmogorov-Smirnov (KS) and the Anderson-Darling (AD) goodness-of-fit statistic for the beta distribution with parameters based on the theoretical mean and variance. Law and Kelton (1991) present 95% critical values for parameters not estimated from the data of 1.358 for an adjusted KS and 2.492 for AD. For component distributions, we test a range of beta distributions with parameters shown in Table 1 and depicted in Figure 3.

Gallagher and Whiteman (2004) demonstrated that (1) and (2) represent the mean and variance of a system in series and that a beta distributions fits unless significant probabilities are near either zero or one. Table 2 shows results for three components combined in parallel. Since the order of components does not matter, the 20 unique combinations of the four distribution parameter sets are depicted. The columns labeled μ and σ^2 are the theoretical parameters while the columns "Mean" and "Variance" are the empirical statistics. The average and maximum absolute differences in mean values are 0.002 and 0.004, respectively, and corresponding variances differences are zero to three decimals and 0.002. Hence, this methodology accurately represents the first two moments.

The shaded KS and AD statistics are greater than the 95% critical value. The beta distribution fails to reject only 50% and 44%. The overall system has a mean very close to one so often the simulation of 2,000 replications has no failures. Since the beta distribution must have zero probability at both zero and one, it has difficulty fitting some of the distributions that have significant likelihood at these limits. The correlation between the theoretical mean and the KS and AD statistics of 50% and 47%, respectfully, is evidence of this modeling challenge. As the mean approaches either limit, the beta distribution fit is less adequate. In Table 2, the goodness-of-fit statistics often exceed the critical value when the theoretical mean is greater than 0.87. Since KS equally weighs each data point, whereas the AD heavily weighs data at the limits, the AD consistently has lower non-rejection rates for these trials.

While (1) through (4) combine either parallel or series independent components, these equations may be applied repeatedly to evaluate systems with combinations of series and parallel structures. Starting from the initial components, the inputs to each component are combined using the appropriate equations until the overall system parameters are determined. If desired, the corresponding beta distribution may be determined with (5) and (8).

Table 1. Beta Parameters and Statistics for Tested Distributions

FORM	ALPHA	BETA	MEAN	VARIANCE
Uniform	1	1	0.5	0.083
Symmetric Mode	3	3	0.5	0.036
Left Skewed	5	2	0.714	0.026
Right Skewed	2	5	0.286	0.026

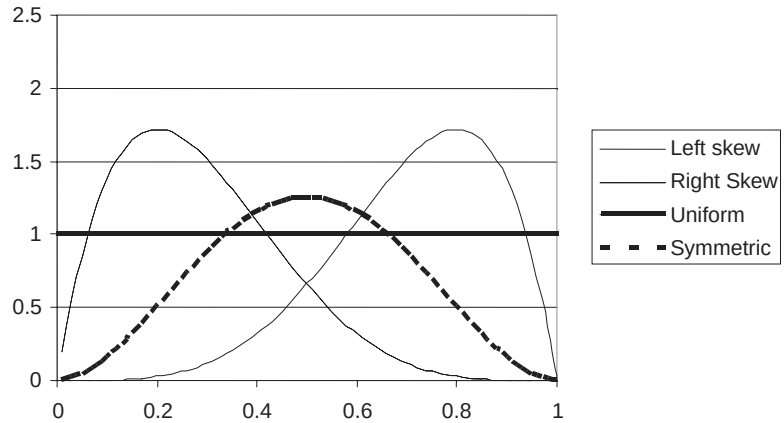

Figure 3. Tested Component Beta Distributions

Table 2. Results for System with 3 Parallel Components

Comp 1	Comp 2	Comp 3	μ	Mean	σ^2	Var	KS	AD
Left Skew	Left Skew	Left Skew	0.977	0.977	0.001	0.001	3.122	15.245
Left Skew	Left Skew	Uniform	0.959	0.959	0.002	0.002	1.497	3.200
Left Skew	Left Skew	Symmetric	0.959	0.958	0.002	0.002	2.241	6.595
Left Skew	Left Skew	Right Skew	0.942	0.942	0.003	0.003	1.050	1.668
Uniform	Uniform	Left Skew	0.929	0.928	0.007	0.007	1.698	4.153
Uniform	Symmetric	Left Skew	0.929	0.930	0.005	0.005	1.447	3.509
Symmetric	Symmetric	Left Skew	0.929	0.928	0.004	0.004	1.581	4.743
Uniform	Left Skew	Right Skew	0.898	0.898	0.009	0.009	1.017	1.645
Symmetric	Left Skew	Right Skew	0.898	0.894	0.006	0.006	1.175	3.037
Uniform	Uniform	Uniform	0.875	0.877	0.021	0.020	2.051	8.307
Uniform	Uniform	Symmetric	0.875	0.871	0.016	0.018	0.981	1.497
Uniform	Symmetric	Symmetric	0.875	0.874	0.012	0.012	0.761	0.628
Symmetric	Symmetric	Symmetric	0.875	0.874	0.008	0.008	0.807	1.162
Left Skew	Right Skew	Right Skew	0.854	0.852	0.009	0.010	0.916	1.210
Uniform	Uniform	Right Skew	0.821	0.825	0.028	0.027	1.240	1.283
Uniform	Symmetric	Right Skew	0.821	0.821	0.019	0.019	0.787	0.907
Symmetric	Symmetric	Right Skew	0.821	0.821	0.012	0.012	0.630	0.359
Uniform	Right Skew	Right Skew	0.745	0.749	0.031	0.030	1.386	3.935
Right Skew	Right Skew	Symmetric	0.745	0.746	0.017	0.017	0.924	0.717
Right Skew	Right Skew	Right Skew	0.636	0.639	0.021	0.021	0.971	1.131

Table 3 presents the results for a three-component system. The first two components, R_1 and R_2, are in parallel followed by a third component, R_3, in series. The reliability function is $R = (1 - (1 - R_1)(1 - R_2))R_3$. We determined the parameters for the combination of R_1 and R_2 with (3) and (4). Then, we apply (1) and (2) to obtain the series parameters of R_3 with the combined results of R_1/R_2. Again, we see the method determines the mean and variance with great accuracy. For these samples, the average errors are extremely small; the maximum absolute difference in means is 0.009 and in variance is only 0.002. The goodness-of-fit test statistics for the beta distribution surpassed the 95% critical values with overall non-rejection rates of 87.5% for KS and 82.5% for AD. In these applications, the theoretical means, which range from 0.14 to 0.66, are not near zero or one and the beta distribution fits very well.

Table 3. Results for Three Components (Two in Parallel Followed by One in Series)

Comp 1	Comp 2	Comp 3	μ	Mean	σ^2	Var	KS	AD
Left Skew	Left Skew	Left Skew	0.656	0.653	0.024	0.024	0.784	0.571
Left Skew	Symmetric	Left Skew	0.612	0.606	0.024	0.025	0.888	1.489
Left Skew	Uniform	Left Skew	0.612	0.616	0.027	0.029	1.511	2.455
Left Skew	Right Skew	Left Skew	0.569	0.564	0.025	0.024	1.227	1.415
Symmetric	Symmetric	Left Skew	0.536	0.541	0.025	0.026	1.674	2.860
Symmetric	Uniform	Left Skew	0.536	0.533	0.032	0.032	0.721	0.647
Uniform	Uniform	Left Skew	0.536	0.527	0.040	0.040	1.034	2.114
Right Skew	Symmetric	Left Skew	0.459	0.460	0.024	0.025	0.634	0.574
Left Skew	Left Skew	Symmetric	0.459	0.459	0.031	0.031	0.765	0.407
Right Skew	Uniform	Left Skew	0.459	0.454	0.038	0.037	1.003	1.167
Left Skew	Left Skew	Uniform	0.459	0.464	0.072	0.072	1.289	2.287
Left Skew	Symmetric	Symmetric	0.429	0.434	0.029	0.029	0.971	1.176
Left Skew	Uniform	Symmetric	0.429	0.427	0.031	0.030	0.735	0.365
Left Skew	Symmetric	Uniform	0.429	0.428	0.065	0.065	0.989	3.111
Left Skew	Uniform	Uniform	0.429	0.426	0.066	0.067	0.805	1.539
Left Skew	Right Skew	Symmetric	0.398	0.407	0.027	0.027	1.186	3.253
Left Skew	Right Skew	Uniform	0.398	0.406	0.058	0.056	1.697	6.118
Symmetric	Symmetric	Symmetric	0.375	0.376	0.026	0.025	0.589	0.438
Symmetric	Uniform	Symmetric	0.375	0.375	0.029	0.030	0.710	0.694
Uniform	Uniform	Symmetric	0.375	0.375	0.034	0.034	0.706	0.583
Symmetric	Symmetric	Uniform	0.375	0.377	0.053	0.054	1.224	3.025
Symmetric	Uniform	Uniform	0.375	0.382	0.058	0.058	1.335	2.076
Uniform	Uniform	Uniform	0.375	0.383	0.063	0.063	1.101	1.458
Right Skew	Right Skew	Left Skew	0.350	0.353	0.020	0.021	0.833	0.739
Right Skew	Symmetric	Symmetric	0.321	0.321	0.022	0.023	0.697	0.471
Right Skew	Uniform	Symmetric	0.321	0.322	0.029	0.029	0.756	0.802
Right Skew	Symmetric	Uniform	0.321	0.323	0.043	0.042	1.144	2.338
Right Skew	Uniform	Uniform	0.321	0.320	0.051	0.050	0.590	0.529
Left Skew	Left Skew	Right Skew	0.262	0.263	0.022	0.022	0.578	0.262
Right Skew	Right Skew	Symmetric	0.245	0.245	0.016	0.015	0.686	0.790
Right Skew	Symmetric	Right Skew	0.245	0.245	0.020	0.020	1.088	0.726
Right Skew	Uniform	Right Skew	0.245	0.242	0.020	0.020	1.512	2.047
Right Skew	Right Skew	Uniform	0.245	0.246	0.029	0.028	1.120	2.771
Left Skew	Right Skew	Right Skew	0.227	0.229	0.018	0.019	0.643	0.413
Symmetric	Symmetric	Right Skew	0.214	0.212	0.016	0.015	0.877	1.421
Symmetric	Uniform	Right Skew	0.214	0.214	0.018	0.018	1.001	0.913
Uniform	Uniform	Right Skew	0.214	0.220	0.020	0.021	1.145	1.646
Right Skew	Symmetric	Right Skew	0.184	0.190	0.013	0.014	1.273	2.425
Right Skew	Uniform	Right Skew	0.184	0.180	0.016	0.016	1.412	2.569
Right Skew	Right Skew	Right Skew	0.140	0.144	0.009	0.009	1.199	1.901

Table 4 presents results from another three-component system. In this case, components 1 and 2 are combined in series. The composite R_1/R_2 is combined with R_3 in parallel. The reliability function is $R = (1 - (1 - R_1R_2)(1 - R_3)$. The errors between the empirical and theoretical moments are always very small. These samples are below the 95% critical values at a rate of 77.5% for the KS test and 65.0% for the AD test. We tracked the number of times that the each case generated a sample of all pass or all failures. Although rare, these counts have correlation coefficients of 43% with KS and 99% with AD.

Table 4. Results for Three Components (Parallel of Two in Series and One)

Comp 1	Comp 2	Comp 3	μ	Mean	σ^2	Var	KS	AD
Right Skew	Right Skew	Right Skew	0.344	0.347	0.024	0.024	1.061	1.143
Right Skew	Symmetric	Right Skew	0.388	0.386	0.024	0.024	0.588	0.308
Right Skew	Uniform	Right Skew	0.388	0.388	0.027	0.027	0.593	0.222
Right Skew	Left Skew	Right Skew	0.431	0.436	0.025	0.024	1.263	1.698
Symmetric	Symmetric	Right Skew	0.464	0.466	0.025	0.023	1.020	1.843
Symmetric	Uniform	Right Skew	0.464	0.461	0.032	0.031	0.782	0.810
Uniform	Uniform	Right Skew	0.464	0.469	0.040	0.040	0.766	1.055
Left Skew	Symmetric	Right Skew	0.541	0.540	0.024	0.023	0.542	0.427
Right Skew	Right Skew	Symmetric	0.541	0.540	0.031	0.030	0.648	0.593
Left Skew	Uniform	Right Skew	0.541	0.534	0.038	0.036	1.328	2.178
Right Skew	Right Skew	Uniform	0.541	0.534	0.072	0.071	1.013	32.461
Right Skew	Symmetric	Symmetric	0.571	0.574	0.029	0.028	0.834	1.045
Right Skew	Uniform	Symmetric	0.571	0.570	0.031	0.031	0.913	0.598
Right Skew	Symmetric	Uniform	0.571	0.574	0.065	0.066	1.094	3.767
Right Skew	Uniform	Uniform	0.571	0.579	0.066	0.068	1.480	17.889
Right Skew	Left Skew	Symmetric	0.602	0.601	0.027	0.026	0.923	0.804
Right Skew	Left Skew	Uniform	0.602	0.600	0.058	0.059	1.372	33.694
Symmetric	Symmetric	Symmetric	0.625	0.624	0.026	0.026	0.711	0.465
Symmetric	Uniform	Symmetric	0.625	0.629	0.029	0.029	1.076	1.442
Uniform	Uniform	Symmetric	0.625	0.627	0.034	0.033	0.670	0.713
Symmetric	Symmetric	Uniform	0.625	0.621	0.053	0.052	1.750	35.440
Symmetric	Uniform	Uniform	0.625	0.620	0.058	0.058	1.058	46.570
Uniform	Uniform	Uniform	0.625	0.629	0.063	0.065	1.288	46.892
Left Skew	Left Skew	Right Skew	0.650	0.653	0.020	0.021	1.050	0.867
Left Skew	Symmetric	Symmetric	0.679	0.681	0.022	0.023	1.012	0.911
Left Skew	Uniform	Symmetric	0.679	0.676	0.029	0.031	0.587	0.612
Left Skew	Symmetric	Uniform	0.679	0.674	0.043	0.044	1.822	37.303
Left Skew	Uniform	Uniform	0.679	0.679	0.051	0.051	0.419	45.705
Right Skew	Right Skew	Left Skew	0.738	0.740	0.022	0.022	0.819	0.825
Left Skew	Left Skew	Symmetric	0.755	0.750	0.016	0.017	1.291	2.186
Right Skew	Symmetric	Left Skew	0.755	0.759	0.020	0.019	0.737	0.908
Right Skew	Uniform	Left Skew	0.755	0.758	0.020	0.020	0.939	15.998
Left Skew	Left Skew	Uniform	0.755	0.760	0.029	0.029	1.556	49.838
Right Skew	Left Skew	Left Skew	0.773	0.772	0.018	0.019	0.992	0.940
Symmetric	Symmetric	Left Skew	0.786	0.789	0.016	0.016	0.962	1.003
Symmetric	Uniform	Left Skew	0.786	0.786	0.018	0.019	1.100	1.247
Uniform	Uniform	Left Skew	0.786	0.783	0.020	0.021	0.594	0.503
Left Skew	Symmetric	Left Skew	0.816	0.816	0.013	0.014	0.393	0.224
Left Skew	Uniform	Left Skew	0.816	0.815	0.016	0.016	1.104	1.953
Left Skew	Left Skew	Left Skew	0.860	0.859	0.009	0.009	0.709	0.842

Again, significant probabilities at the distribution limits of zero and one cannot be modeled well with the beta distribution.

We mentioned earlier, a reliability structure can be related to an equivalent failure effect (unreliability) construct where the effect, $E = 1 - R$. For a failure-effect construct, all the components in parallel reliability structures must fail, so parallel reliability structures relate to failure-effect (unreliability) series. Similarly, any component in a reliability series failing equates to a failure effect so reliability series relate to failure-effect parallel structures. In addition, the complement for each component distribution applies. For example, the system used in Table

3, viewed as a reliability system, directly relates to the system in Table 4, viewed as a failure effect.

$$\begin{aligned} R &= (1 - (1 - R_1)(1 - R_2))R_3 \\ &= (1 - E_1E_2)(1 - E_3) \\ &= 1 - E \end{aligned}$$

The cases arranged by theoretical means, increasing in Table 3 and decreasing in Table 4. In each corresponding case, the distributions are the complement (right skew versus left skewed), the means are the complements, and the variances of the reliability/effect are the same. Undem (2000) describes how we gain a great deal of insight when we examine our task based on either reliability or effect (survivability and targeting in his vernacular).

So far we have assumed independent components. In actuality, we may encounter situations in which a single action affects more than one component. Ross (2000) shows that the actual reliability may be calculated by adjusting for the correlation coefficients of all combinations (pairwise, three-wise, etc.) of component reliabilities. We can bound the correlated probability, one bound is the independent case, as developed in this article, and the other bound is perfectly correlated case where the construct is modified so that correlated components are replaced with one composite component.

SUMMARY

We define effect so that the outcomes map to a sample space and probability theory may be applied. This enables accounting for uncertainty while considering potential actions or evaluating results based on limited observables. We develop a statistical distribution for the probability of effect for systems functioning or behavior that can be described as a combination of series and parallel constructs. The overall estimate's mean and variance may be estimated from the individual factors' means and variances under the assumption of independence. A conservative variance estimate, depicting the most possible uncertainty, may be obtained from individual means. We demonstrate that our approach matches the overall mean and variance of the series, parallel, or combined systems. Except when the mean approaches the limits of zero or one, the beta distribution is very reasonable distribution of the uncertainty in the overall outcome.

Disclaimer

The views expressed in this article are those of the authors and do not reflect the official policy or position of the United States Air Force, Department of Defense, or the US Government.

REFERENCES

Defense Intelligence Agency Manual (DIAM 65-3-1), *Standard Coding Systems Functional Classification Handbook,* Bolling Air Force Base, Maryland, April, 2001.

Department of Defense (DoD), Dictionary of Military Terms, Web Site http://www.dtic.mil/doctrine/jel/doddict/data/e/01823.html, accessed 20 August, 2007.

DoD Architecture Framework Working Group, DoD Architicture Framework, Version 1.0, Volume II: Product Descriptions, 9 February 2004.

Gallagher, Mark A., and Phillip (Bud) Whiteman, "Probability Distribution Function for Damage Expectancy," *Military Operations Research,* Vol. 9, No. 3, 2004, pp. 5–15.

Gallagher, Mark A., Jeffery D. Weir, and Wesley D. True, "Relating Weapon System Test Sizes to Warfighting Capability," *Military Operations Research,* Vol. 3, No. 3, 1997, pp. 5–12.

Joint Publication 5-0, *Joint Operations Planning,* Joint Chiefs of Staff, Pentagon, Washington, District of Columbia, December 26, 2006.

Joint Technical Coordinating Group for Munitions Effectiveness (JTCG/ME), Joint Service Target Data Standardization Group, *Guidance for Generating, Documenting Reviewing, and Approving JCTG/ME Target Description and Vulnerability Data,* Aberdeen Proving Ground, MD, Draft, September, 2006.

Joint Warfighting Center, Joint Concept Development and Experimentation

Directorate, Standing Joint Force Headquarters, *Commander's Handbook for an Effects-Based Approach to Joint Operations,* Suffolk, Virginia, February 24, 2006.

Kapur, K. C. and L. R. Lamberson, *Reliability of Engineering Design,* John Wiley & Sons, New York, New York, 1977.

Law, Averill M. and W. David Kelton, *Simulation Modeling & Analysis,* 2nd Edition, McGraw-Hill, Inc., New York, New York, 1991.

Ross, Sheldon M., *Introduction to Probability Models,* 7th Edition, Academic Press, San Diego, California, 2000.

Undem, Halvor A. "A Random Variable Approach To Nuclear Targeting And Survivability," *Military Operations Research,* Vol. 5, No. 2, 2000, pp. 19–36.

United States Joint Forces Command (USJFCOM), United States Joint Forces Command Glossary, Web Site http://www.jfcom.mil/about/glossary.htm#D, accessed 30 November, 2005.

ABSTRACT

Hyperspectral anomaly detection is a useful means for using hyperspectral imagery to locate unusual objects. Current anomaly detection methods commonly use non-robust statistical methods that may lead to inaccurate detection results. This research explores the use of different multivariate outlier detection methods for the anomaly detection problem. Theoretically, these methods are better suited than existing anomaly detection methods for finding anomalous objects in a hyperspectral image. This hypothesis is tested by applying a range of outlier detection methods to both simulated and real-world image data. Test results indicate that multivariate outlier detection can achieve superior detector performance relative to benchmark anomaly detection methods.

INTRODUCTION

Hyperspectral imaging of the Earth's surface provides a unique means for identifying objects of interest. Specifically, the unique spectral signatures of the materials in an image can be measured and compared to reference signatures to positively identify an object. If reference signatures are not available, hyperspectral image data can still be used to locate objects that are anomalous to the predominant background materials contained in the image.

The basic hyperspectral anomaly detection problem is an exercise in detecting outliers in a multivariate dataset. Though numerous methods have been proposed in the technical literature to detect multivariate outliers, there is surprisingly little evidence that these methods have been used for hyperspectral anomaly detection. Rather, the majority of anomaly detection methods found in the literature rely on non-robust statistical methods known to be unreliable in the presence of outlying observations. Use of such methods that do not accommodate the effects of outliers can lead to the undesirable effects of masking and swamping. Masking refers to the phenomenon of strong outliers skewing statistical estimates to the degree that weaker outliers do not appear abnormal. Swamping, on the other hand, refers to the case in which outliers distort statistical estimates in such a way that otherwise normal observations appear to be outlying. If the goal of an analysis is to correctly identify anomalies, we can say that masking leads to a decrease in true positives and swamping leads to an increase in false positives. In almost all applications, either one of these conditions is to be avoided.

The focus of this research effort is to study the ability of different multivariate outlier detection methods to find hyperspectral anomalies while minimizing the masking and swamping effects. To this end, we have selected four outlier detection methods from the literature and adapted them for use as hyperspectral anomaly detectors. These detectors are tested against both simulated hyperspectral data and real-world imagery, and their performance is compared to that of benchmark anomaly detection methods from the literature.

In the remainder of this paper, we first provide background information on basic hyperspectral concepts and related anomaly and multivariate outlier detection literature. We then give an overview of the outlier detection algorithms we evaluated, followed by a summary of the simulated and real-world performance tests. The paper concludes with significant insights gained from these tests, as well as recommendations for further research.

BASIC HYPERSPECTRAL CONCEPTS

To gain a basic understanding of hyperspectral imagery, we can begin with a discussion of the common digital camera that has become ubiquitous in modern society. Conceptually, when we use a digital camera to take a color photograph, the camera divides the imaged scene into a two-dimensional grid of pixels. For each pixel, three pieces of information are collected. Namely, the amount of energy emanating from the pixel in the red, green, and blue portions of the electro-magnetic (EM) spectrum. This information is stored in three separate two-dimensional arrays. For any given pixel, combining its respective red, green, and blue information produces the true color of the pixel. Of course, viewing the array of colored pixels on a computer screen or in its printed form reveals the scene originally photographed.

A Comparison of Multivariate Outlier Detection Methods For Finding Hyperspectral Anomalies

Timothy E. Smetek

Air Force Institute of Technology
timothy.smetek@afit.edu

Kenneth W. Bauer Jr.

Air Force Institute of Technology
kenneth.bauer@afit.edu

APPLICATION AREAS:
ISR and Intelligence Analysis

OR METHODOLOGIES:
Multivariate Analysis and Pattern Recognition

If we image a scene for the purpose of identifying different objects that it may contain, a simple color image produced by a digital camera may suffice; however, a true-color image has its limitations. To address these limitations, hyperspectral sensors collect information beyond the visible region of the electro-magnetic spectrum. Just as a digital camera produces three images for wavelength bands corresponding to red, green, and blue light, a hyperspectral sensor produces images for many different contiguous wavelength bands, typically spanning the visible to near-infrared regions of the EM spectrum. The number of image bands collected by a sensor can range from twenty to over 500. When the separate two-dimensional image bands are "stacked" on top of each other, the resulting three-dimensional array is referred to as the image cube. A notional schematic of this imaging concept is shown in Figure 1.

From Figure 1 we can see that there is a pixel in each image band that corresponds to the same spatial location of the scene. For example, the pixel in row *m*, column *n* of band 1 refers to the same spatial location of the scene as the pixels in row *m*, column *n* of every other band in the image cube. The sensor reading for a pixel in row *m*, column *n*, and band λ, can be referred to by the variable $x_{mn\lambda}$. For a given pixel address (*m, n*), we can form the vector:

$$\begin{Bmatrix} x_{mn1} \\ x_{mn2} \\ \vdots \\ x_{mnP} \end{Bmatrix} \tag{1}$$

where *P* is the total number of image bands in the image cube. This vector is often referred to as a pixel vector. If we take the transpose of all the pixel vectors in the image and place them in a *MNxP* array, where *M* is the total number of rows and *N* the total number of columns, we then have a data matrix, **X,** that is commonly used in multivariate statistical analysis. Using **X,** we are free to analyze the image data using multivariate analysis methods such as principal component analysis, cluster analysis, maximum likelihood classification, discriminant analysis, and others.

As shown in Figure 1, we can also plot the elements of a pixel vector against their respective image band numbers. Such a plot reveals the spectral signature of the material contained in pixel (*m, n*). Due to the chemical and physical properties of different materials, these spectral

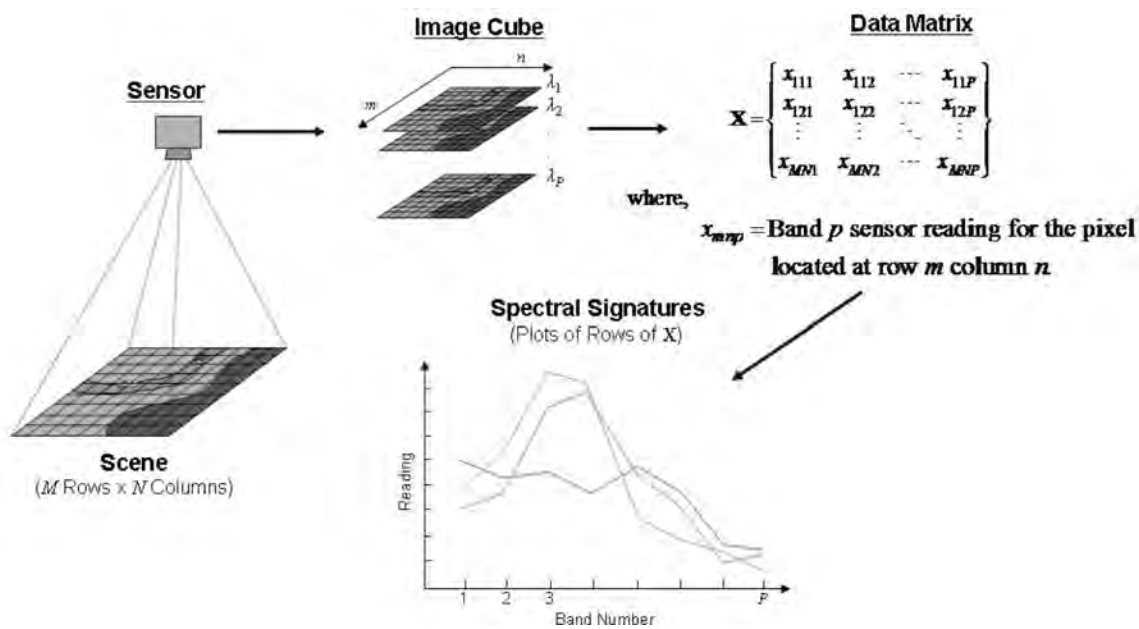

Figure 1. The basic hyperspectral imaging process and data representation.

signatures can be used to identify the materials in an image scene. Thus, if we have reference library signatures for known materials, it is possible to compare pixel signatures to library signatures to locate and identify objects of interest. Unfortunately, there are several complicating factors that make this signature matching strategy difficult to execute. First and foremost, library signatures typically describe the manner in which a material reflects a known energy source, whereas a hyperspectral sensor detects the amount of energy radiating from a pixel's spatial location. The process of converting a radiance signature to a reflectance signature, or vice a versa, is a complicated process that requires knowledge of the atmospheric conditions, sun angle, sensor location, object orientation, and other parameter values that existed at the time the hyperspectral image was acquired. Additional obstacles to successful signature matching include material weathering, objects that are smaller than a pixel's spatial dimension, the inherent variability of an object's spectral signature, sensor noise, and other inaccuracies imposed by the image acquisition process.

As an alternative to signature matching, the data matrix, **X,** can also be searched for pixel vectors that are anomalous relative to the majority of the pixel vectors contained in the image cube. In the lexicon of hyperspectral analysis, this process is referred to as anomaly detection. The advantage of anomaly detection relative to signature matching is that the sensor data can be analyzed without concern for radiance-reflectance conversion. The disadvantage, however, is that anomalous pixel vectors may not actually be objects of interest and must be further analyzed to determine their true identity. Nonetheless, anomaly detection provides a feasible means for locating potential objects of interest when little is known about an object's spectral signature or the conditions under which an image was acquired. To make an anomaly detection analysis most valuable, the number of false alarms should be kept to a minimum since each anomalous pixel vector must be further analyzed—either visually or by another sensor—to confirm its status as an object of interest.

The preceding discussion has provided an admittedly brief overview of basic hyperspectral concepts. The intent of this overview is to provide the proper context for the remainder of this paper. For a more detailed discussion of hyperspectral imagery and its applications, the reader is referred to the article by Landgrebe (2002), as well as the texts by Landgrebe (2003), Chang (2003), and Richards and Jia (1999).

RELATED LITERATURE

There are two primary bodies of literature upon which this present research effort is built: hyperspectral anomaly detection and multivariate outlier detection. In the following paragraphs, these research areas are summarized in order to provide vectors for more in-depth study of these fields. For more detailed discussions on anomaly detection, the reader is referred to the review provided by Stein et al. (2002), while the text by Barnett and Lewis (1994), or the survey provided by Beckman and Cook (1983), serve as good entry points to the study of outlier detection. A more detailed survey of the anomaly detection and multivariate outlier detection literature can also be found in Smetek and Bauer (2007).

Anomaly Detection Literature

Anomaly detection methods can be categorized under two general strategies: local detection, and global detection. Local anomaly detectors are characterized by the use of processing windows that are passed over every pixel in an image to find anomalies. In general, the pixel vectors contained in the processing window are used to characterize the local background materials. The pixel vector in the center of the window is then tested relative to this background to determine if it is an anomaly. The original detection method that fostered the local window scheme is the RX detector proposed by Reed and Yu (1990). For the RX detector, assuming Gaussian data, the test statistic computed for each pixel vector, **x,** relative to its local processing window is given by:

$$RX(\mathrm{x}) = (x - \hat{\mu})^T \left(\frac{N}{N+1} \mathrm{S} + \frac{1}{N+1} (\mathrm{x} - \hat{\mu})(\mathrm{x} - \hat{\mu})^T \right)^{-1} (\mathrm{x} - \hat{\mu}) \tag{2}$$

where,

$\hat{\mu}$ = the window mean vector,
S = the window covariance matrix, and
N = the number of pixel vectors in the processing window.

Asymptotically, as N becomes large, (2) reduces to the following:

$$RX(\mathrm{x}) = (\mathrm{x} - \hat{\mu})^T \mathrm{S}^{-1} (\mathrm{x} - \hat{\mu}) \tag{3}$$

which is simply the Mahalanobis Squared Distance (MSD) for **x** relative to the mean vector and covariance matrix of the processing window. Again using the Gaussian assumption, (3) can be compared to an appropriate quantile of the χ^2-distribution with p degrees of freedom—where p is the dimensionality of the data—to assess if **x** is an anomalous pixel vector.

Several limitations are known to exist with the RX detector. Specifically, the algorithm has difficulty locating large anomalies, its performance degrades with the signal-to-noise ratio of the sensor, and the Gaussian assumption for the processing window data is not always accurate. RX-based detectors that attempt to correct these limitations include those proposed by Chang and Chiang (2001), Hsueh and Chang (2004), Riley et al. (2004), Kwon and Nasrabadi (2005), and Gaucel et al. (2005). Of particular interest are modifications proposed by Schaum (2004), West et al. (2005), and Schaum (2006). These methods attempt to account for heterogeneous materials in the processing window that may produce inaccurate background mean and covariance estimates.

In addition to these RX modifications, several local detectors have been proposed that dispense with the metric in (3) while retaining the processing window strategy. These methods include multiple—window detectors given by Kwon et al. (2003), Liu and Chang (2004), Rosario (2004), and Goovaerts et al. (2004). Local detectors have also been proposed that attempt to use Markov Random Fields (MRFs) to capture contextual information provided by a pixel's surrounding neighbors. Examples of these MRF detectors can be found in Schweizer and Moura (2000), Schweizer and Moura (2001), and Hazel (2000).

Where local anomaly detectors attempt to classify a pixel based on local window statistics, global detectors strive to find pixels that are anomalous relative to the entire image scene. By using this philosophy, global detectors are theoretically better-suited to find large anomalies and avoid false alarms resulting from scene clutter. In general, global detectors can be divided into two groups: mixture model methods and distribution-based methods.

In the case of mixture model detectors, the fundamental assumption is the existence of M distinct background material spectra—referred to as endmember spectra—with characteristic signatures given by the vectors, s_m, $m = 1, \ldots, M$. Thus, each pixel vector, **x,** in the image can be represented as a linear combination of the M endmember spectra plus an additive noise vector, **n:**

$$\mathrm{x} = \sum_{m=1}^{M} \alpha_m s_m + \mathrm{n}. \tag{4}$$

where the α_m are referred to as the abundance fractions since they relate the relative abundance of each endmember material contained in the pixel. The endmember vectors may be actual signatures found in the image, or may be derived in such a way that they span the subspace defined by the image. Once the endmembers are found, anomalies are detected by finding the pixels that produce the greatest residuals when fit by (4). Alternatively, endmembers that represent interesting signatures are identified from the M endmembers, and those image pixels that are most similar to the some interesting object's endmembers are identified as anomalies. Implementations of mixture model anomaly detectors are given by Grossman et al. (1998), Stein et al. (2002), and Clare et al. (2003).

The fundamental challenges of global mixture model anomaly detectors are determining the appropriate number of endmembers and actually selecting the endmember spectra themselves. To avoid these issues, the distribution-based view of global anomaly detection can be adopted. The basic premise of these methods is that each of the background materials in an image

can be represented by a multivariate distribution. Detecting anomalies under this assumption entails estimating the distributions contained in the image and determining the likelihood that each pixel in the image comes from one of the distributions. Pixels with low likelihood relative to all distributions are considered anomalies. The primary characteristics that differentiates distribution-based detectors is the method used to determine the constituent distributions of an image scene and the form of the distributions. Schaum and Stocker (1997) and Stein et al. (2002) outline a detector that assumes the component distributions are Gaussian and uses stochastic expectation-maximization to estimate the distribution parameters. The method proposed by Carlotto (2005) also assumes a mixture of Gaussian distributions, but uses *k*-means clustering to identify the sample of pixels for each distribution. In a similar manner, Catterall (2004) uses *k*-means to determine which pixel vectors belong to which distributions, but assumes Multivariate Normal Inverse Gaussian (MNIG) distributions when estimating the parameters. In order to avoid the issue of density estimation altogether, Chiang et al. (2001) and Achard et al. (2004) use projection pursuit to find univariate projections of the image data to locate pixel vectors that induce the most skewness or kurtosis in the projected data.

To conclude this discussion on existing anomaly detection methods, it is significant to note the apparent absence in the existing literature of attempts to use multivariate outlier detection to find hyperspectral anomalies. This is an interesting omission considering that hyperspectral anomalies could be considered nothing more than outlying observations in a multivariate data set. With this idea in mind, we now focus attention on the significant methods that have been proposed in the multivariate outlier detection literature.

Multivariate Outlier Detection Literature

Detecting unusual, or outlying, observations has been a subject of interest for hundreds of years. However, due to the computational challenges posed by multivariate data, multivariate outlier detection first received serious attention in the early 1970s with the survey conducted by Gnanadesikan and Kettenring (1972). Since that time, a number of outlier detection algorithms have been proposed in the pursuit of several ideal characteristics: 1) the detector should have a high breakdown point, defined as the percent of the dataset that can be outlying while still allowing the algorithm to detect the outliers; 2) the detector should be computationally efficient with large datasets in high dimension; and 3) the detector should be affine equivariant so that detection results do not change due to translations, rotations, or scaling of the data. Needless to say, these goals are somewhat lofty, and no single method has succeeded in satisfying the complete set.

In order to digest the range of multivariate outlier detection methods that have been proposed, we categorize them into two groups: robust distance methods, and non-traditional methods. Robust distance methods—by far the most popular of the two detection strategies—attempt to assess if an observation is outlying using the MSD in a similar manner to the RX detector mentioned previously; however, rather than using the classical mean and covariance estimate for the data when computing the distance, robust estimates are used instead. For the case of non-traditional methods, alternative characteristics of outlying observations, besides large MSDs, are exploited to reveal the outliers. Existing methods that fall in these two categories are summarized in the following paragraphs.

The majority of robust distance detectors found in the literature use some variant of the Minimum Volume Ellipsoid (MVE) or Minimum Covariance Determinant (MCD) estimators proposed by Rousseeuw (1983) to obtain robust mean and covariance estimates. The MVE estimation method involves finding the ellipsoid of minimum volume that encompasses at least $h = [(n + p + 1)/2]$ of the observations, where n is the sample size, p is the dimension of the data, and h is often referred to as the "half-sample." The mean vector is taken to be the center of the ellipsoid and the covariance estimate defines the ellipsoid. MCD estimation—which has better convergence and efficiency characteristics than the MVE estimator—consists of finding the subset of h observations whose covariance matrix has the mini-

mum determinant. The MCD mean estimate is then the centroid of the half-sample and the covariance estimate is the covariance of the half-sample, scaled for consistency. The MVE and MCD estimates are attractive for robust-distance outlier detection because they have breakdown points near 50% and are affine equivariant. However, finding these estimates entails solving combinatorial optimization problems whose approximate solutions are generally found using heuristic search. Robust distance detectors that use approximate MVE estimates are proposed by Rousseeuw and Leroy (1987), Rousseeuw and van Zomeren (1990), Hadi (1992), Hadi (1994), and Atkinson (1994). MCD-based detectors include the Feasible Solution Algorithm of Hawkins (1994), the compound method of Rocke and Woodruff (1996), the FAST-MCD method of Rousseeuw and van Driessen (1999), an iterative deletion method proposed by Viljoen and Venter (2002), and a robust clustering method given by Hardin and Rocke (2004). This last method is somewhat interesting in that it is apparently the only multivariate outlier detection method that assumes the majority of the dataset is composed of observations from multiple "good" populations vice a single population. By "good" we mean that observations from the population can be expected in the dataset and should not be considered outlying.

A significant limitation of MVE and MCD-based methods is the computational complexity of even approximate solution methods. To overcome this weakness, several attempts have been made to develop robust distance methods that are computationally fast at the expense of being less theoretically formal than MVE and MCD methods. These detectors include the Smallest Half-Volume and Resampling by Half-Means methods of Egan and Morgan (1998), the Blocked Adaptive Computationally efficient Outlier Nominator (BACON) detector of Billor et al. (2000), and the Closest Distance to Center method of Chiang et al. (2003).

An alternative approach to robust distance detection is to apply weights to the observations to obtain robust mean and covariance estimates, with more weight given to observations near the center of the data. Examples of this detection approach include an *M*-estimation method proposed by Campbell (1980), and the Stahel-Donoho Estimator method developed by Maronna and Yohai (1995). These methods are theoretically interesting, but are difficult to implement due to the nonlinear optimization problems that must be solved to obtain their respective robust estimates.

Turning our attention to non-traditional outlier detection methods, these types of detectors attempt to exploit some alternative characteristic of outlying observations in their detection scheme besides large Mahalanobis distances. Several methods based on principal component analysis are given by Gnanadesikan and Kettenring (1972), Chiang et al. (2003), and Gao et al. (2005). Kim (2000) gives a decomposition of the Mahalanobis distance and uses it to generate scatter plots that can be informally analyzed to reveal outliers. Pan et al. (2000) use projection pursuit to analyze the data in higher-dimensional space, while Juan and Prieto (2001) study the distribution of the data projected on the unit hypersphere to detect concentrations of outlying observations.

To conclude this discussion on multivariate outlier detection research, two points should be noted. First, virtually all the methods discussed in the previous paragraphs assume Gaussian or elliptically distributed data. The primary reasons for this assumption is to validate statistical tests of significance used for formal outlier detection, and to ensure the data has an elliptical structure that lends itself to a particular type of analysis. Second, as mentioned previously, all the methods found in the literature, with the exception of Hardin and Rocke's clustering method, assume the data comes from a single population contaminated by outliers from other populations. If a dataset is known to contain observations from multiple "good" populations, applying single population detection methods is not appropriate. To apply multivariate outlier detection methods to hyperspectral data, these issues must be considered.

ALGORITHM OVERVIEW

In the preceding sections we have attempted to introduce the basic concepts of hyperspectral imagery, as well as the existing literature pertaining to hyperspectral anomaly detection and multivariate outlier detection. We now return to the

purpose of this research which is to compare the effectiveness of different outlier detection methods in finding hyperspectral anomalies. In this section, we outline the four outlier detection methods used in our assessment. These detectors include the FAST-MCD method of Rousseeuw and van Driessen (1999), the BACON method of Billor et al. (2000), the non-traditional method proposed by Juan and Prieto (2001), and a modification of the Stahel-Donoho Estimator (SDE) method of Maronna and Yohai (1995). These methods were selected as potential anomaly detectors for two reasons. First, they collectively represent the different types of multivariate outlier detectors discussed earlier. Second, these methods appear best able to accommodate extremely large hyperspectral datasets that typically contain tens of thousands of observations in high-dimensional space.

Each algorithm is outlined in the following paragraphs. It should be noted, however, that the algorithms are only described in enough detail to convey their respective solution strategies. The reader interested in implementing these methods should consult the original articles for the necessary details.

The FAST-MCD Detector

The primary objective of the FAST-MCD detector is to rapidly search for a solution to the following non-linear optimization problem:

$$\min \det(S) = \det\left(\frac{\sum_{i=1}^{n} t_i (x_i - \sum_{j=1}^{n} t_j x_j / \sum_{j=1}^{n} t_j)(x_i - \sum_{j=1}^{n} t_j x_j / \sum_{j=1}^{n} t_j)^T}{\sum_{i=1}^{n} t_i - 1} \right) \quad (5)$$

s.t.

$$\sum_{i=1}^{n} t_i = h = \left[\frac{(n + p + 1)}{2} \right],$$

$$t_i \in \{0,1\} \forall i = 1,\ldots,n.$$

where $\mathbf{x}_i$ is an observation vector, n is the total number of observations in the dataset, det(•) is the determinant operator, and $(\bullet)^T$ is the transpose operator. The search is conducted by first selecting a user-specified number of random subsets of size h from the original dataset. For each subset, a *C*-step procedure is performed consisting of the following: 1) the Mahalanobis squared distances are computed for all observations in the dataset using the mean vector and covariance matrix of the subset data; 2) the distances are sorted; and 3) the h observations from the original dataset with smallest squared distances are used to form a new subset. Rousseeuw and van Driessen (1999) prove that repeated applications of the *C*-step procedure to a dataset will produce a new subset of size h that has a covariance determinant less than or equal to that of the preceding estimate.

After applying the *C*-step procedure to each random subset until convergence of the respective covariance determinant, the subset that produced the smallest covariance determinant is identified. The mean vector of this subset is used for the robust mean estimate of the original dataset, and the covariance matrix of the subset is used as the robust estimate of the data's shape matrix. This shape matrix is then scaled to be consistent with Gaussian data in the sense that the median of the Mahalanobis squared distances obtained using the scaled covariance matrix is equal to the 0.5-quantile of a Chi-squared distribution with p degrees of freedom. The resulting scaled matrix becomes the robust covariance estimate. These robust estimates are used to compute robust Mahalanobis squared distances for each observation in the dataset. Any observation whose squared distance exceeds an appropriate quantile of the Chi-squared distribution with p degrees of freedom is considered an outlier.

To allow the FAST-MCD method to handle very large datasets, Rousseeuw and van Driessen also propose a nesting scheme that initiates the search by selecting a random sample of the original data and forming the initial subsets from this random sample. As the search proceeds, more and more of the original data is included in the search until the final solution is obtained. The FAST-MCD method is imple-

mented in S-Plus 4.5 as the cov.mcd function and in SAS/IML 7 as the MCD function.

The BACON Detector

The BACON detector proposed by Billor et al. (2000) is a robust distance detector designed to rapidly identify outliers in very large datasets. The algorithm is relatively simple to implement with the added advantage that it is very fast relative to the other detectors we consider, even for extremely large datasets.

BACON attempts to find outlying observations by first identifying a basic subset of "clean" observations close to the centroid of the data. The user has the option of using either a robust, non-affine equivariant or a non-robust, affine equivariant method to find this subset. Once determined, the basic subset is used to estimate a mean vector and shape matrix for the dataset. The shape matrix is multiplied by a small-sample correction factor derived by Billor et al. from a Monte Carlo simulation study. Using the mean vector and scaled shape matrix, Mahalanobis squared distances are computed for all observations in the dataset. Any observations whose squared distances are less than an appropriate quantile of the Chi-squared distribution with p degrees of freedom are then used to form a new basic subset. This process is repeated until the basic subset fails to increase in size between iterations. Any observations not in the basic subset when the algorithm terminates are considered outliers.

The Juan-Prieto Detector

The outlier detector proposed by Juan and Prieto (2001)—hereafter referred to as the Juan-Prieto detector—is a non-traditional outlier detection method that avoids the computation of Mahalanobis distances altogether. Thus, the method offers a good contrast to the other methods we consider. The Juan-Prieto detector is also designed to locate concentrated outliers, which, intuitively, would seem to match well with the problem of finding interesting objects in a hyperspectral scene.

The underlying statistical theory exploited by the Juan-Prieto detector is that p-dimensional Gaussian data projected onto the p-dimensional unit hypersphere has a Uniform distribution. Further, the angles between each normalized vector and a reference direction will have a Beta distribution. These properties are also reasonably robust to departures from normality if the data is elliptically symmetrical. With this theory in mind, the Juan-Prieto detector begins by normalizing all the observation vectors so that they have a magnitude of one, and thus lie on the p-dimensional unit hypersphere. A reference direction is then chosen using a non-linear optimization method suggested by Juan and Prieto, and the angles between the reference direction and each normalized vector are computed. To determine if these angles have the prescribed Beta distribution, they are entered as arguments to the inverse of the appropriate Beta distribution function. If the angles indeed have the proper distribution, the outputs to the inverse distribution function should, in turn, have a Uniform distribution. This hypothesis is tested by analyzing the maximum spacing between the ordered inverse function outputs. If the maximum spacing is not consistent with a Uniform distribution, all corresponding observations beyond the maximum spacing in the ordered inverse function outputs are considered outliers.

The Modified Stahel-Donoho Estimator (SDE) Detector

The original SDE detector proposed by Maronna and Yohai (1995) is a robust distance method that arrives at mean vector and covariance matrix estimates using a robust estimation method originally proposed by Stahel (1981) and Donoho (1982). The SDE mean vector, **T,** and covariance matrix, **S,** are given by:

$$\mathrm{T}(\mathrm{X}) = \frac{\sum_{i=1}^{n} w_i \mathrm{x}_i}{\sum_{i=1}^{n} w_i} \qquad (6)$$

and

$$S(X) = \frac{\sum_{i=1}^{n} w_i (x_i - T(X))(x_i - T(X))^T}{\sum_{i=1}^{n} w_i} \quad (7)$$

where the w_i are weights whose magnitudes depend on the degree to which the corresponding observation is outlying. Though different weight functions can be employed, Maronna and Yohai demonstrate empirically that the following function provides good statistical efficiency of the estimator:

$$w_i = I(r_i \leq c) + (c/r_i)^q I(r_i \rangle c) \quad (8)$$

Where

$I(g)$ = the indicator function, and
r_i = the measure of "outlyingness" for observation i.

The parameters c and q in (8) are constants. Maronna and Yohai suggest values for these constants based on experimental tests that attempt to minimize a "bias" measure for the covariance estimate. Details of this experiment are beyond the scope of this paper, but can be found in the original article by Maronna and Yohai. The r_i metric in (8) for an observation vector, $\mathbf{x}_i$, is defined as:

$$r_i = \sup_{\|a\|=1} \left\{ \frac{|a^T x_i - \operatorname{med}_j(a^T x_j)|}{\operatorname{med}_k |a^T x_k - \operatorname{med}_j(a^T x_j)|} \right\} \quad (9)$$

The interpretation of (9) is we are looking for some projection vector, **a,** on the p-dimensional unit hypersphere that maximizes the standardized distance between the projection of $\mathbf{x_i}$ onto **a** and the centroid of the projected dataset onto **a.** To ensure a robust estimate of r_i, the median of the projected data is used to estimate the centroid, and the median absolute deviation (MAD) is used to estimate the standard deviation. The rationale for using (9) to measure "outlyingness" is that for elliptically symmetric data, an outlier in p-dimensional space will be an outlier in some univariate projection of the data.

Once the robust estimates of (6) and (7) are obtained, they can be used to compute robust Mahalanobis distances for all the observations in the dataset. Maronna and Yohai suggest that these distances are F-distributed, and provide a suitable critical value for screening them for outliers. Hence, implementing the SDE outlier detector entails: 1) computation of the r_i for each observation; 2) using these values to compute (6) and (7); 3) using the robust estimates to compute Mahalanobis squared distances for the observations; and 4) using the appropriate critical value to screen the distances for outliers. The practical challenge in using this detector, however, is solving the non-linear optimization problem given by (9). Due to the non-differentiable objective function, derivative-free optimization methods must be used to search for a local solution.

Rather than using penalty or barrier function methods to solve (9), Maronna and Yohai suggest generating random points on the unit hypersphere that have a Uniform distribution. Each point, or vector, is then substituted into (9) to find an approximate solution to the maximization problem. As an alternative to random vector generation, we propose using number theoretic methods (NTM) to generate points that are uniformly scattered—as opposed to uniformly distributed—on the unit hypersphere. We favor this method because NTM point generation requires fewer points to evenly cover the unit hypersphere than random point generation, as explained in Fang and Wang (1994). Thus, given the same number of points generated by the two methods, we can be more confident of an "even" search of the feasible region with NTM generation than with random generation.

By modifying Maronna and Yohai's SDE detector using NTM point generation, we define the SDE-NTM generator as follows:

1) Generate a set of uniformly scattered points, or vectors, on the p-dimensional unit hypersphere using the TFWW method outlined in Fang and Wang (1994).
2) Use the vectors from Step 1 to find an approximate solution to (9) for each observation.
3) Use the r_i's from Step 2 to compute the mean vector and covariance estimates given by (6) and (7), respectively.

4) Compute the Mahalanobis squared distance for each observation relative to the robust mean and covariance estimates computed in Step 3.
5) Scale the squared distances from Step 4 by the median of the squared distances, and declare as outliers any observation whose scaled squared distance, d^*, exceeds $F(\alpha/n; p, n\text{-}2p)/F(0.5; p, n\text{-}2p)$, where α is a specified significance level and $F(\bullet; a, b)$ is the F-distribution function with a and b degrees of freedom.

The critical value given in Step 5 is based on empirical simulation studies conducted by Maronna and Yohai that indicate the following:

$$d^*_{(i)} \approx F(i/(n+1);p;n-k)/F(0.5;p;n-k) \tag{10}$$

where,

$$d^*_{(i)} = \text{the } i\text{th ordered } d^*, \text{ and}$$

$$k \in [p, 2p].$$

Though the SDE-NTM detector offers a more efficient procedure for finding approximate solutions to (9) relative to the original SDE detector, the method is still computationally expensive, particularly in high-dimensions. To reduce the number of unnecessary computations, we suggest computing and storing uniformly scattered sets of points for different combinations of dimensionality and numbers of points.

Summary

The preceding paragraphs outlined the four multivariate outlier detection methods used in the comparison tests described in the following sections. Again, these methods were selected based on their perceived ability to handle very large datasets, as well as the different detection strategies they employ. By comparing the relative performance of this diverse set of detectors, it is hoped that useful insights may be obtained as to how best multivariate outlier detection may be used to find hyperspectral anomalies.

SIMULATED DATA TESTS

To assess the relative anomaly detection performance of the FAST-MCD, BACON, Juan-Prieto, and SDE-NTM detectors, we first applied them to controlled datasets consisting of simulated hyperspectral signatures contaminated with known quantities of outlying signatures. To produce the simulated signatures, we collected signatures of known materials from a hyperspectral image of Fort A.P. Hill, Virginia, taken with the airborne COMPASS sensor. The signatures we collected correspond to grass, road, dead grass, and shadow. The mean spectral signatures for each material are shown in Figure 2; the error bars in the plots indicate the standard deviation of the signatures at each band. To ensure our results were not image dependent, we also collected grass, asphalt, gravel, and water signatures from an image of the National Mall in Washington, D.C., acquired from the airborne AVIRIS sensor. These signatures are also shown in Figure 2.

For each material, we used the collected signatures to obtain a mean vector and covariance matrix estimate. These estimates were then used to generate random, multivariate Gaussian signatures with the same mean vector and covariance matrix as the real-world signatures. Because the hyperspectral analysis community is divided on the validity of the Gaussian assumption, we also used the reference signatures to generate multivariate t-distributed signatures with twelve degrees of freedom. We chose this alternative distribution based on the research of Manolakis and Marden (2002), Kerekes and Manolakis (2004), and Manolakis et al. (2005).

With this strategy for generating simulated hyperspectral signatures in mind, our simulated data tests were executed in the following manner:

1) From the available materials, one material was chosen as the background material and another was chosen as the outlying material.
2) Two thousand background signatures were generated using the Gaussian distribution.
3) A specified number of outlying signatures were generated using the Gaussian distribution.

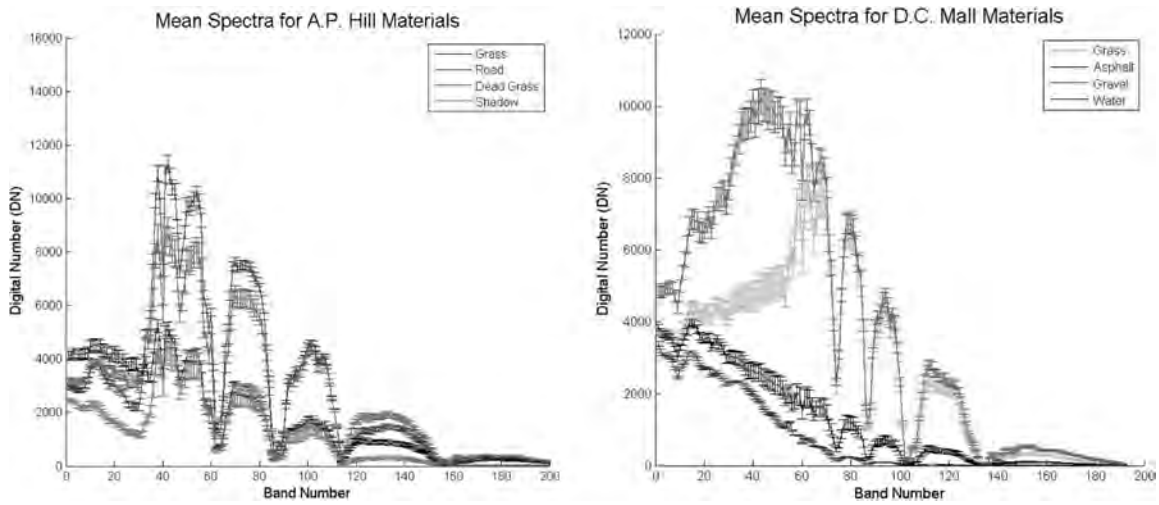

Figure 2. Mean spectra for materials used to generate simulated hyperspectral data. Error bars indicate standard deviation of data.

4) The background and outlier materials were combined into a single dataset.
5) Each of the four outlier detection methods was applied to the dataset to detect the outliers, with the number of true-positives and false-alarms recorded for each method. We also applied a classical Mahalanobis distance detector to the dataset to better assess the benefits of using multivariate outlier detection methods.
6) Steps 2 through 5 were repeated 30 times, keeping the number of outliers constant. The mean and standard deviation of the number of true-positives and false-alarms were computed for each method across these 30 repetitions.
7) Steps 2 through 6 were then repeated for another level of contamination. The level of contamination was varied from 0 to 500 outliers in increments of 50 outliers.
8) Steps 1 through 7 were repeated using a different combination of background and outlier materials while ensuring both materials came from the same image.
9) Steps 1 through 8 were repeated using the multivariate *t*-distribution.

The combinations of background and outlier materials used in our tests are listed in Table 1. We used these combinations because it was shown in Smetek and Bauer (2007) that they suffer significantly from the masking effect when using classical Mahalanobis distance detection.

The outcome of our tests are summarized in Tables A-1 through A-4 of the Appendix. Tables A-1 and A-2 show the mean true-positives obtained by each detector for the Gaussian and multivariate *t*-distributed data, respectively. Similarly, Tables A-3 and A-4 show the

Table 1. Background/Outlier material combinations used for simulated data tests.

Fort A.P. Hill Combinations		D.C. Mall Combinations	
Background	Outlier	Background	Outlier
Grass	Road	Grass	Asphalt
Grass	Shadow	Grass	Water
Dead Grass	Shadow	Asphalt	Water
Road	Shadow	Gravel	Asphalt

mean number of false-positives for the two distributions. For each mean value, the standard error is also reported as a measure of detector performance variability. To keep these tables as concise as possible, we have only included results from a subset of the contamination levels tested; however, we feel they are sufficient to show the relative performance of the detectors.

From Tables A-1 and A-2, several conclusions can be made. First, it is clear that the classical, non-robust Mahalanobis distance detector suffers significantly from masking, as indicated by the low number of true positives across all the material combinations, contamination levels, and distributions. Second, we see that the BACON, FAST-MCD, and SDE-NTM detectors successfully identify all outliers in the cases tested. This finding is true for both the Gaussian data and the multivariate *t*-data, which confirms similar conclusions reported in Smetek and Bauer (2007). The ability of these detectors to successfully find outliers in heavy-tailed distributions is important since it provides an alternative to the challenging task of correctly identifying a specific distribution from the multivariate *t*-distribution family. A third observation from Tables A-1 and A-2 is the inability of the Juan-Prieto detector to find outliers when the contamination level is relatively low. The likely cause of this limitation is that relatively few outliers are not likely to affect the uniformity of the data when projected onto the unit hypersphere.

Turning to the false positive data reported in Tables A-3 and A-4, it is seen that when all detectors are applied to the Gaussian data, the number of false positives is close to zero for all levels of contamination and material combinations. The reason for this seemingly ideal false positive rate is the use of a Bonferoni significance level of α/n used to threshold the respective test statistics for the different detectors, where $\alpha = 0.05$ and n is the total number of observations. For all cases tested, the expected number of false alarms for the significance level used is less than one.

In the case of the multivariate *t*-data, the false alarm data is somewhat more interesting. First, we note that the false alarms for the BACON detector remain close to zero. In contrast, the false alarms for the FAST-MCD and SDE-NTM methods are significantly higher. The reason for this difference is that both the FAST-MCD and SDE-NTM methods arrive at a covariance estimate by first estimating the shape matrix of the data by trimming away all the observations far from the center. Because "good" observations may also be trimmed, the shape matrix underestimates the true variance of the good data. Hence, either the shape matrix or the Mahalanobis distances derived from it must be scaled before testing the distances for outliers. We have found that the scaling process used for both the FAST-MCD and SDE-NTM methods still tend to underestimate the true variance in the data, particularly when compared to the scaling method used by the BACON detector. Hence, it is reasonable to expect more false alarms with the FAST-MCD and SDE-NTM methods relative to the BACON detector, particularly with heavy-tailed data.

A final observation of note in the false alarm data is the decreasing number of false alarms for the FAST-MCD and SDE-NTM detectors as the level of contamination increases. We hypothesize that this phenomenon occurs for the following reason: as the contamination level increases, it is more likely that outliers are still contained in the set of observations used to estimate the shape matrix, since neither the FAST-MCD nor the SDE-NTM methods are guaranteed to generate a "clean" estimate. Though few in number, these outliers are sufficient to artificially increase the variance of the data. This increased variance will, in turn, result in a lower false alarm rate. A similar affect was demonstrated in Smetek and Bauer (2007).

REAL-WORLD IMAGE TESTS

To confirm that our simulated test results translate to actual hyperspectral images we applied the BACON, FAST-MCD, and SDE-NTM detectors to real-world hyperspectral scenes. We omitted the Juan-Prieto detector from this test due to its limited ability to detect small numbers of outliers, as indicated by the simulated data tests—initial hyperspectral image tests also supported this omission. In addition to the three multivariate outlier detectors, we

also applied the RX anomaly detector—outlined earlier in this report—and a Cluster detector to the image scenes. The Cluster detector is a global anomaly detector based on the Cluster-Based Anomaly Detector (CBAD) proposed by Carlotto (2005). The performance of the RX and Cluster detectors provide a performance benchmark used to assess the merit of the multivariate outlier methods.

Three hyperspectral image scenes were used for this test. Scene 1 was taken from the Forest Radiance I dataset acquired by the HYDICE sensor. The scene contains a number of manmade square panels of varying sizes and materials arranged in a grid pattern in a relatively uncluttered grass field. Scene 2 is also taken from the Forest Radiance I dataset, but is somewhat more complex. This scene contains a number of vehicles with different paint schemes parked in a grass field, and also contains additional background materials such as road, trees, dirt, and shadow. Scene 3 is the Fort A.P. Hill image scene from which we obtained the signatures for the simulated data tests. This image contains manmade objects of varying size and compositions strewn throughout a grass field. The scene also contains road, trees, dirt, shadow. Taken collectively, these three scenes offer a range of objects, sensors, and scene complexity to adequately test and compare the detection methods.

Using the fore-mentioned detectors and image scenes, the real-world image tests proceeded in the following manner. Each detector was applied to each of the scenes to identify pixel vectors that appeared anomalous. Using ground-truth images, Operating Characteristic (OC) curves were constructed depicting the true-positive fraction of the detectors as a function of the false-positive fraction. A particular point on an OC curve indicates the number of true positives detected at a corresponding false-positive fraction. In addition to the OC curves, a binary object image was produced for each scene-detector combination indicating which pixels the respective detector labeled as anomalies. These object images can be visually compared to the ground-truth images for a qualitative assessment of detector performance. The OC curves for the three scenes are shown in Figure 3, and the object images for each scene are given in Figures 4 though 6.

Prior to applying the detectors to the image scenes, several preprocessing steps where implemented. First, a principal components analysis (PCA) was applied to each image to reduce the dimensionality the data to ten features. Though technically not required, this data reduction significantly reduces the computational time required for all detectors. For the RX detector, no other preprocessing was required. Because the BACON, FAST-MCD, and SDE-NTM, detectors are designed to operate on datasets from a single, homogeneous population, each image was clustered into k groups of homogenous pixels using the k-means clustering algorithm. Each detection method was then applied to each of the k clusters to find anomalies. The value of k for each image was set to the number of background materials in the image based on visual inspection. It should be noted that, for a given image, the same clustering was used for all algorithms to account for the stochastic nature of k-means clustering.

From the OC curves and object images, we note several significant observations. First, the BACON detector has the best OC performance of all the methods tested across the three image scenes. This superior performance is verified by visual inspection of the object images relative to the respective ground truth image. Though the BACON detector is not perfect—as evident by the complete omission of some objects of interest in Scene 2—it appears to be most effective at detecting man-made objects while minimizing the number of false alarms. A second observation is that clustering methods, in general, perform considerably better than the local RX detector when presented with complex scenes. Though the RX detector performed well against Scene 1, it was ineffective in detecting the larger objects in Scenes 2 and 3 while at the same time producing more false alarms. A third insight drawn from the detection results is the extremely high number of false alarms generated by the FAST-MCD and SDE-NTM methods, as shown in Figures 4 through 6. These object images would appear to disagree with the OC curves which seem to

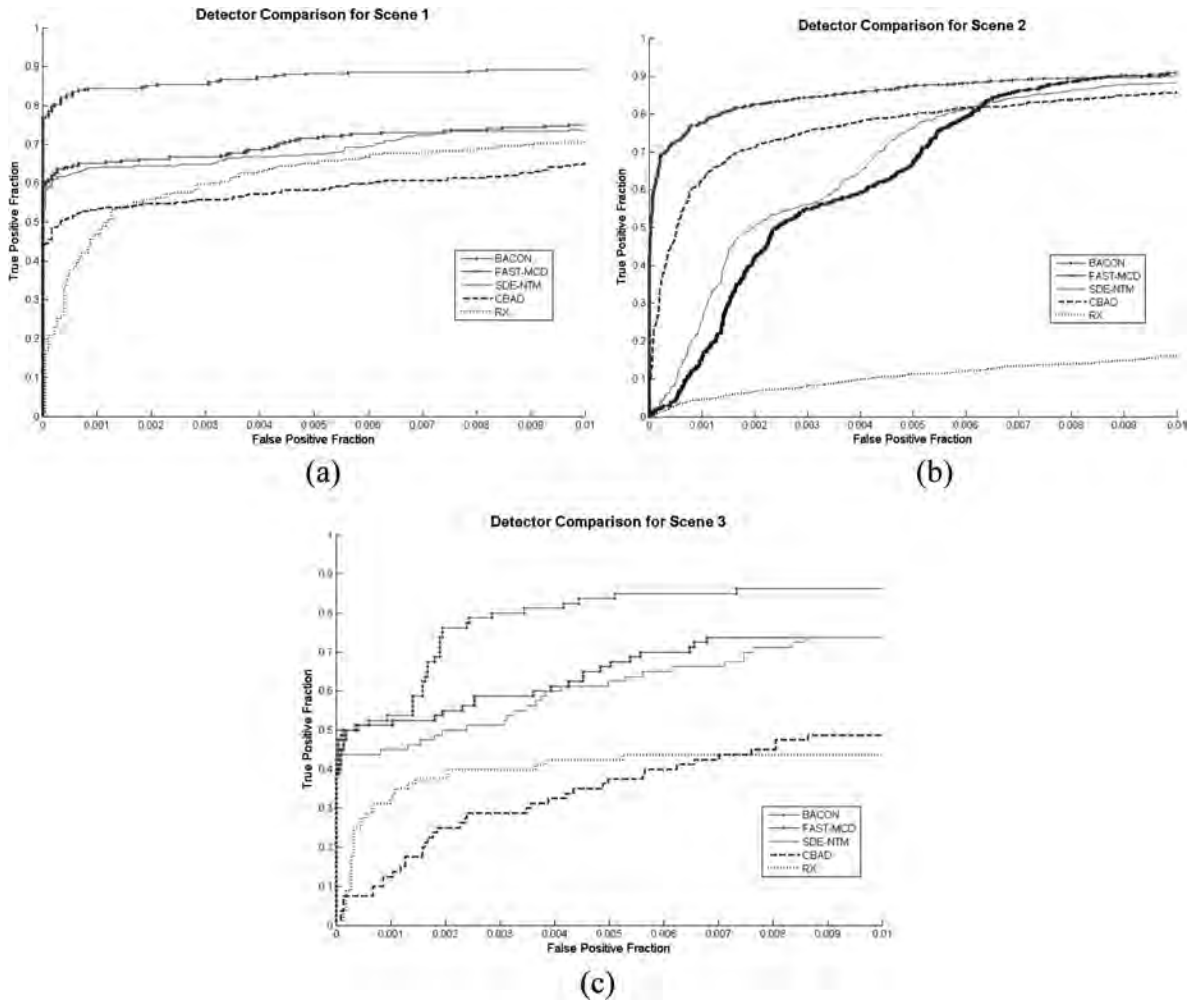

Figure 3. Operating Characteristic (OC) curves for (a) Scene 1, (2) Scene 2, and (3) Scene 3.

indicate that these two methods are superior to the CBAD and RX detectors. The reason for this apparent contradiction is that the significance level used by the detectors to generate the object images results in a false positive fraction for the FAST-MCD and SDE-NTM methods that is beyond the range shown in the OC curves. In these omitted regions of the OC curves, the CBAD detector is actually outperforming the FAST-MCD and SDE-NTM methods.

So why do the FAST-MCD and SDE-NTM methods produce such a large number of false alarms? For the same reason given in the simulated data tests, these detectors are underestimating the variance in the data, thereby increasing the false alarm rates. In other words, the scaling factors used with the shape matrix estimates for these two detectors are not sufficiently inflating the matrices to adequately represent the variance in the data.

To demonstrate the validity of this assertion, we used the scaling method of the BACON detector with the FAST-MCD and SDE-NTM detectors and applied these modified detectors to Scene 2. The resulting OC curves and object images are shown in Figures 7 and 8. Inspection of these figures reveals a marked improvement in the performance of the modified detectors that is almost identical to the performance of the BACON detector. In fact, modifying FAST-MCD and SDE-NTM in

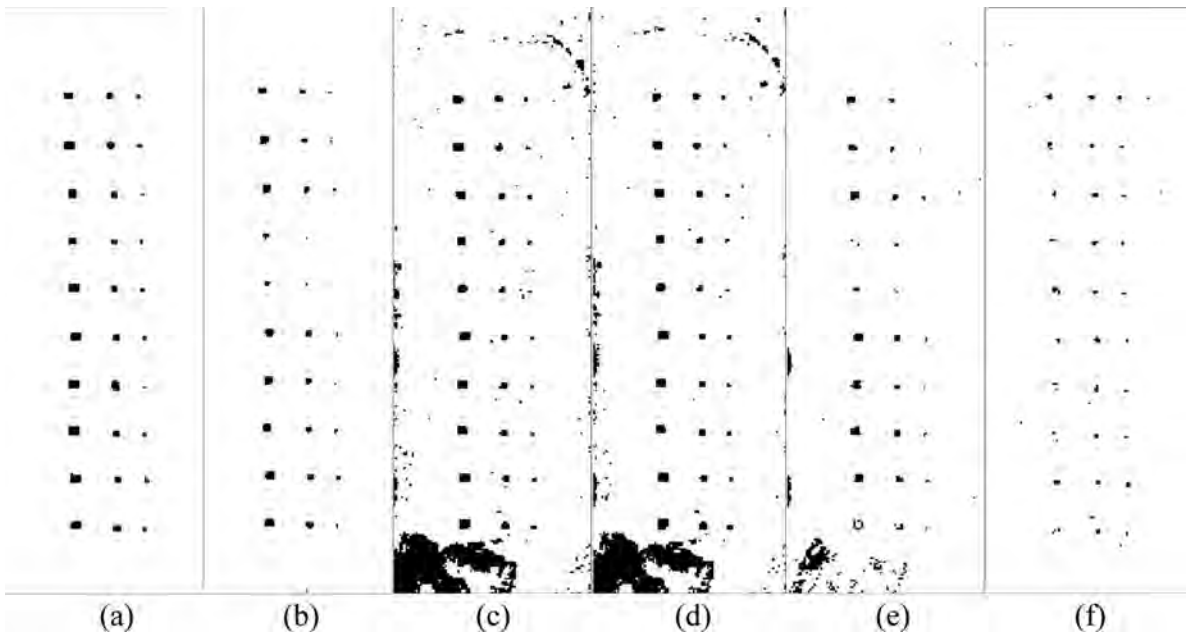

Figure 4. Object images for Scene 1. (a) Truth Image. (b) BACON Detector. (c) FAST-MCD Detector. (d) SDE-NTM Detector. (e) Cluster Detector. (f) RX Detector.

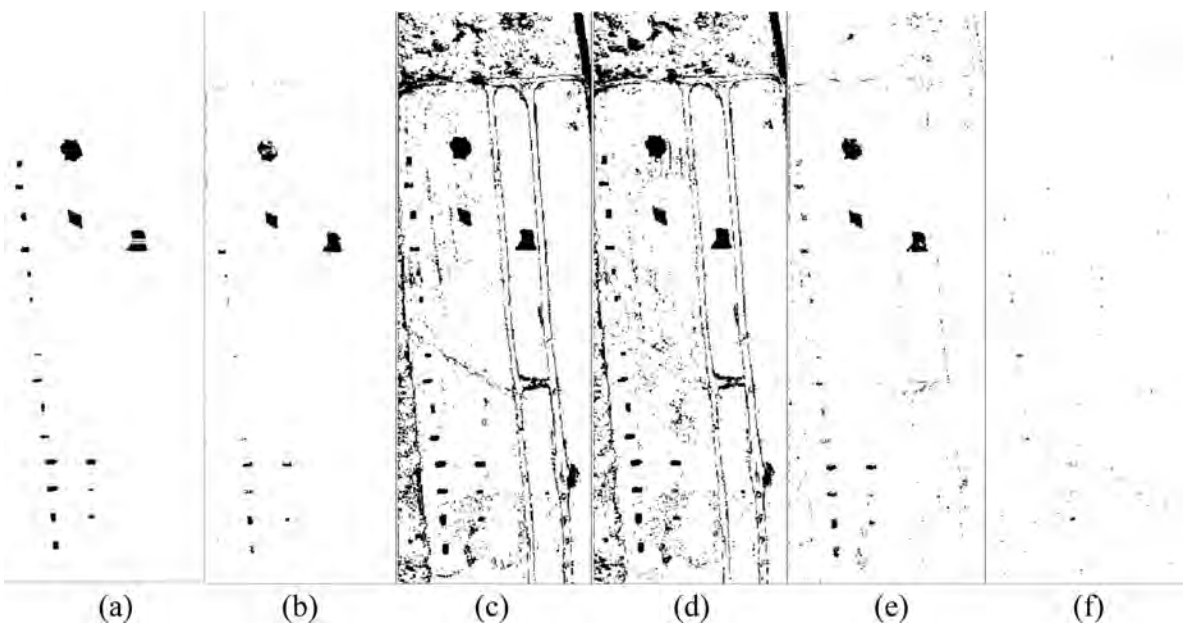

Figure 5. Object images for Scene 2. (a) Truth Image. (b) BACON Detector. (c) FAST-MCD Detector. (d) SDE-NTM Detector. (e) Cluster Detector. (f) RX Detector.

this manner essentially produces a BACON detector that uses either the FAST-MCD or SDE-NTM robust mean and covariance estimates to form the initial basic subset. Further research is required to determine if other scaling methods found in the literature also lead to better detection performance of the FAST-MCD and SDE-NTM methods.

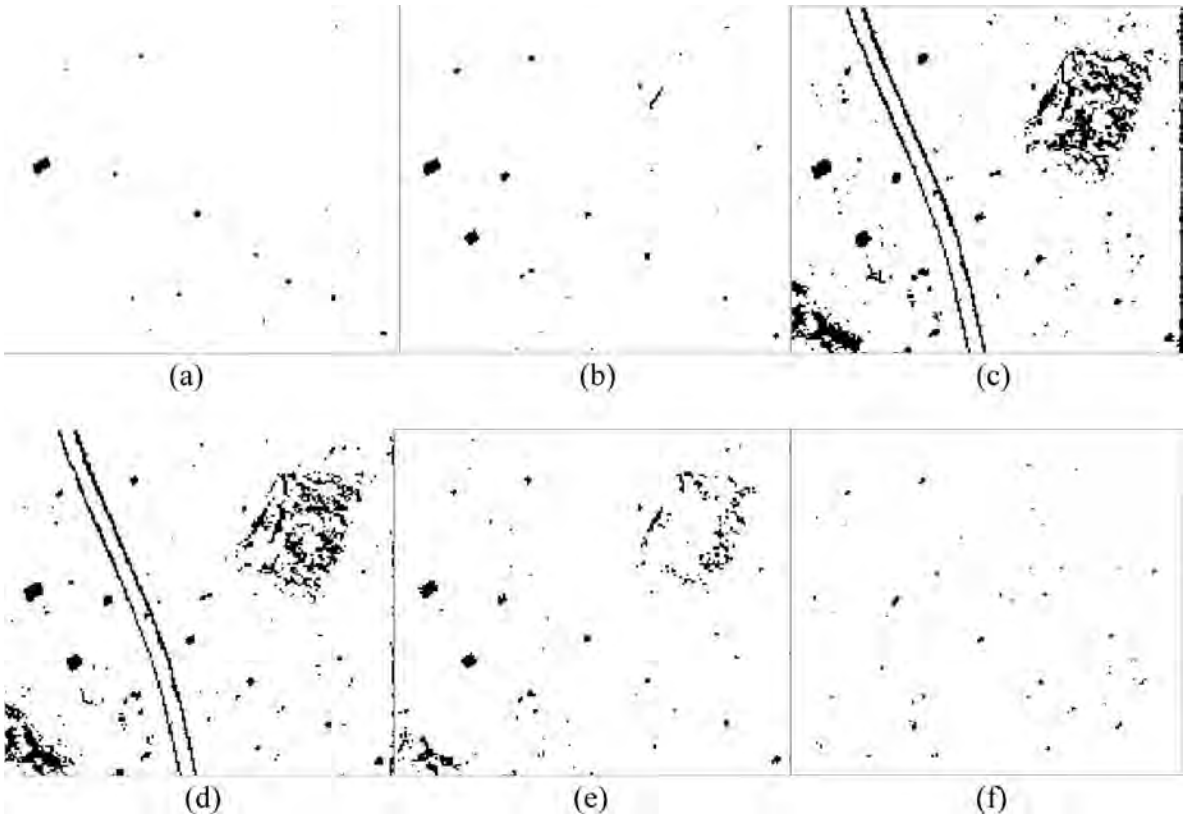

Figure 6. Object Images for Scene 3. (a) Truth Image. (b) BACON Detector. (c) FAST-MCD Detector. (d) SDE-NTM Detector. (e) Cluster Detector. (f) RX Detector.

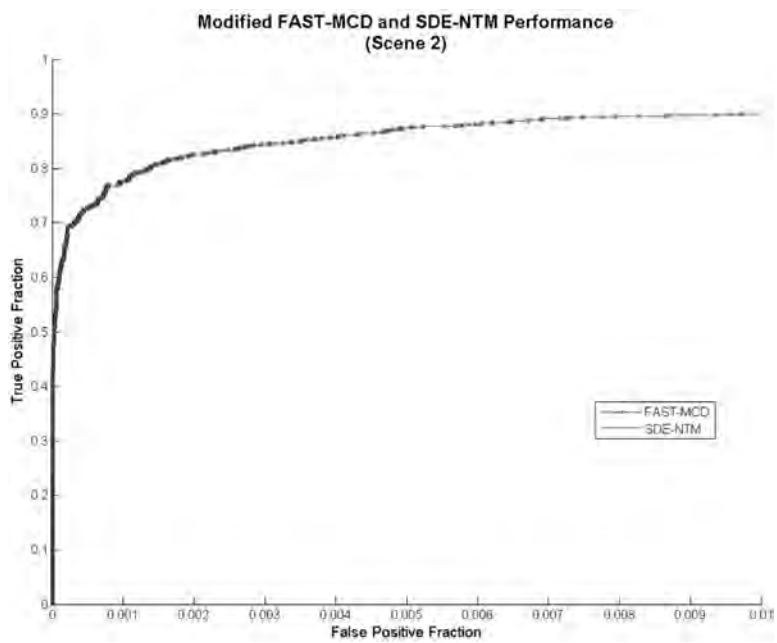

Figure 7. OC curve for modified FAST-MCD and SDE-NTM detectors.

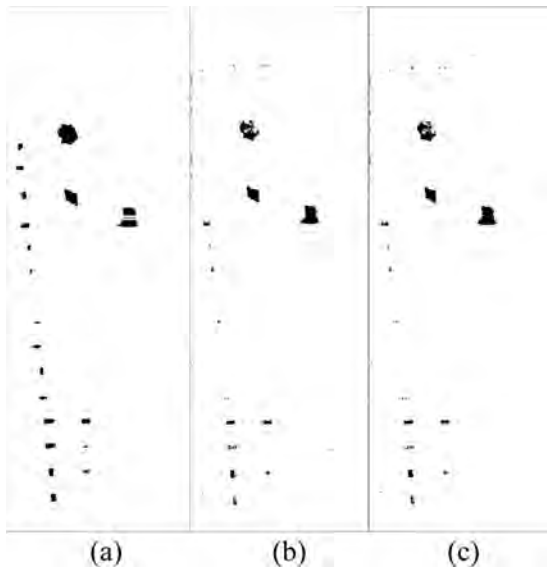

Figure 8. Object images for modified FAST-MCD and SDE-NTM detectors. (a) Truth Image. (b) Modified FAST-MCD. (c) Modified SDE-NTM.

CONCLUSIONS

The objective of this research is to compare the relative performance of different multivariate outlier detection methods when used to detect hyperspectral image anomalies. To this end, we applied the BACON, FAST-MCD, Juan-Prieto, and SDE-NTM detectors to simulated and actual anomaly detection problems. These tests indicated that multivariate outlier detection methods are superior to non-robust benchmark anomaly detectors in locating anomalies. In the case of the BACON detector, these anomalies were found while maintaining a low number of false alarms. The FAST-MCD and SDE-NTM methods, though effective in finding anomalies, require modification from their original forms to reduce the their respective false alarms rates to acceptable levels. In addition to these and other conclusions discussed in the previous sections, we noted the following:

- The BACON detector is the least computationally expensive outlier detection method tested. For the relatively large Scene 2 image, BACON detection took a matter of seconds to complete, while FAST-MCD detection took several minutes. SDE-NTM detection was the longest of the methods, taking on the order of tens of minutes to complete.
- All methods benefit from PCA data reduction in terms of improved computation speed. Though exploratory tests indicate that detection accuracy may improve if the original, full-dimensional dataset is used, the added computation time may be prohibitive.
- Of the methods tested, the BACON detector appears most practical for hyperspectral anomaly detection due to its superior accuracy, ease of implementation, and computational speed.

To extend this research in the pursuit of an accurate, operationally useful hyperspectral anomaly detector, we recommend the following areas for further study:

1) Investigate methods to automate the clustering process used to group the hyperspectral scene into homogeneous groups. For the present research, we simply guessed at the proper number of clusters based on visual inspection of the image. Automation of this process will result in an anomaly detector that is less dependent on user interaction.
2) Explore using the r_i values produced during the SDE-NTM detection process to reveal outliers. If the distribution of these values can be determined, a useful detection method may result.
3) Conduct a thorough investigation of the false-negatives resulting from the anomaly detection process to better understand why they elude detection. Such an investigation may suggest ways to improve detection algorithms or to better understand operational limits of such methods.
4) Implement a more rigorous algorithm testing methodology to better assess the performance of anomaly detectors against a broad range of imagery and operational conditions. Development of such a testing scheme is complicated by the expense of collecting hyperspectral imagery; however, the use of simulated hyperspectral imagery may alleviate this problem.

By exploring these research areas further, it is our goal to produce an anomaly detection scheme that is useful against a broad range of

image scenes and that requires minimal subjective input from the user. We feel such an algorithm is most beneficial in an operational military environment and will make the wealth of information provided by hyperspectral data more accessible to the military and civilian community.

DISCLAIMER

The views expressed in this article are those of the authors and do not reflect the official policy of the United States Air Force, Department of Defense, or the U.S. Government.

REFERENCES

Achard V., Landrevie A. and Fort J.C., 2004, Anomalies Detection in Hyperspectral Imagery Using Projection Pursuit Algorithm, SPIE Conference on Image and Signal Processing for Remote Sensing X, 5573:193–202

Atkinson Anthony C., 1994, Fast Very Robust Methods for the Detection of Multiple Outliers, Journal of the American Statistical Association, 89:1329–1339

Barnett Vic and Lewis Toby, 1994, Outliers in Statistical Data, 3rd Ed., John Wiley & Sons, Inc., Chichester, UK

Beckman R.J. and Cook R.D., 1983, Outlier. . . s, Technometrics, 25:119–163

Billor Nedret, Hadi Ali S. and Velleman Paul F., 2000, BACON: Blocked Adaptive Computationally Efficient Outlier Nominators, Computational Statistics & Data Analysis, 34:279–298

Campbell N.A., 1980, Robust Procedures in Multivariate Analysis I: Robust Covariance Estimation, Applied Statistics, 29:231–237

Carlotto Mark J., 2005, A Cluster-based Approach for Detecting Man-made Objects and Changes in Imagery, IEEE Transactions on Geoscience and Remote Sensing, 43:374–387

Catterall Stephen, 2004, Anomaly Detection Based on the Statistics of Hyperspectral Imagery, SPIE Conference on Imagery Spectroscopy X, 5546:171–178

Chang Chein-I, 2003, Hyperspectral Imaging: Techniques for Spectral Detection and Classification, Kluwer Academic/Plenum Publishers, New York

Chang Chein-I and Chiang Shao-Shan, 2001, Real-time Processing Algorithms for Target Detection and Classification in Hyperspectral Imagery, IEEE Transactions on Geoscience and Remote Sensing, 39:760–768

Chiang Leo H., Pell Randy J. and Seasholtz Mary Beth, 2003, Exploring Process Data with the Use of Robust Outlier Detection Algorithms, Journal of Process Control, 13: 437–449

Chiang Shao-Shan, Chang Chein-I and Ginsberg I.W., 2001, Unsupervised Target Detection in Hyperspectral Images Using Projection Pursuit, IEEE Transactions on Geoscience and Remote Sensing, 39:1380–1391

Clare Phil, Bernhardt Mark, Oxford William, Murphy Sean, Godfree Peter and Wilkinson Vicky, 2003, A New Approach to Anomaly Detection in Hyperspectral Images, SPIE Conference on Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery IX, 5093:17–28

Donoho D. L., 1982, Breakdown Properties of Multivariate Location Estimators, PhD Qualifying Paper, Harvard University, Cambridge, MA

Egan William J. and Morgan Stephen L., 1998, Outlier Detection in Multivariate Analytical Chemical Data, Analytical Chemistry, 70: 2372–2379

Fang Kai-Tai and Wang Y., 1994, Number Theoretic Methods in Statistics, Chapman and Hall, London

Gao Shaogen, Li Guoying and Wang Dongqian, 2005, A New Approach for Detecting Multivariate Outliers, Communications in Statistics-Theory and Methods, 34:1857–1865

Gaucel J.-M., Guillaume M. and Bourennane S., 2005, Whitening Spacial Correlation

Filtering for Hyperspectral Anomaly Detection, 2005 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '05), 1:333–336

Gnanadesikan R. and Kettenring J.R., 1972, Robust Estimates, Residuals, and Outlier Detection with Multiresponse Data, Biometrics, 28:81–124

Goovaerts Pierre, Jacquez Geoffrey, Warner Amanda, Crabtree Bob and Marcus Andrew, 2004, Detection of Local Anomalies in High Resolution Hyperspectral Imagery Using Geostatistical Filtering and Local Spatial Statistics, 2003 IEEE Workshop on Advances in Techniques for Analysis of Remotely Sensed Data, 1: 385–394

Grossman John M., Bowles Jeffrey, Haas Daniel, Antoniades John A., Grunes Mitchell R., Palmadesso Peter, Gillis David, Tsang Kwok Y., Baumbeck Mark, Daniel Mark, Fisher John and Triandaf Ioana, 1998, Hyperspectral Analysis and Target Detection System for the Adaptive Spectral Reconnaissance Program, SPIE Conference on Algorithms for Multispectral and Hyperspectral Imagery, IV, 3372:2–13

Hadi Ali S., 1992, Identifying Multiple Outliers in Multivariate Data, Journal of the Royal Statistical Society, Series B, 54:761–771

Hadi Ali S., 1994, A Modification of a Method for the Detection of Outliers in Multivariate Samples, Journal of the Royal Statistical Society, Series B, 56:393–396

Hardin Johanna and Rocke David M., 2004, Outlier Detection in the Multiple Cluster Setting Using the Minimum Covariance Determinant Estimator, Computational Statistics & Data Analysis, 44:625–638

Hawkins Douglas M., 1994, The Feasible Solution Algorithm for the Minimum Covariance Determinant Estimator in Multivariate Data, Computational Statistics & Data Analysis, 17:197–210

Hazel Geoffrey, 2000, Multivariate Gaussian MRF for Multispectral Scene Segmentation and Anomaly Detection, IEEE Transactions on Geoscience and Remote Sensing, 38: 1199–1211

Hsueh Mingkai and Chang Chein-I, 2004, Adaptive Causal Anomaly Detection for Hyperspectral Imagery, 2004 IEEE International Geoscience and Remote Sensing Symposium (IGARSS '04), 5:3222–3224

Juan Jesus and Prieto Francisco J., 2001, Using Angles to Identify Concentrated Multivariate Outliers, Journal of the American Statistical Association, 43:311–322

Kerekes John P. and Manolakis Dimitris, 2004, Improved Modeling of Background Distributions in an End-to-End Spectral Imaging System Model, Proceedings of the 2004 IEEE International Geoscience and Remote Sensing Symposium, 2:972–975

Kim Myung Geun, 2000, Multivariate Outliers and Decompositions of Mahalanobis Distance, Communications in Statistics-Theory and Methods, 29:1511–1526

Kwon Heesung, Der S.Z. and Nasrabadi Nasser M., 2003, Adaptive Anomaly Detection Using Subspace Separation for Hyperspectral Imagery, Optical Engineering, 42:3342–3351

Kwon Heesung and Nasrabadi Nasser M., 2005, Kernel RX-Algorithm: A Nonlinear Anomaly Detector for Hyperspectral Imagery, IEEE Transactions on Geoscience and Remote Sensing, 43:388–397

Landgrebe David A., 2002, Hyperspectral Image Data Analysis, IEEE Signal Processing Magazine, 19:17–28

Landgrebe David A., 2003, Signal Theory Methods in Multispectral Remote Sensing, John Wiley & Sons, Inc., Hoboken, New Jersey

Liu Weimin and Chang Chein-I, 2004, A Nested Spatial Window-Based Approach to Target Detection for Hyperspectral Imagery, 2004 IEEE International Geoscience and Remote Sensing Symposium (IGARSS '04), 1:266–268

Manolakis D. and Marden D., 2002, Non Gaussian Models for Hyperspectral Algorithm Design and Assessment, IEEE International Geoscience and Remote Sensing Symposium, 2002 (IGARSS '02), 1:1664–1666

Manolakis D., Rossacci M., Cipar J., Lockwood R., Cooley T. and Jacobson J., 2005, Statistical Characterization of Natural Hyperspectral Backgrounds Using t-Elliptically Contoured Distributions, SPIE Conference on Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XI, 5806:56–65

Maronna Ricardo A. and Yohai Victor J., 1995, The Behavior of the Stahel-Donoho Robust Multivariate Estimator, Journal of the American Statistical Association, 90:330–341

Pan Jian-Xin, Fung Wing-Kam and Fang Kai-Tai, 2000, Multiple Outlier Detection in Multivariate Data Using Projection Pursuit Techniques, Journal of Statistical Planning and Inference, 83:153–167

Reed Irving S. and Yu Xiaoli, 1990, Adaptive Multiple-Band CFAR Detection of an Optical Pattern with Unknown Spectral Distribution, IEEE Proceedings on Acoustics, Speech, and Signal Processing, 38:1760–1770

Richards John A. and Jia Xiuping, 1999, Remote Sensing Digital Image Analysis: An Introduction, 3rd Ed., Springer-Verlag, Berlin

Riley Ronald, Newsom Rob K. and Andrews Aaron K, 2004, Anomaly Detection in Noisy Hyperspectral Imagery, SPIE Conference on Imaging Spectrometry X, 5546:159–170

Rocke David M. and Woodruff David L., 1996, Identification of Outliers in Multivariate Data, Journal of the American Statistical Association, 91:1047–1061

Rosario Dalton S., 2004, Highly Effective Logistic Regression Model for Signal (Anomaly) Detection, IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '04), 1:817–820

Rousseeuw Peter J., 1983, Multivariate Estimation with High Breakdown Point, Fourth Pannonian Symposium on Mathematical Statistics and Probability,

Rousseeuw Peter J. and Leroy Annick M., 1987, Robust Regression and Outlier Detection, John Wiley & Sons, Inc., New York

Rousseeuw Peter J. and van Driessen Katrien, 1999, A Fast Algorithm for the Minimum Covariance Determinant Estimator, Technometrics, 41:212–223

Rousseeuw Peter J. and van Zomeren Bert C., 1990, Unmasking Multivariate Outliers and Leverage Points, Journal of the American Statistical Association, 85:633–639

Schaum Alan P., 2004, Joint Subspace Detection of Hyperspectral Targets, Proceedings of the 2004 IEEE Aerospace Conference, 3:1818–1824

Schaum Alan P., 2006, A Remedy for Nonstationarity in Background Transition Regions for Real Time Hyperspectral Detection, Proceedings of the 2006 IEEE Aerospace Conference,

Schaum Alan P. and Stocker Alan D., 1997, The Stochastic Mixing Model, 1997 International Symposium on Spectral Sensing Research,

Schweizer Susan M. and Moura Jose M.F., 2000, Hyperspectral Imagery: Clutter Adaptation in Anomaly Detection, IEEE Transactions on Information Theory, 46:1855–1871

Schweizer Susan M. and Moura Jose M.F., 2001, Efficient Detection in Hyperspectral Imagery, IEEE Transactions on Image Processing, 10:584–597

Smetek Timothy E. and Bauer Kenneth W., 2007, Finding Hyperspectral Anomalies Using Multivariate Outlier Detection, Proceedings of the 2007 IEEE Aerospace Conference (To Be Published),

Stahel W.A., 1981, Robuste Schatzungen: Infinitesimale Optimalitat und Schatzungen von Kovarianzmatrizen, PhD Thesis, Zurich, Switzerland

Stein David W.J., Beaven Scott G., Hoff Lawrence E., Winter Edwin M., Schaum Alan P. and Stocker Alan D., 2002, Anomaly Detection for Hyperspectral Imagery, IEEE Signal Processing Magazine, 19:58–69

Viljoen H. and Venter J.H., 2002, Identifying Multivariate Discordant Observations: A Computer-Intensive Approach, Computational Statistics & Data Analysis, 40:159–172

West Jason E., Messinger David W., Ientilucci Emmett J., Kerekes John P. and Schott John R., 2005, Matched Filter Stochastic Background Characterization for Hyperspectral Target Detection, SPIE Conference on Algorithms and Technologies for Multispectral, Hyperspectral, and Ultraspectral Imagery XI, 5806:1–12

DESCRIPTOR LIST

Combat Identification, Applied Statistics, Hyperspectral Image Analysis, Anomaly Detection, Outlier Detection, Multivariate Data Analysis.

APPENDIX

This appendix contains the results from the simulated data tests. Tables A-1 and A-2 summarize the mean number of true-positives detected using Gaussian and Multivariate *t*-distributed data, respectively. Similarly, Tables A-3 and A-4 summarize the mean number of false-positives detected using Gaussian and Multivariate *t*-distributed data, respectively.

Table A-1. Mean true-positives detected as a function of contamination level for Gaussian Data.

Background/ Outlier	Number Of Outliers	True-Positives by Method									
		Classical		BACON		F.MCD		Juan-Prieto		SDE-NTM	
		Mean	S.E.	Mean	S.E.	Mean	S.E.	Mean	S.E.	Mean	S.E.
Grass/Road (A.P. Hill)	50	46.3	2.1	50.0	0.0	50.0	0.0	0.0	0.0	50.0	0.0
	100	12.1	2.8	100.0	0.0	100.0	0.0	0.0	0.0	100.0	0.0
	300	2.9	1.5	300.0	0.0	300.0	0.0	286.4	6.8	300.0	0.0
	500	1.5	1.1	500.0	0.0	500.0	0.0	481.0	13.9	500.0	0.0
Grass/ Shadow (A.P. Hill)	50	45.0	2.3	50.0	0.0	50.0	0.0	0.0	0.0	50.0	0.0
	100	14.2	3.0	100.0	0.0	100.0	0.0	0.0	0.0	100.0	0.0
	300	4.4	1.8	300.0	0.0	300.0	0.0	283.9	10.5	300.0	0.0
	500	2.2	1.1	500.0	0.0	500.0	0.0	468.6	17.1	500.0	0.0
Dead Grass/ Shadow (A.P. Hill)	50	45.3	2.0	50.0	0.0	50.0	0.0	0.0	0.0	50.0	0.0
	100	15.7	2.7	100.0	0.0	100.0	0.0	0.0	0.0	100.0	0.0
	300	4.8	1.4	300.0	0.0	300.0	0.0	284.1	10.4	300.0	0.0
	500	2.4	1.8	500.0	0.0	500.0	0.0	471.4	15.7	500.0	0.0
Road/ Shadow (A.P. Hill)	50	49.2	0.9	50.0	0.0	50.0	0.0	0.0	0.0	50.0	0.0
	100	72.9	3.2	100.0	0.0	100.0	0.0	0.0	0.0	100.0	0.0
	300	68.4	5.3	300.0	0.0	300.0	0.0	7.5	41.1	300.0	0.0
	500	43.4	5.4	500.0	0.0	500.0	0.0	21.9	84.5	500.0	0.0
Grass/ Asphalt (D.C. Mall)	50	11.7	3.1	50.0	0.0	50.0	0.0	0.0	0.0	50.0	0.0
	100	0.0	0.2	100.0	0.0	100.0	0.0	70.0	43.1	100.0	0.0
	300	0.0	0.0	300.0	0.0	300.0	0.0	297.5	2.6	300.0	0.0
	500	0.1	0.3	500.0	0.0	500.0	0.0	491.8	5.3	500.0	0.0
Grass/Water (D.C. Mall)	50	0.0	0.0	50.0	0.0	50.0	0.0	0.0	0.0	50.0	0.0
	100	0.0	0.0	100.0	0.0	100.0	0.0	86.7	34.6	100.0	0.0
	300	0.0	0.0	300.0	0.0	300.0	0.0	300.0	0.0	300.0	0.0
	500	0.0	0.0	500.0	0.0	500.0	0.0	500.0	0.0	500.0	0.0
Asphalt/ Water (D.C. Mall)	50	27.1	2.7	50.0	0.0	50.0	0.0	0.0	0.0	50.0	0.0
	100	0.0	0.0	100.0	0.0	100.0	0.0	88.9	30.2	100.0	0.0
	300	0.0	0.0	300.0	0.0	300.0	0.0	299.2	1.1	300.0	0.0
	500	0.0	0.0	500.0	0.0	500.0	0.0	496.8	2.8	500.0	0.0
Gravel/ Asphalt (D.C. Mall)	50	20.5	2.6	50.0	0.0	50.0	0.0	0.0	0.0	50.0	0.0
	100	0.0	0.0	100.0	0.0	100.0	0.0	89.2	30.3	100.0	0.0
	300	0.0	0.0	300.0	0.0	300.0	0.0	299.4	0.9	300.0	0.0
	500	0.0	0.0	500.0	0.0	500.0	0.0	497.8	1.8	500.0	0.0

Table A-2. Mean true-positives detected as a function of contamination level for Multivariate *t*-Data.

Background/ Outlier	Outliers Present	True-Positives by Method									
		Classical		BACON		F.MCD		Juan-Prieto		SDE-NTM	
		Mean	S.E.	Mean	S.E.	Mean	S.E.	Mean	S.E.	Mean	S.E.
Grass/Road (A.P. Hill)	50	45.6	1.9	50.0	0.0	50.0	0.0	0.0	0.0	50.0	0.0
	100	12.4	2.3	100.0	0.0	100.0	0.0	4.9	18.9	100.0	0.0
	300	5.8	1.6	300.0	0.0	300.0	0.0	286.4	5.8	300.0	0.0
	500	4.1	2.0	500.0	0.0	500.0	0.0	475.2	13.3	500.0	0.0
Grass/ Shadow (A.P. Hill)	50	44.2	1.9	50.0	0.0	50.0	0.0	0.0	0.0	50.0	0.0
	100	15.2	2.8	100.0	0.0	100.0	0.0	3.2	12.5	100.0	0.0
	300	8.7	2.8	300.0	0.0	300.0	0.0	278.5	11.8	300.0	0.0
	500	8.2	2.4	500.0	0.0	500.0	0.0	465.0	23.4	500.0	0.0
Dead Grass/ Shadow (A.P. Hill)	50	44.3	2.5	50.0	0.0	50.0	0.0	0.0	0.0	50.0	0.0
	100	16.8	3.3	100.0	0.0	100.0	0.0	0.0	0.0	100.0	0.0
	300	8.7	2.4	300.0	0.0	300.0	0.0	282.4	8.7	300.0	0.0
	500	7.6	2.5	500.0	0.0	500.0	0.0	468.6	17.4	500.0	0.0
Road/ Shadow (A.P. Hill)	50	48.4	1.0	50.0	0.0	50.0	0.0	0.0	0.0	50.0	0.0
	100	63.9	4.1	100.0	0.0	100.0	0.0	0.0	0.0	100.0	0.0
	300	67.6	5.8	300.0	0.0	300.0	0.0	32.2	83.8	300.0	0.0
	500	59.3	4.0	500.0	0.2	500.0	0.0	85.8	174.8	500.0	0.0
Grass/ Asphalt (D.C. Mall)	50	9.7	2.2	50.0	0.0	50.0	0.0	0.0	0.0	50.0	0.0
	100	0.1	0.3	100.0	0.0	100.0	0.0	71.0	43.6	100.0	0.0
	300	0.2	0.6	300.0	0.0	300.0	0.0	295.9	3.2	300.0	0.0
	500	0.2	0.4	500.0	0.0	500.0	0.0	487.4	7.9	500.0	0.0
Grass/Water (D.C. Mall)	50	0.0	0.0	50.0	0.0	50.0	0.0	0.0	0.0	50.0	0.0
	100	0.0	0.0	100.0	0.0	100.0	0.0	90.0	30.5	100.0	0.0
	300	0.0	0.0	300.0	0.0	300.0	0.0	300.0	0.0	300.0	0.0
	500	0.0	0.0	500.0	0.0	500.0	0.0	500.0	0.2	500.0	0.0
Asphalt/ Water (D.C. Mall)	50	23.4	2.9	50.0	0.0	50.0	0.0	0.0	0.0	50.0	0.0
	100	0.1	0.4	100.0	0.0	100.0	0.0	88.6	30.1	100.0	0.0
	300	0.0	0.0	300.0	0.0	300.0	0.0	298.6	1.1	300.0	0.0
	500	0.0	0.2	500.0	0.0	500.0	0.0	493.4	3.6	500.0	0.0
Gravel/ Asphalt (D.C. Mall)	50	18.8	2.2	50.0	0.0	50.0	0.0	0.0	0.0	50.0	0.0
	100	0.0	0.0	100.0	0.0	100.0	0.0	82.2	37.4	100.0	0.0
	300	0.0	0.0	300.0	0.0	300.0	0.0	298.7	1.9	300.0	0.0
	500	0.0	0.0	500.0	0.0	500.0	0.0	495.9	3.1	500.0	0.0

Table A-3. Mean false-positives detected as a function of contamination level for Gaussian Data.

Background/ Outlier	Outliers Present	False-Positives by Method									
		Classical		BACON		F.MCD		Juan-Prieto		SDE-NTM	
		Mean	S.E.	Mean	S.E.	Mean	S.E.	Mean	S.E.	Mean	S.E.
Grass/Road (A.P. Hill)	0	0.0	0.0	0.0	0.0	0.1	0.3	0.0	0.0	0.0	0.2
	50	0.1	0.3	0.0	0.0	0.1	0.4	0.0	0.0	0.1	0.3
	100	0.0	0.0	0.0	0.0	0.1	0.3	0.0	0.0	0.0	0.2
	300	0.0	0.2	0.0	0.0	0.1	0.3	0.0	0.0	0.0	0.0
	500	0.2	0.5	0.0	0.0	0.1	0.3	0.0	0.0	0.0	0.0
Grass/ Shadow (A.P. Hill)	0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.2
	50	0.0	0.0	0.0	0.0	0.2	0.4	0.0	0.0	0.1	0.3
	100	0.0	0.2	0.0	0.0	0.1	0.3	0.0	0.0	0.1	0.3
	300	0.0	0.0	0.0	0.0	0.1	0.3	0.0	0.0	0.0	0.0
	500	0.1	0.3	0.0	0.0	0.0	0.2	0.7	1.9	0.0	0.0
Dead Grass/ Shadow (A.P. Hill)	0	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.0
	50	0.0	0.0	0.0	0.0	0.1	0.3	0.0	0.0	0.0	0.2
	100	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.2
	300	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
	500	0.0	0.2	0.0	0.0	0.1	0.3	1.3	6.6	0.0	0.0
Road/ Shadow (A.P. Hill)	0	0.0	0.2	0.0	0.0	0.1	0.3	0.0	0.0	0.1	0.3
	50	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.2
	100	0.0	0.0	0.0	0.0	0.2	0.5	0.0	0.0	0.1	0.3
	300	0.0	0.0	0.0	0.0	0.1	0.3	1.5	4.8	0.0	0.2
	500	0.1	0.3	0.0	0.0	0.2	0.6	8.8	25.9	0.0	0.2
Grass/ Asphalt (D.C. Mall)	0	0.1	0.3	0.0	0.0	0.2	0.4	0.0	0.0	0.1	0.3
	50	0.0	0.0	0.0	0.0	0.1	0.3	0.0	0.0	0.0	0.2
	100	0.0	0.0	0.0	0.0	0.1	0.3	2.6	5.7	0.0	0.0
	300	0.2	0.5	0.0	0.0	0.1	0.3	0.0	0.0	0.0	0.2
	500	0.2	0.6	0.0	0.0	0.1	0.3	0.0	0.0	0.0	0.0
Grass/Water (D.C. Mall)	0	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.0
	50	0.0	0.2	0.0	0.0	0.1	0.3	0.0	0.0	0.0	0.0
	100	0.0	0.0	0.0	0.0	0.1	0.3	0.0	0.0	0.0	0.0
	300	0.3	0.5	0.0	0.0	0.1	0.3	0.0	0.0	0.0	0.0
	500	0.6	0.8	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.0
Asphalt/ Water (D.C. Mall)	0	0.1	0.3	0.0	0.0	0.2	0.4	0.0	0.0	0.1	0.3
	50	0.0	0.0	0.0	0.0	0.1	0.4	0.0	0.0	0.0	0.0
	100	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.2
	300	0.1	0.3	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.2
	500	0.1	0.3	0.0	0.0	0.1	0.3	0.0	0.0	0.0	0.0
Gravel/ Asphalt (D.C. Mall)	0	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.2
	50	0.0	0.0	0.0	0.0	0.0	0.2	0.0	0.0	0.0	0.2
	100	0.1	0.3	0.0	0.0	0.1	0.3	0.0	0.2	0.1	0.3
	300	0.2	0.4	0.0	0.0	0.1	0.3	0.0	0.0	0.0	0.0
	500	0.2	0.4	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Table A-4. Mean false-positives detected as a function of contamination level for Multivariate *t*-Data.

Background/ Outlier	Outliers Present	False-Positives by Method									
		Classical		BACON		F.MCD		Juan-Prieto		SDE-NTM	
		Mean	S.E.	Mean	S.E.	Mean	S.E.	Mean	S.E.	Mean	S.E.
Grass/Road (A.P. Hill)	0	14.8	3.5	0.0	0.0	39.4	6.5	0.0	0.0	27.9	4.6
	50	10.6	2.8	0.0	0.0	38.7	7.3	0.0	0.0	25.2	5.3
	100	10.9	3.5	0.1	0.3	37.1	6.7	1.0	3.1	23.9	6.2
	300	11.1	3.0	0.0	0.2	30.4	5.9	0.0	0.0	14.6	3.4
	500	14.1	3.3	0.0	0.0	28.5	5.3	0.0	0.0	11.6	3.9
Grass/ Shadow (A.P. Hill)	0	14.6	3.4	0.0	0.2	40.5	7.2	0.0	0.0	26.8	5.1
	50	10.0	2.9	0.0	0.2	39.8	7.1	0.0	0.0	26.3	5.5
	100	10.6	2.8	0.0	0.0	38.8	6.2	0.9	4.6	24.5	4.8
	300	10.5	2.9	0.1	0.3	30.1	6.3	0.0	0.0	15.8	4.5
	500	10.1	3.5	0.0	0.0	26.5	6.1	0.0	0.0	10.0	3.9
Dead Grass/ Shadow (A.P. Hill)	0	14.4	3.4	0.0	0.0	41.3	8.0	0.0	0.0	28.0	6.3
	50	9.9	3.1	0.0	0.0	38.9	6.9	0.0	0.0	25.8	5.4
	100	10.0	3.2	0.0	0.0	36.5	6.6	0.1	0.4	23.2	6.0
	300	10.0	3.2	0.0	0.2	31.2	6.0	0.0	0.0	16.3	3.8
	500	11.1	3.4	0.0	0.2	26.0	7.1	0.0	0.0	10.8	3.7
Road/ Shadow (A.P. Hill)	0	14.2	3.1	0.0	0.2	41.6	6.0	0.0	0.0	27.9	5.6
	50	6.1	1.9	0.0	0.2	38.8	6.4	0.0	0.0	25.6	5.6
	100	5.5	2.0	0.0	0.0	40.5	5.4	0.0	0.0	25.0	5.0
	300	4.3	1.7	0.0	0.0	31.9	6.3	2.8	7.4	16.5	4.4
	500	3.8	2.5	0.0	0.0	27.6	5.9	21.6	62.6	11.4	3.3
Grass/ Asphalt (D.C. Mall)	0	13.4	2.8	0.1	0.3	39.7	6.9	0.0	0.0	27.3	4.5
	50	11.9	2.9	0.0	0.0	39.4	6.2	0.0	0.0	25.8	4.5
	100	11.0	2.8	0.0	0.0	36.1	6.5	1.1	3.2	23.0	4.8
	300	16.0	3.6	0.0	0.2	33.1	4.6	0.0	0.0	17.0	4.0
	500	19.5	3.4	0.0	0.2	26.1	4.3	0.0	0.0	10.5	3.8
Grass/Water (D.C. Mall)	0	14.2	3.1	0.0	0.0	37.2	6.0	0.0	0.0	25.7	4.6
	50	10.7	2.9	0.0	0.0	36.6	6.1	0.0	0.0	23.7	4.3
	100	12.9	3.0	0.0	0.0	35.4	5.5	0.0	0.0	22.8	4.8
	300	17.4	2.3	0.0	0.0	31.0	5.3	0.0	0.0	15.9	3.1
	500	25.2	5.2	0.0	0.0	27.3	6.8	0.0	0.0	10.3	4.2
Asphalt/ Water (D.C. Mall)	0	14.0	3.3	0.0	0.2	40.3	5.6	0.0	0.0	26.9	4.8
	50	11.4	2.7	0.0	0.2	38.2	6.2	0.0	0.0	25.6	4.6
	100	13.0	3.1	0.0	0.0	35.5	5.9	0.1	0.5	23.5	4.5
	300	17.5	3.6	0.0	0.2	33.2	4.9	0.0	0.0	17.8	3.8
	500	20.1	4.4	0.1	0.3	26.0	5.1	0.0	0.0	11.3	3.9
Gravel/ Asphalt (D.C. Mall)	0	14.5	3.3	0.1	0.3	39.3	5.6	0.0	0.0	27.2	4.8
	50	10.7	3.0	0.0	0.0	38.3	6.1	0.0	0.0	24.7	5.4
	100	11.8	3.3	0.0	0.0	34.6	6.4	0.1	0.6	22.2	6.0
	300	16.1	3.2	0.0	0.2	32.1	6.8	0.0	0.0	16.5	4.4
	500	21.0	3.5	0.0	0.2	26.7	5.1	0.0	0.0	10.4	3.5

Publications Order Form

1703 North Beauregard Street, Suite 450, Alexandria, VA 22311-1745
703-933-9070 – Fax: 703-933-9066 – www.mors.org

20% Member Discount on MORS books and MORS Polo Shirts!

Name:		Date:
SHIPPING address preference: ❒ **Home** ❒ **Work**	Organization:	
Home address:	Address:	
City – State – Zip	City – State – Zip	
(A phone number must be entered or the order will not be processed. Thank you)	Phone:	Email:

***The 20% Member discount does not apply to the *PHALANX*, *MOR* Journal and back issues.**

UNITED STATES ADDRESSES		INTERNATIONAL ADDRESSES		TOTAL
**PHALANX*	*(1 year) – $ 40 - (2 years) – $ 70*	*PHALANX*	*(1 year) – $ 75 - 2 years – $ 140*	$
**MOR* Journal	*(1 year) – $ 70 - (2 years) – $ 130*	*MOR* Journal	*(1 year) – $ 135 - 2 years – $ 260*	$

*Back issues of *PHALANX* ($15(12) hard copy $10(8) electronic copy) and *MOR* ($25(20) hard copy and $20(16) electronic copy) are available. Prices are per copy, freight included—call 703-933-9070 for availability. — Subscriptions are Non-taxable

Please order back issues here:	*PHALANX* ❒	*MOR* ❒	Vol		No.		PDF ❒	Paper ❒	$

DESCRIPTION	PRICE	QTY	Member Disc 20%	TOTAL
Analysis for Military Decisions, E.S. Quade	$ 22		$ 17.60	
Mathematical Models of Target Coverage and Missile Allocation, Eckler & Burr	$ 21		$ 16.80	
Methods of Operations Research, Morse and Kimball (eds.)	$ 25		$ 20.00	
Military Modeling for Decision Making, 3rd Edition, Wayne P. Hughes (ed.)	$ 40		$ 32.00	
✪ ***Methods for Conducting Military Operational Analysis,*** Loerch and Rainey (eds.)	$ 75		$ 60.00	
Operational Research in the RAF, Air Ministry	$ 23		$ 18.40	
Search and Screening, Bernard Koopman	$ 24		$ 19.20	
Warfare Modeling, Bracken, Kress and Rosenthal (eds.)	$ 35		$ 28.00	
MORS Member Polo Shirt ***(Freight included)*** — ***SM–MED–LG–XL–2XL–3XL*** — ***SIZE(S):***	$ 35		$ 28.00	
OTHER: *(Please list)*	$		$	
FREIGHT to US Addresses - **$10 PER** book ordered FREIGHT to all other address - **$15.00 PER** book ordered			**$10 x QTY** **$15 x QTY**	$
✪ ***Please note:*** International freight charges for *Methods for Conducting Military Operational Analysis,* Loerch and Rainey (eds.) is **$30.00 per book**			**TOTAL SALES**	$
			(Sold in Virginia) 5% TAX	$
❒ **CASH** ❒ **CHECK #_____** ❒ **VISA** ❒ **MC** ❒ **AMEX**			**GRAND TOTAL**	$

Credit Card #:	**Expiration date:**
Name on Credit Card (print):	**Phone:**
If billing address is the same as shipping address check. ❒ **Billing address:**	**Zip code:**
Signature: X **Date:**	**Date mailed/ Picked up:** — **SOLD BY:**

11/10/2008

ABSTRACT

We consider two types of non-reactive aerial sensors, which are subject to false-positive and false-negative errors. The sensors search for threat objects such as ballistic missile launchers or improvised explosive devices. The objects are located in a certain area of interest, which is divided into a grid of area-cells. The grid is defined such that each area-cell may contain at most one object. The objective of a sensor is to *determine* if a certain area-cell is likely or unlikely to contain an object. An area-cell is said to be *determined* if the searcher can ascertain with a given high probability these events. Since definitive identification of a threat object, and subsequent handling of that threat, are done by limited number of available ground combat units, the *determination* of an area-cell can help field commanders better allocate and direct these scarce resources. We develop two models, one for each type of sensor, that describe the search process and maximize the expected number of *determined* area-cells.

INTRODUCTION

Advents in sensing, unmanned aerial vehicles (UAVs), and satellite technologies are expected to increase the military use of aerial or space sensors for detecting threat objects such as improvised explosive devices or missile launchers. These advanced technologies may generate powerful and effective sensors, which necessitate operational concepts in order to facilitate their efficient utilization. In this paper we address operational concepts associated with employing sensors in persistent search missions over an extended search area. Specifically, we consider the problem of efficiently allocating *non-reactive* sensors across a search area of interest. The sensors are non-reactive in the sense that the search plan is set in advance, and it is not updated in real time during the search process following new information (e.g., pre-programmed "send-and-forget" UAVs). The details of the operational search setting are given in Section 2.

The theory of optimal search has a history of principal importance in military operations. The theory has fundamental applications to anti-submarine warfare, counter-mine warfare, and search and rescue operations. The books [6] and [10] are classical references in this area; with [11] a valuable recent reference. Discrete search problems of the type addressed in this paper are not new. Optimal whereabout search, where we seek to maximize the probability of determining which box contains a certain object, is studied in [1] and [5]. Chew [3] considers an optimal search with stopping rule where all search outcomes are independent, conditional on the location of the searched object and the search policy. Wegener [12] investigates a search process where the search time of a cell depends on the number of searches so far. A minimum cost search problem is discussed in [8], where only one search mode is considered and the sensor has perfect specificity. The paper [9] deals with discrete search with multiple sensors in order to maximize the probability of successful search of a single target during a specified time period. Other discrete search problems are studied in [2, 7, 13]. However, all of the aforementioned references assume that the sensor has perfect specificity, that is, there are no false positive detections. Our models, which are based on [4], relax this assumption.

The main contribution of this paper, in addition to the relaxation of the perfect specificity assumption, is the development of two novel sensor models (*smart* and *dummy* sensors; see Section 2), and their application to a variety of scenarios. For the scenarios examined, the results and analysis indicate that,

- The level of initial intelligence regarding the area of interest has a significant effect on the optimal employment of the sensors and on the expected number of *determined* area-cells, and this effect is quantified.
- The optimal employment of a sensor follows a greedy rule where search effort is first invested in area-cells that are more likely to be determined than others.
- The *smart* sensor significantly outperforms the *dummy* sensor in situations of minimum uncertainty regarding the presence or absence of the threat object. In other situations the effect is not significant.

Efficient Employment of Non-Reactive Sensors

Moshe Kress

Naval Postgraduate School
mkress@nps.edu

Roberto Szechtman

Naval Postgraduate School
rszechtm@nps.edu

Jason S. Jones

US Navy
jason.s.jones1@navy.mil

APPLICATION AREA:
Operations Research and Intelligence and Unmanned Systems

OR METHODOLOGY:
Probablistic Operations Research

- The effectiveness of a sensor is determined by the relative values of its sensitivity and specificity and not by the absolute values of these parameters, except when either of these two parameters is very small, in which case sensor's effectiveness is very sensitive to the values of the other parameter.

The paper is organized as follows. In "Operational Setting"we describe the operational setting and in "Models" we formulate the models for the dummy and smart sensors. In "Results and Analysis" we analyze the models with respect to various scenarios, and in "Conclusions" we discuss the conclusions of the paper.

OPERATIONAL SETTING

Targets (e.g., missile launchers) are scattered in an area of interest and the objective of the field commander is to detect as many as possible of them. The area of interest is divided into a grid of area-cells such that each area-cell may contain at most one target. A sensor is assigned to search a certain area-cell for a certain time period during which it can make a finite number of discrete observations or *looks.* The result of each look is either a *detect* result or a *no detect* result. The sensor is imperfect - it is subject to false-positive and false-negative errors - and therefore the sensor's cues may be erroneous. The information provided by the sensor is used by the field commander to decide on further tactical or operational actions. Our goal is to help the field commander to determine the best search plan such that the information provided by the search results - his awareness regarding which area-cells are likely to contain targets and which area-cells are likely to be empty - is maximized. This *informational* MOE is described next.

An area-cell is said to be *determined* if it can be ascertained, with a given (high) probability, whether it contains a target or not. Specifically, given two probability thresholds, selected by the commander and reflecting his attitude regarding uncertainty, an area-cell is determined to be empty if the post-search probability that a target is in that cell is lower than the lower threshold. The area-cell is determined as containing a target if that posterior probability is higher than the higher threshold. The objective is to maximize the expected number of area cells that are determined. This type of information - classifying area-cells as being very likely or very unlikely to contain targets - can help field commanders filter a sizable area of interest down to only those area-cells that are likely to contain a target, and therefore better focus their operational effort.

The sensors we consider are *non reactive;* the assignment of looks to area-cells is made in advance and it does not change dynamically following information (*detection* and *no-detection* results) obtained during the search. This situation is applicable in particular to pre-programmed UAVs whose way-points and search pattern cannot be modified during the search mission.

We consider two types of sensors: *dummy* and *smart.* The *dummy* sensor evaluates the detection/no-detection results of a certain area-cell only at the end of the search, after all assigned looks have been exhausted. Based on the resulting posterior probability and the two probability thresholds, the searcher decides at that point if the area-cell is determined or not. This sensor represents a *batch* handling of the sensor data; the searcher examines the sensor's results and decides upon them only after the search process is over. The *smart* sensor examines the detection/no-detection results and computes the probability of a target continuously during the search. If at any point during the search this probability crosses either of the two thresholds, the area-cell is determined before all the looks are exhausted.

MODELS

We start this section by describing the basic framework shared by the two models. Specifically, we assume that one sensor is assigned an area of interest to search, which is partitioned into a grid of I area-cells. We assume that the area of interest can be partitioned in such a way that each area-cell i, for $i = 1, 2, \ldots, I$, contains at most one threat object. The sensor has a finite number of L looks that it can apply to the search. These looks are allocated to the various

area-cells prior to the start of the search mission.

We suppose that there is some initial intelligence about the presence of threat objects, which is manifested by a prior probability. Let $\theta = (\theta_1, \ldots, \theta_I)$ be the parameter that describes the presence/absence of threat objects; that is, $\theta_i = 1$ if there is a threat object in area-cell i, and $\theta_i = 0$ otherwise. The intelligence is captured by the prior probability mass function of θ_i,

$$\pi_i^{(0)} = P(\theta_i = 1),$$

for $i = 1, \ldots, I$. Following a single look at an area-cell, the sensor returns either a *detection* or a *no-detection* signal. The sensor is characterized by its sensitivity and specificity; for each area-cell i we have

$$p_i = P(\text{sensor indicates detection} | \theta_i = 1),$$

which is called the *sensitivity* of the sensor. The *specificity* of the sensor is $1 - q_i$, where

$$q_i = P(\text{sensor indicates detection} | \theta_i = 0).$$

Although the p_i's and q_i's may depend on the area-cell, we assume that they do not depend on the number of looks. Without loss of generality we take $p_i > q_i$, because we can reverse the cue if $p_i < q_i$. We explicitly assume that $p_i \neq q_i$, for otherwise the sensor does not provide any valuable information.

After the sensor looks at an area-cell, the intelligence regarding the likelihood of a threat object gets updated, and we obtain a *posterior* probability. More specifically,

$$\pi_i^{(1)}(\omega) = \begin{cases} \dfrac{p_i \pi_i^{(0)}}{p_i \pi_i^{(0)} + q_i(1 - \pi_i^{(0)})}, \\ \text{if } \omega = \text{sensor indicates detection} \\ \dfrac{(1 - p_i)\pi_i^{(0)}}{(1 - p_i)\pi_i^{(0)} + (1 - q_i)(1 - \pi_i^{(0)})}, \\ \text{if } \omega = \text{sensor indicates no-detection} \end{cases} . \quad (1)$$

In this way, for area-cell i we have a sequence of posteriors $\pi_i^{(1)}, \pi_i^{(2)}, \ldots$, adapted to the sequence of signals generated by the sensor in that area-cell.

We assume that the collection of look results are independent for a given area-cell; this assumption asserts that there is not systematic bias in the sensor. The results for different area-cells may be dependent. As the number of looks for a area-cell i increases, the posterior approaches 1 (if $\theta_i = 1$) or 0 (if $\theta_i = 0$). In reality, one would stop looking when the posterior becomes sufficiently close to 1 or 0. This motivates the introduction of two thresholds, which are subjective measures set by an individual involved in the search mission, such as the watch officer in the tactical operations center, or the field commander in charge of attacking these threat objects. An area-cell is considered to be *determined* if the posterior has crossed either an upper threshold or a lower threshold. If the posterior has crossed the upper threshold β, then the conclusion is that the area-cell is most likely to contain a threat object. Conversely, if the posterior has crossed the lower threshold α, then the area-cell is most likely to be clear. To make the problem non-trivial, we assume throughout the paper that $0 \leq \alpha < \beta \leq 1$.

In most realistic situations, the number of looks available is not large relative to the number of area-cells I and therefore an optimal resource (looks) allocation is needed. Specifically, the decision variables for both the dummy and smart sensor models are the number of looks allocated to area-cell i, denoted by l_i. The measure of effectiveness is the expected number of area-cells determined with at most L looks. Observe that

$$E(\#\text{ area-cells determined with } l_1, \ldots, l_I \text{ looks}) = \sum_{i=1}^{I} P(\text{area-cells } i \text{ determined in } l_i \text{ looks}).$$

It follows that the optimization problem is to choose $l_1, \ldots, l_I$ that

$$\text{maximize} \sum_{i=1}^{I} P(\text{area-cell } i \text{ determined in } l_i \text{ looks}) \quad (2)$$

subject to

$$\sum_{i=1}^{I} l_i \le L$$

$$l_i \ge 0 \text{ and integer for } i = 1,\ldots,I.$$

In order to solve this problem we need to find P(area-cell i determined in l_i looks) for the dummy and smart sensors. This is the subject of the next two subsections.

Remark 1 *The objective function in Problem 2 is non-linear, and even not necessarily concave; see Figure 3. Once P(area-cell i determined in l_i looks) is found for each number of looks* 1, 2, . . ., *L, Problem 2 can be implemented and solved - we used GAMS to illustrate the results in this paper.*

Dummy Sensor

The dummy sensor is characterized by the fact that in each area-cell the sensor checks its status (i.e. posterior probability) only after the allocated l looks are exhausted. If at that point the posterior is larger than β or smaller than α then the area-cell is declared *determined*. A smarter sensor would watch the posterior continuously and determine the area-cell as soon as the posterior crosses a threshold. Indeed, this is the characterization of the *smart* sensor discussed in the next subsection.

Let D_i = number of detections in area-cell i. Conditioning on θ_i we have

$$P(D_i = d) = \binom{l}{d} p_i^d (1-p_i)^{l-d} \times \pi_i^{(0)} + \binom{l}{d} q_i^d (1-q_i)^{l-d} \times (1-\pi_i^{(0)}), \quad (3)$$

for $d = 0, 1, \ldots, l$. When $D_i = d$ the dummy posterior $\psi_i^{(l)}(d)$ is given by, after some algebra,

$$\psi_i^{(l)}(d) = \frac{p_i^d (1-p_i)^{l-d} \pi_i^{(0)}}{p_i^d (1-p_i)^{l-d} \pi_i^{(0)} + q_i^d (1-q_i)^{l-d} (1-\pi_i^{(0)})} \quad (4)$$

Next we ask: How many detections will cause the dummy posterior to be outside either threshold? In other words, for what values in the range of D_i do we have $\psi_i^{(l)}(d) \ge \beta$ or $\psi_i^{(l)}(d) \le \alpha$? Solving for d in Equation (4) we obtain $\alpha < \psi_i^{(l)}(d) < \beta$ if and only if $a_i < d < b_i$, where

$$a_i = \frac{\log\left(\frac{\alpha}{1-\alpha}\frac{1-\pi_i^{(0)}}{\pi_i^{(0)}}\right) + l\log\left(\frac{1-q_i}{1-p_i}\right)}{\log\left(\frac{p_i(1-q_i)}{(1-p_i)q_i}\right)} \quad (5)$$

and

$$b_i = \frac{\log\left(\frac{\beta}{1-\beta}\frac{1-\pi_i^{(0)}}{\pi_i^{(0)}}\right) + l\log\left(\frac{1-q_i}{1-p_i}\right)}{\log\left(\frac{p_i(1-q_i)}{(1-p_i)q_i}\right)} \quad (6)$$

Hence

$$P(\text{area-cell i not determined in } l \text{ looks}) = P(a_i < D_i < b_i) \quad (7)$$

where the probability mass function of D_i is computed according to Equation (3). It is beneficial to view the interval (a_i, b_i) as a *no determination* region. Naturally,

$$P(\text{area-cell i determined in } l \text{ looks}) = 1 - P(\text{area cell } i \text{ not determined in } l \text{ looks}),$$

can then be employed in the optimization problem.

As an example of the dummy sensor model, consider Figure 1. For $l = 1$, it is impossible to determine the area cell because the posterior is always inside the thresholds; for $l = 2$, having two detections cause the posterior to be above β and so the area-cell is determined. Observe, however, that the probability of determining the area-cell is less for $l = 3$ (.47) than for $l = 2$ (.48), because getting a no-detection after two detections decreases the posterior and pushes it back to within the thresholds.

Smart Sensor

The smart sensor monitors the posterior continuously and therefore may determine an

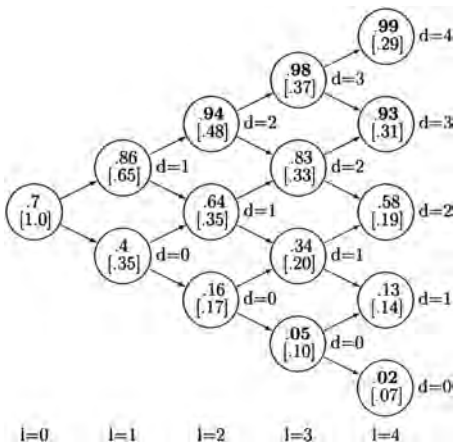

Figure 1. Probability transitions for the dummy sensor model, for $\pi^{(0)} = 0.7$, $p = 0.8$, $q = 0.3$, $\alpha = 0.1$, and $\beta = 0.9$. Inside each node, the top number is the posterior, and the number in brackets is the probability of arriving at the node; nodes with **bold** posteriors occur when the area-cell is *determined.*

area-cell as soon as the posterior crosses a threshold, before the looks allocated to that cell are actually exhausted. The subsequent looks are essentially redundant. Although there are several approaches to compute the probability of detection, a simple approach is to use dynamic programming. Define $V_{i,l}(\pi_i)$ as the probability of determining the presence, or absence, of a threat object after l looks in area-cell i, given the current prior probability is π_i. We have the boundary conditions

$$V_{i,l}(\pi_i) = 1, \text{ if } \pi_i \geq \beta \text{ or } \pi_i \leq \alpha,$$

else if $l = 0$

$$V_{i,0}(\pi_i) = 0.$$

For $l \geq 1$, the recursion is given by

$$V_{i,l}(\pi_i) = (p_i\pi_i + q_i(1 - \pi_i))V_{i,l-1}\left(\frac{p_i\pi_i}{p_i\pi_i + q_i(1-\pi_i)}\right) + ((1-p_i)\pi_i + (1-q_i)(1 - \pi_i))V_{i,l-1}\left(\frac{(1-p_i)\pi_i}{(1-p_i)\pi_i + (1-q_i)(1-\pi_i)}\right).$$

Given a prior $\pi_i^{(0)}$ and l looks, we start the above recursion with $V_{i,l}(\pi_i^{(0)})$.

Another (computationally faster) approach to compute $V_{i,l}(\pi_i^{(0)})$ is to notice that $\Pi_i = (\pi_i^{(k)} : k \geq 0)$ is a Markov chain defined on [0, 1]. Let $\tau = \inf\{k \geq 0 : \pi_i^{(k)} \notin (\alpha, \beta)\}$ be the first look at which the posterior crosses either threshold. We have

$$V_{i,l}(\pi_i^{(0)}) = P(\tau \leq l). \tag{8}$$

Let $B = (B_{xy} : \alpha < x, y < \beta)$ be the restriction of the transition kernel of Π_i to (α, β). Then

$$P(\tau > l) = \sum_{\alpha<y<\beta} B^l_{\pi_i^{(0)}y},$$

which together with Eq. (8) leads to $V_{i,l}(\pi_i^{(0)})$.

An example of the smart sensor search process is shown in Figure 2. In this example $B_{.7,.86}{}^1 = .65$, $B_{.7,.4}{}^1 = .35$ and $B_{.7,y}{}^1 = 0$ for all other values of y. Also $P(\tau > 2) = .52$ and $P(\tau > 3) = .42$. The difference between the dummy and smart sensors is that the smart sensor determines the area-cell when arriving at a node whose posterior is outside the thresholds. So while the probability that a dummy sensor determines an area-cell after 3 looks is .37 (see Figure 1) the smart sensor determines it with probability .48 + .10 = .58.

RESULTS AND ANALYSIS

In this section we present the results and their analysis for both the dummy and smart

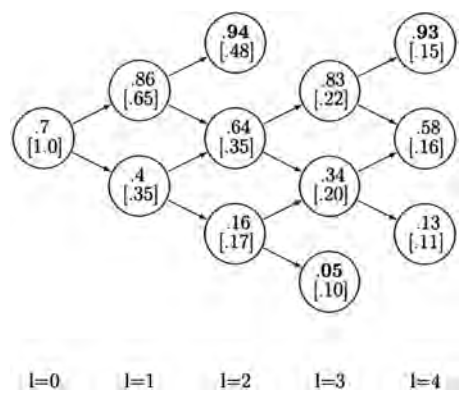

Figure 2. Probability transitions for the smart sensor model, for $\pi^{(0)} = 0.7$, $p = 0.8$, $q = 0.3$, $\alpha = 0.1$, and $\beta = 0.9$. Inside each node, the top number is the posterior, and the number in brackets is the probability of arriving at the node; nodes with **bold** posteriors occur when the area-cell is *determined.*

sensor models. In the first subsection we discuss the single area-cell scenario, while in the second subsection we optimize sensor employment in multiple area-cells.

Single area-cell Scenario

We analyze the effect of the model parameters on the probability of determining a single area-cell, so the optimization problem (2) does not come into play. To simplify notation, we drop the subindex i in the discussion that follows in this subsection.

First, we consider the dummy sensor. From the definition of a and b, it is easy to see that they are linear functions of l with positive slope since $p > q$, and that the difference $b - a$ is constant in l. So, we have the following result (see the Appendix for all proofs).

Proposition 1 *The probability of determining an area-cell approaches 1 as the number of looks grows to infinity.*

Observe that the number of integers that lie in the open interval (a, b) is not necessarily constant as a function of the number of looks l. That is, on the sample paths where the dummy posterior is outside the thresholds, a sensor signal may push the dummy posterior back into (α, β), thus increasing the probability of not determining the area-cell. Looking at Equation (7), this suggests that P(area-cell i determined in l looks) may not be monotonic in the number of looks l. Indeed, for certain parameter settings, increasing the number of looks actually lowers the probability of determining an area-cell. Figure 3 illustrates this situation: When the number of looks goes from 3 to 4, the probability of determining the area-cell decreases; the same happens when going from 5 to 6 looks, 7 to 8 looks, etc. This phenomenon is demonstrated also in Figure 1, as discussed above. From the definition of a in (5) and of b in (6), it follows that the cardinality of the open interval (a, b) is generally not continuous in the model parameters. Ultimately, this causes Figure 3 - Figure 8 to be jagged for the dummy sensor.

An important issue is the effect of the prior intelligence on the probability of determining an area-cell. Note that the prior $\pi^{(0)}$ is the mixture parameter of the Binomial mixture in Equation (3), and it appears in the definition of a and b. When the number of looks is large, the area-cell is determined with very high probability, regardless of the prior. Hence, for the purpose of our analysis we assume that l is not too large and consider three ranges of $\pi^{(0)}$:

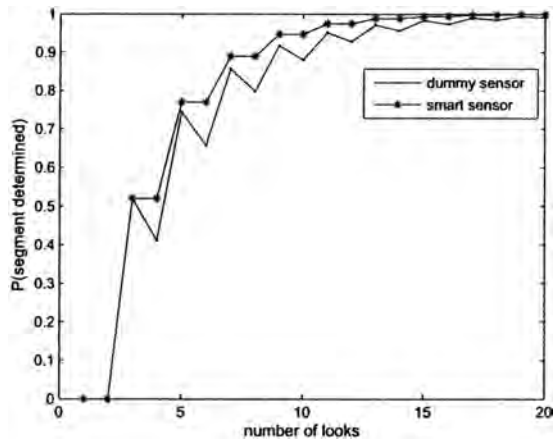

Figure 3. Probability of determining an area-cell as a function of number of looks, for $\pi^{(0)} = 0.5$, $p = 0.8$, $q = 0.2$, $\beta = 0.95$, and $\alpha = 0.05$.

- $\pi^{(0)}$ is close to the lower threshold α. In this case $D \underset{\approx}{D} Bin(l, q)$ (where $\underset{\approx}{D}$ means approximately distributed, and $Bin(l, q)$ is a binomial distribution with l looks and probability of detection q). Also, a and b are in the highest part of their range; that is, we determine the area-cell for small values of D. But this is precisely what happens when $D \underset{\approx}{D} Bin(l, q)$ and q is not too large: D is most likely to be a small number. Hence, when $\pi^{(0)} \approx \alpha$ and q not too large, the probability of determining the area-cell is large.
- $\pi^{(0)}$ is close to the upper threshold. In this case $D \underset{\approx}{D} Bin(l, p)$ so that D puts most of its mass in the higher end of its range if p is large. Also, a and b are in the low part of their range, so that we determine the area-cell for large values of D. Putting the last two observations together, we conclude that the probability of determining an area-cell is large when $\pi^{(0)}$ is close to β and p is large.
- $\pi^{(0)}$ is not close to either threshold, p and q are mid range. In this situation $\pi^{(0)}$ is a mixture of binomials, and since p and q are mid range, D is most likely to take values in the middle of its range. Also, a and b are in the

middle part of their range. Putting these two arguments together, we conclude that the probability of determining an area-cell will be small under these circumstances.

Figure 4 illustrates the above analysis. Other than confirming our explanation of the effect of the prior, Figure 4 is rather striking because of its jumps; these are due to the change in the number of integers that lie in the interval (a, b) as we change the prior, a phenomenon observed and discussed earlier.

Now we address the effect of sensitivity (p) and specificity ($1 - q$). The basic question regarding the parameters p and q is: What is a good dummy sensor with respect to these parameters? Naturally, $p = 1$ and $q = 0$ is *the* perfect sensor, but this situation is unattainable in practice. As Figure 5 suggests,

- it is the difference $p - q$ that makes a sensor better or worse, regardless of the absolute values of the parameters; the probability of determining an area-cell increases with $p - q$.
- for low values of p, the probability of determining an area-cell is very sensitive to q; that is, a small increase in q causes the probability of determining an area-cell to decrease significantly. The reason for this behavior is that (a, b) expands to include all the integers in $[0, l]$ as q gets closer to p small.
- for high values of q, the probability of determining the area-cell is very sensitive to p. As p moves from q to 1, the no determination region (a, b) moves away from $[0, l]$, thus causing the area-cell to be determined.

Regarding the smart sensor, the following proposition summarizes some of its properties.

Proposition 2 *The smart sensor has a determination probability that is non-decreasing, approaches one as the number of looks increases, and is not smaller than the determination probability of the dummy sensor for the same number of looks.*

Figure 3 illustrates the last proposition. We explain the observation that the probability of determining an area-cell remains constant when going from 3 to 4 looks, from 5 to 6 looks, etc, by the fact that for the parameters settings of Figure 3, on any sample path where the posterior is within (α, β) prior to looking at the area-cell, the posterior remains within (α, β) regardless of the sensor signal (detection or no-detection). Like in the dummy sensor case, a

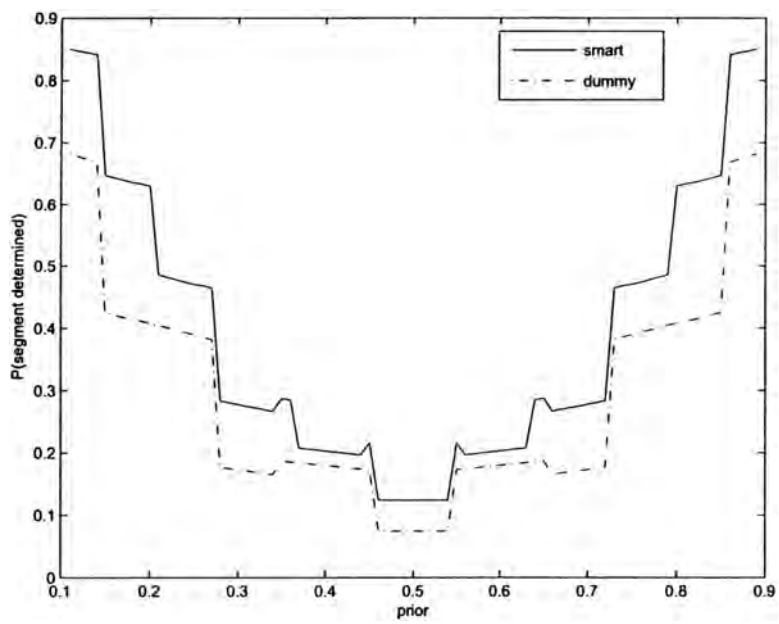

Figure 4. Probability of determining an area-cell as a function of the initial prior, for $l = 11$, $p = 0.6$, $q = 0.4$, $\beta = 0.9$, and $\alpha = 0.1$.

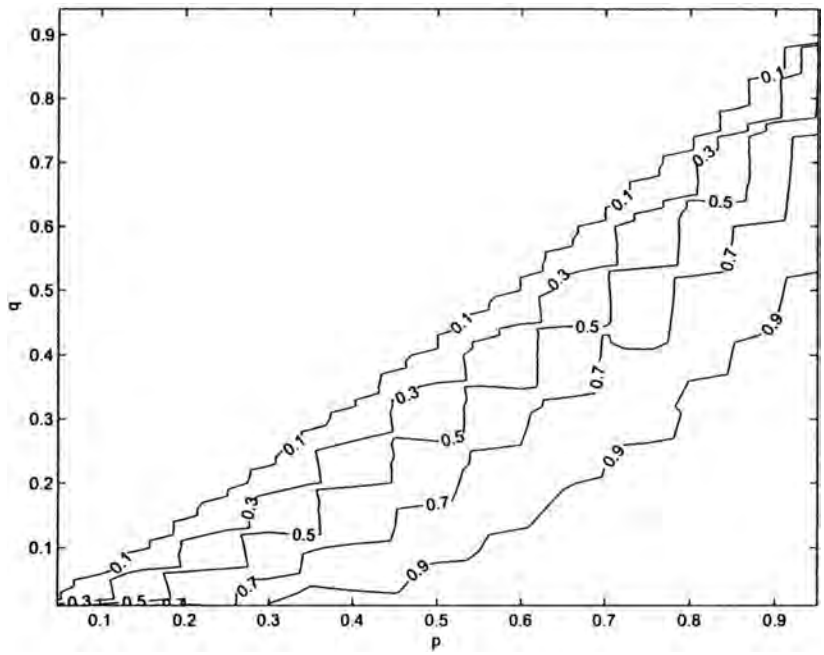

Figure 5. Contour plot of the probability of determining an area-cell for the dummy sensor, as a function of the sensitivity and specificity, for $l = 11$, $\pi^{(0)} = .8$, $\beta = 0.9$, and $\alpha = 0.1$.

determination probability that is not everywhere differentiable causes Figure 3 - Figure 8 to be non-smooth for the smart sensor as well.

The effect of prior intelligence on the smart sensor is illustrated in Figure 4, with the following interpretation:

- As $\pi^{(0)}$ gets close to either threshold, the probability of determining the area-cell approaches 1. That is, when $\pi^{(0)} \approx \alpha$ and α is small, according to Equation (1), $\pi^{(1)}$ (no-detection) ≈ 0 (so we cross the lower threshold) with probability equal to P(no-detection signal) $\approx 1 - q$, so that $P(\tau = 1) \approx 1 - q$; if we get a detection in the first look, the same analysis shows that $P(\tau = 2) \approx q(1 - q)$. Proceeding in that fashion we see that P(area-cell determined for l small) ≈ 1 when $\pi^{(0)} \approx \alpha$ and $\alpha \approx 0$.
- The analysis for the upper threshold β is analogous when $\pi^{(0)} \approx \beta$ and $\beta \approx 1$, and we have P(area-cell determined for l small) ≈ 1.
- A remarkable feature of Figure 4 is that the smart sensor significantly outperforms the dummy sensor when the prior $\pi^{(0)}$ is close to either threshold. The reason for this behavior is that the dummy sensor only checks the value of the posterior when all the looks have been exhausted, by which time it is possible that $\psi^{(l)}$ is within the thresholds. In the next section we discuss how this phenomenon carries over to the multiple area-cell situation.

The effect of sensitivity and specificity with respect to the smart sensor (see Figure 6) is similar to the dummy sensor: The difference $p - q$ is the important measure concerning sensor performance, and the probability of determining an area-cell increases with $p - q$. Like in the dummy sensor model, the probability of determining the area-cell is very sensitive to q when p is small, and very sensitive to p when q is large.

Multiple area-cells Scenario

When there is more than one area-cell, we solve Problem (2) to obtain an efficient allocation of looks. Evidently, the effectiveness of the sensors increases with the number of looks, because the uncertainty regarding the presence or absence of threat objects in each area-cell is

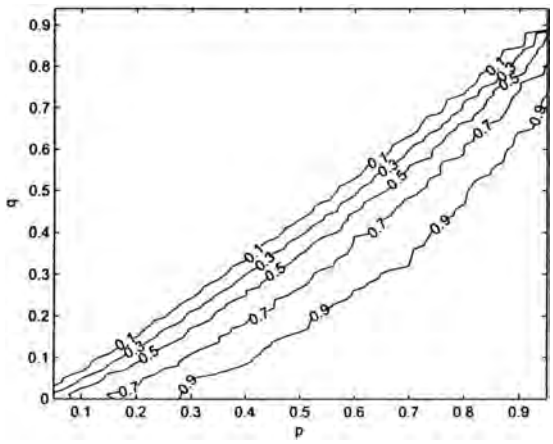

Figure 6. Contour plot of the probability of determining an area-cell for the smart sensor, as a function of the sensitivity and apecificity, for $l = \pi^{(0)} = .8$, $\beta = 0.9$, and $\alpha = 0.1$.

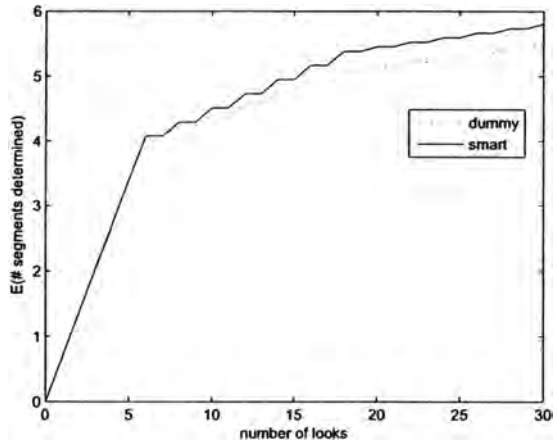

Figure 7. Expected number of area-cells determined as a function of the number of looks, for $I = 6$, $L = 30$, $\pi^{(0)} = 0.8$, $p = 0.8$, $q = 0.2$, $\beta = 0.9$, and $\alpha = 0.1$.

revealed as we increase the number of looks. In view of Propositions 1 and 2, for any *fractional* allocation $\hat{}l_1 = t_1L$, $\hat{}l_2 = t_2L \ldots$, $\hat{}l_I = t_IL$ such th at all the t's are positive and $t_1 + \ldots + t_I = 1$, we have

$$E(\#\text{area-cells determined with } \hat{l}_1,\ldots,\hat{l}_I \text{ looks}) \to I,$$

as $L \to \infty$. Since the optimal allocation is no worse than the $\hat{l}_1, \ldots, \hat{l}_I$ allocation, we have

Proposition 3 *Suppose that $l_1^*(L), \ldots, l_I^*(L)$ is an optimal solution to Problem (2) when there are L looks available. Then, for both the dummy and smart sensor models,*

$$E(\#\text{ area-cells determined with } l_1^*(L),\ldots,l_I^*(L) \text{ looks}) \to I,$$

as $L \to \infty$.

In words, Proposition 3 states that the expected number of area-cells determined under an optimal allocation of looks approaches the total number of area-cells, as the number of looks available grows. Also, since the smart sensor cannot be worse than the dummy sensor, the expected number of area-cells determined by the smart sensor is never smaller than the expected number of area-cells determined by the dummy sensor. This observation and Proposition 3 are demonstrated in Figure 7, where $I = 6$ and all area-cells have the same prior, sensitivity and specificity probabilities. For $L \leq 7$ we have that the dummy and smart sensor models yield the same result, this is because each of the 6 area-cells gets no more than 2 looks, and P(area-cell determined) is the same for both sensors in this situation (cf. Figure 3). As the number of looks available increases, the smart sensor has a larger number of area-cells determined than the dummy sensor, in accordance with Proposition 2.

Next we examine the effect of the prior, sensitivity and specificity probabilities ($\pi^{(0)_i}$, p_i and $1 - q_i$, respectively) on the the optimal *allocation* of looks. The question is: What area cells get a large (or small) number of looks at optimality? Due to the high dimensionality of the problem, it is impossible to run a full factorial experiment. Therefore, we settle with solving Problem (2) under various representative scenarios that capture the main effects of the above parameters.

Concerning the effect of the prior (see Figure 8), we have the same conclusion as for the single area-cell scenario, namely: The priors close to the thresholds α and β lead to larger expected number of area-cells determined, and the smart sensor significantly improves on the dummy sensor in situations of good prior intelligence. Figure 8 shows these properties for $I = 6$ area-cells when the priors of all area-cells shift together from the lower threshold to the upper threshold. We solved the optimization

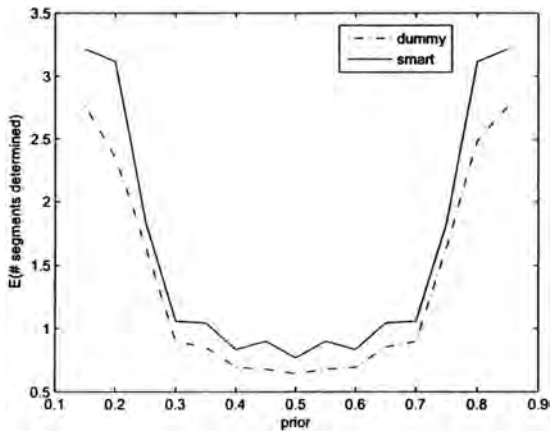

Figure 8. Expected number of area-cells determined as a function of the initial prior, for $I = 6$, $L = 30$, $\pi^{(0)} = 0.8$, $p = 0.8$, $q = 0.2$, $\beta = 0.9$, and $\alpha = 0.1$.

problem (2) under several other representative configurations, with all the results supporting the above conclusions.

Table 1 summarizes the results of a more detailed analysis. For each one of the two sensors we consider two levels of effectiveness, manifested by the sensitivity and specificity of the sensor, and two configurations of prior probabilities. A relatively *ineffective* sensor has $p = .6$ and $q = .4$ for all area-cells, while for a relatively *effective* sensor these parameters are .7 and .3, respectively. For each sensor and each level of effectiveness we consider two spatial configurations of the prior probabilities: (1) *Uniform Worse-Case* configuration where $\pi_i^{(0)} = .5$, $i = 1, \ldots, 6$, and (2) *Mixed* configuration where the prior is close to the upper threshold for two area-cells, the prior is far from both thresholds for two area-cells, and the prior is close to the lower threshold for two area-cells. For each sensor, level of effectiveness and spatial prior configuration, Table 1 presents the optimal allocation and the maximum expected number of determined area-cells.

The take-away of Table 1 is:

- When the sensors are relatively ineffective, and and the prior configuration is uniform ($\pi_i^{(0)} = 0.5$, $i = 1, \ldots, 6$), both the dummy and smart sensors allocate all the looks to just two area-cells. This happens because in this situation it takes a large number of looks to make the probability of determining an area-cell lift above its zero-value floor. If the prior configuration is mixed, there are four area-cells with prior probabilities close to a threshold. Hence, the model allocates the looks in a greedy fashion so as to determine these four area-cells. In accordance to our single area-cell analysis, the smart sensor significantly outperforms the dummy sensor.
- When the sensors are relatively effective and the prior configuration is uniform, both the smart and dummy sensors uniformly allocate the looks among the 6 area-cells. This occurs because it takes a small number of looks to have the initial shoot up in the probability of determining an area-cell. Observe that the expected number of determined cells shows a remarkable increase from the ineffective-sensor uniform-prior situation; we will have more to say about this issue when we analyze the effect of the p, q configuration.
- When the sensors are effective and the prior has a mixed spatial configuration, it takes a few looks to determine the area-cells whose

Table 1. Effect of the prior for $L = 30$, $I = 6$, $\alpha = 0.1$, and $\beta = 0.9$

		Effect of the prior							
p, q	*Spatial Configuration*	E(# det.)	l_1	l_2	l_3	l_4	l_5	l_6	*Sensor*
$p = .6, q = .4$	$\pi^{(0)} = (.5, .5, .5, .5, .5, .5)$	0.7705	16	14	0	0	0	0	smart
		0.6447	16	14	0	0	0	0	dummy
	$\pi^{(0)} = (.8, .8, .5, .5, .2, .2)$	2.4488	8	8	0	0	8	6	smart
		1.9936	10	8	0	0	6	6	dummy
$p = .7, q = .3$	$\pi^{(0)} = (.5, .5, .5, .5, .5, .5)$	3.6186	5	5	5	5	5	5	smart
		3.3540	5	5	5	5	5	5	dummy
	$\pi^{(0)} = (.8, .8, .5, .5, .2, .2)$	4.6857	3	3	9	9	3	3	smart
		4.2253	5	5	9	9	1	1	dummy

prior is close to a threshold. Hence, the optimal allocation specifies a large number of looks for the cells whose prior is far from either threshold.

Regarding the sensitivity and specificity of the sensors, in all the scenarios we assume a worse-case prior $\pi^{(0)} = 0.5$ for all area cells $i = 1, \ldots, 6$, and consider two general cases of sensing situations - relatively effective and relatively ineffective. We consider three types of spatial configurations of these sensing capabilities over the 6 area-cells. Recall that a sensor becomes more effective as $p_i - q_i$ increases (see Figures 5 and 6). Thus, in the *effective* case we assume that the average value of $p_i - q_i$ (denoted as $\bar{p} - \bar{q}$) is 0.7, while in the *ineffective* case this average difference is 0.2. The three spatial configurations are: (1) *Mixed* - Three area-cells with relatively large difference and three with relatively small difference, (e.g., $p_i = .9$, $q_i = .1$, $i = 1, 2, 3$; $p_i = .8$, $q_i = .8$, $i = 4, 5, 6$), (2) *Uniform* - All six area-cells have the same difference (e.g., $p_i = .9$, $q_i = .2$, $i = 1, \ldots, 6$), (3) *Monotonic* - the sensitivity is monotonic decreasing, the specificity is monotonic increasing but $p_i - q_i$ remains constant for every $i = 1, \ldots, 6$.

Table 2 summarizes the results of the optimization models for both sensors. From this table we can draw the following conclusions:

- For an effective sensor ($\bar{p} - \bar{q} = 0.7$) almost all the area-cells are determined. While the smart sensor is obviously better, the difference in the expected number of determined area-cells between the two sensors is 5% or less. Also, the spatial configuration has only a small effect on that measure.
- For the ineffective sensor ($\bar{p} - \bar{q} = 0.2$) the performance of the sensors is quite poor (one or two determined area cells) and it depends on the spatial configuration. Both sensors perform best when the spatial configuration is mixed, and worst when it is uniform. The difference in the expected number of determined area-cells between these two spatial configurations is about 100% for both sensors.
- The optimal allocation of looks depends both on the effectiveness of the sensor and the spatial configuration of its effectiveness across the area-cells, but not on the type of sensor (smart or dummy). When the sensor is effective, looks are spread out more or less evenly across the area-cells, unless the sensitivity and specificity are high (e.g., .9 each), in which case one look will suffice. When the sensor is ineffective ($\bar{p} - \bar{q} = 0.2$), then the search effort is concentrated in a few area-cells, which are most likely to become determined after a considerable number of looks (e.g., 11 looks in the mixed case). The other area cells are ignored.

CONCLUSIONS

In this paper we developed two bayesian-oriented models that describe the performance

Table 2. Sensitivity and specificity effects for $L = 30$, $I = 6$, $\pi^{(0)} = 0.5$, $\alpha = 0.1$, and $\beta = 0.9$

		Effect of sensitivity and specificity							
$\bar{p}$ - $\bar{q}$	*Spatial Configuration*	E (# det.)	l_1	l_2	l_3	l_4	l_5	l_6	*Sensor*
0.7	p = (.9, .9, .9, .8, .8, .8)	5.9757	1	1	1	9	8	10	smart
	q = (.1, .1, .1, .2, .2, .2)	5.8818	1	1	1	10	8	8	dummy
	p = (.9, .9, .9, .9, .9, .9)	5.9255	5	5	5	5	5	5	smart
	q = (.2, .2, .2, .2, .2, .2)	5.6277	5	5	5	5	5	5	dummy
	p = (.96, .91, .86, .81, .76, .71)	5.9598	5	5	8	5	5	2	smart
	q = (.26, .21, .16, .11, .06.01)	5.7693	6	5	6	5	6	2	dummy
0.2	p = (.8, .8, .8, .5, .5, .5)	2.0485	11	11	8	0	0	0	smart
	q = (.5, .5, .5, .4, .4, .4)	1.6179	11	11	8	0	0	0	dummy
	p = (.8, .8, .8, .8, .8, .8)	0.9845	15	15	0	0	0	0	smart
	q = (.6, .6, .6, .6, .6, .6)	0.8334	15	15	0	0	0	0	dummy
	p = (.96, .86, .76, .66, .56, .46)	1.3654	10	13	0	0	0	7	smart
	q = (.76, .66, .56, .46, .36, .26)	1.2320	10	13	0	0	0	7	dummy

of two types of imperfect sensors - *dummy* and *smart* - and presented optimal employment schemes for these sensors in a variety of scenarios. We have shown and quantified the advantage of he smart sensor over the dummy one, which underscores the importance of continuous monitoring of sensor data, in particular in the presence of prior intelligence. We have demonstrated the importance of this prior intelligence on the effectiveness of the search; on the optimal employment of the sensors and on the expected number of *determined* area-cells. The optimal employment of sensors is *greedy* in the sense that search efforts must be allocated to area-cells where they can produce definitive information in the form of determined area-cells. Finally, we demonstrated that for realistic sensing capabilities, the effectiveness of a sensor is determined by the relation between its sensitivity and specificity, rather than the absolute values of these parameters.

The models developed in this paper may be extended to other types of sensors - in particular *reactive* sensors that may facilitate dynamic employment during the search mission.

APPENDIX

Proof of Proposition 1. Since we focus on just one area-cell, we drop the subindex i from the notation. Consider a collection of random variables that describes the number of detections, indexed by the number of looks: $D_1, \ldots, D_l$. We wish to show that

$$P(D_l \in (a(l),b(l))) \to 0$$

as $l \to \infty$. Since

$$P(D_l \in (a(l),b(l))) = E[P(D_l \in (a(l),b(l))|\theta)],$$

it suffices to show that both $P(D_l \in (a(l), b(l))|\theta = 1)$ and $P(D_l \in (a(l), b(l))|\theta = 0)$ converge to 0 as $l \to \infty$. Observe that, by the independence of looks assumption, the Central Limit Theorem implies

$$P\left(\frac{D_l - l_p}{\sqrt{l_p(1-p)}} \in (u,v) \,\middle|\, \theta = 1\right) \to P(Z \in (u,v))$$

as $l \to \infty$, where Z is a normally distributed random variable with mean 0 and variance 1. Hence

$$P(a(l) < D_l < b(l)|\theta = 1) = P\left(\frac{a(l) - l_p}{\sqrt{l_p(1-p)}} < \frac{D_l - l_p}{\sqrt{l_p(1-p)}} < \frac{b(l) - l_p}{\sqrt{l_p(1-p)}} \,\middle|\, \theta = 1\right) \to 0,$$

as $l \to \infty$, since $l^{-1/2}b(l) - l^{-1/2}a(l) \to 0$ as $l \to \infty$. Analogously, it is possible to show that

$$P(a(l) < D_l < b(l)|\theta = 0) \to 0,$$

as $l \to \infty$. Hence we conclude that

$$P(D_l \notin (a(l),b(l))) \to 1 \tag{9}$$

as $l \to \infty$. $\otimes$

Proof of Proposition 2. We now argue that the smart sensor cannot be worse than the dummy sensor. Because $P(\tau \le l)$ is non-decreasing in l, Equation (8) shows that $V_{i,l}(\pi_i^{(0)})$ is non-decreasing in l too. Moreover,

$$P(\text{area-cell } i \text{ determined in } l \text{ looks by dummy sensor})$$

$$= \sum_{k=0}^{\infty} P(\text{area-cell } i \text{ determined in } l \text{ looks by dummy sensor}|\tau = k)\, P(\tau = k)$$

$$= \sum_{k=0}^{l} P(\text{area-cell } i \text{ determined in } l \text{ looks by dummy sensor}|\tau = k)\, P(\tau = k)$$

$$\le \sum_{k=0}^{l} P(\tau = k)$$

$$= V_{i,l}(\pi_i^{(0)}).$$

The above shows that the smart sensor cannot do worse than the dummy sensor. This, together with Equation (9) shows that

$$P(\text{area-cell } i \text{ determined in } l \text{ looks by smart sensor}) \to 1$$

as $l \to \infty$. $\otimes$

REFERENCES

[1] R. Ahlswede and I. Wegener. *Search Problems.* John Wiley and Sons, 1987.

[2] W. L. Black. Discrete sequential search. *Information and Control,* 8:159–162, 2004.

[3] M. C. Chew. A sequential search procedure. *Ann. Math. Statist.,* 38:494–502, 1967.

[4] J. S. Jones. Modeling detection strategies to battle improvised explosive devices, 2006. Master's Thesis.

[5] J. B. Kadane. Optimal whereabout search. *Operations Research,* 19:894–904, 1971.

[6] B. O. Koopman. *Search and Screening.* Center for Naval Analysis, Alexandria, Virginia, 1946.

[7] D. Matula. A periodic optimal search. *Amer. Math. Monthly,* 71:15–21, 1964.

[8] S. M. Ross. *Introduction to Stochastic Dynamic Programming.* Academic Press, 1983.

[9] N-O. Song and D. Teneketzis. Dicrete search with multiple sensors. *Mathematical Methods of Operations Research,* (60):1–13, 2004.

[10] L. D. Stone. *Theory of Optimal Search.* Academic Press, New York, 1975.

[11] A. R. Washburn. *Search and Detection.* Institute for Operations Research and the Management Sciences, 4th edition, 2002.

[12] I. Wegener. The discrete sequential search problem with nonrandom cost and overlook probabilities. *Mathematics of Operations Research,* 5:373–380, 1980.

[13] I. Wegener. Optimal search with positive switch cost is np-hard. *Information Processing Letters,* 21:49–52, 1980.

BE A MORS MEMBER … Find Rewards!!

Military Operations Research Society

MORS Membership Benefits

YOUR PROFESSIONAL SOCIETY SUPPORTING NATIONAL AND INTERNATIONAL SECURITY ANALYSIS

- Association with highly-respected experts supporting our security
- Participation in diverse and advancing analytical events
- Tutorials to maintain and advance your knowledge and skills
- Quarterly membership networking events
- Professional development opportunities including membership
- Books and merchandise discounts
- Full write privileges on MORS wiki for event and policy development
- A subscription to the MORS quarterly bulletin *PHALANX*
- A MORS Lapel Pin
- MORS Membership Card
- Reduced Meeting Fees

Save **$405.00** a year on meeting fees as a GOVERNMENT MORS member!*

Save **$570.00** a year on meeting fees as a NON-GOVERNMENT MORS member!*

*Savings are calculated assuming participation at five MORS special meetings and the annual Symposium!

More savings with tutorial attendance!

MEMBER REFERRAL INCENTIVE

As a member of MORS, you are very valuable to us and so are your friends and co-workers. We encourage you to share the benefits of MORS with them. In appreciation, **we will extend your membership for three months for each referred individuals that joins MORS member.** In addition, You will also be entered to win a free registration to the 77th MORS Symposium! Simply write your name in the blank after "Who referred you?," and suggest joining to others to get the benefits of MORS—your society!

Membership Dues:
1 year = $75.00
2 years = $140.00
3 years = $210.00

International addresses add $25 per year

Student Membership is *only* $25.00 a year!

Military Operations Research Society (MORS)

Who referred you? _____

1703 North Beauregard Street
Suite 450
Alexandria, VA 22311-1745

Phone: 703-933-9070
Fax: 703-933-9066
E-mail: tiffanie@mors.org

Name _____ Phone _____

Organization _____

Address _____

Email Address _____

Credit Card # _____ Exp. date _____

Billing Address (including zip) _____

Signature _____

Membership Dues:
1 year = $75.00 _____
2 years = $140.00 _____
3 years = $210.00 _____
Students 1 yr = $25.00 _____
International address add $25 per year _____

MOR Journal Subscriptions:
1 year = $70.00 _____
2 years = $130.00 _____
International Address add $65 per year

☐ Visa ☐ MasterCard
☐ Check #_____ ☐ American Express

Total: _____

ABSTRACT

A new measure based upon an actor's reachability to other individuals within the network is presented as an improved means to study adversarial, clandestine networks. This measure was initially motivated by (1) characteristics of existing power and status measures and (2) the requirement to quickly estimate the importance of actors within networks whose data is obtained via distant and error-prone methods. Related measures are reviewed and the underlying assumptions and theoretical bases are presented. The measure is then applied to exemplar networks and analyzed. MATLAB code is available upon request from the corresponding author.

INTRODUCTION

There exist a number of centrality measures that rely upon the structural characteristics of a social network to assess the importance of its actors (cf. Wasserman and Faust, 1994). The majority of social network analysis (SNA) measures perform calculations upon the mathematical representation of the sociogram, the *sociomatrix*, referred to by (**X**). The sociomatrix is a two-way, numerical matrix, "indexed by the sending actors (the rows) and the receiving actors (the columns) . . .," which is equivalent to the adjacency matrix of a graph (Wasserman and Faust, 1994, pg. 77).

In the context of evaluating clandestine networks, efficiently calculated measures that perform well despite limited information are of increasing interest, particularly to counter-terrorism efforts. The measure developed in this paper was specifically designed to serve as a 'screening' tool to identify individuals within such a network who may potentially serve important roles in achieving organizational objectives. Consequently, those actors are deemed of interest, serving as candidates for increased intelligence, surveillance, and analysis resources. As in all operations, military or otherwise, measures are critical.

In the context of typical network analyses, determining and examining the nature of such roles is often predicated upon network composition and analytical objectives. For example, network data that captures directed relationships (e.g., who works for whom, familial relationship, etc.) invokes the notions of prestige and power. A prestigious actor is "one who is the object of extensive ties" (Wasserman and Faust, 1994, pg. 174). Alternatively, a powerful actor is one that "influences the behavior (either overtly or covertly) of others in accordance with his own intentions," implying a focus upon measuring 'ties emanated' (Goldhamer and Shils, 1939, pg. 171). Symmetric, or undirected, network data (e.g., friendships, telephone calls, etc.) simply fall within the study of, and have a variety of accompanying measures to assess, actor centrality (cf. Wasserman and Faust, 1994, Chp. 5). The measure presented is easily applied to all three of these categories of actor role analyses (prestige, power, and centrality) and assumes that actor importance is positively correlated with an actor's position, taking into account the category under examination.

We propose the use of a new SNA measure-reach-based assessment of position (RBAP). RBAP was initially motivated by the concepts underlying the status measure of Katz (1953). It is argued that subtle theoretical changes, and the resulting computational modifications, to Katz's measure provide a more suitable approach to analyzing clandestine networks heavily reliant upon secrecy for their operational success (cf. Post, 2005; Baker and Faulkner, 1993). Note that applying Katz's measure to the transpose of the sociomatrix also permits the study of power relations. The same approach may also be taken when applying RBAP, providing flexibility in the types of analyses that may be performed (cf. Valente and Foreman, 1998).

Lastly, RBAP shares the concept of influence attenuation as a function of path length as seen in Katz (1953) and related eigenvalue-based centrality measures (Bonacich, 1987; Bonacich and Lloyd, 2001, pg. 195). Differences in and rationale behind the implementation of attenuation in RBAP are presented, as well as a novel use of this factor to further facilitate the process of screening individuals.

Screening actors of interest given assumptions of the flow (e.g., directed or undirected), the analytical context (e.g., prestige, power, or centrality), and attenuation may be based simply upon an individual's RBAP score relative to all other actors, as accomplished in traditional applications of SNA measures. However, comparing the

Reach-Based Assessment of Position

Lt Col J. Todd Hamill, USAF

USSTRATCOM/J811
HamillJ@stratcom.mil

Dr. Richard F. Deckro

AFIT/ENS
Richard.Deckro@afit.edu

Dr. Robert F. Mills

AFIT/ENG
Robert.Mills@afit.edu

Dr. James W. Chrissis

AFIT/ENS
James.Chrissis@afit.edu

Application Areas:
Social Science Methods; Counter-Terrorism; Information Operation/ Information Warfare; ISR and Intelligence Analysis

OR Methodologies:
Network Methods and Social Network Analysis

changes in scores as a result of varying the attenuation factor offers an additional opportunity to identify actors of interest by assessing their position within any of the analytic categories, from both a local and a global perspective.

The following sections briefly review related works, present the assumptions, theory, and underlying mathematical model of RBAP, and conclude with some examples. Note that the actors of interest for the exemplar networks may be obvious. However, the size of some real-world terrorist networks will likely not afford the analyst this luxury and will often require computational methods to aid experts and decision makers (cf. MIPT, 2006). Consequently, measures that can be calculated efficiently, such as RBAP, are vital to enabling the analysis of large and ever changing networks.

BACKGROUND

Given a dichotomous (e.g., 0–1) representation of a clandestine network, RBAP seeks to identify actors that are able to reach others (power), be reached by others (status), or control influence among all other actors (centrality) within the network as efficiently as possible. This invokes a common underlying assumption prevalent in many SNA measures-that influence or information propagates through a network via shortest, or geodesic, paths. The geodesic path is defined as "the (not necessarily unique) shortest path through the network from one vertex to another" (Newman, 2003, pg. 173). The definition implies that there could be multiple shortest paths of a given distance between any two actors, a phenomena leveraged in RBAP as well as the classic betweenness centrality measure (Wasserman and Faust, 1994, pg. 188–91).

From an interpersonal communications point of view, flow via the shortest path may minimize the likelihood and impact of errors or misperceptions that often plague human interaction. However, as several authors have contended, communication or influence between individuals within a clandestine network may not necessarily flow along the shortest path. For example, regarding the impetus behind their centrality measure that accounts for all possible paths between any two individuals, Stephenson and Zelen (1989) point out, "it is quite possible that information will take a more circuitous route either by random communication or may be intentionally channeled through many intermediaries in order to 'hide' or 'shield' information in a way not captured by geodesic paths" (Stephenson and Zelen, 1989, pg. 3).

Other works suggest that when an organization is faced with tradeoffs between efficiency and concealment, the subsequent network structure evolves in a manner that may be contrary to classical sociological expectations (Krebs, 2002; Baker and Faulkner, 1993, pg. 856). However, communication may still follow the shortest path relative to the secretive network, despite the fact that such a path could be shorter if the network were operating and evolving freely without recourse. If secure communications are required, it is assumed that longer communication chains offer more opportunity for interception of message traffic and concomitant operational risk, as well as increased chances for losses in information or influence. Hence, communication among paths other than the geodesics is potentially contrary to the organizational goals of secrecy (e.g., Post, 2005, Chapter 2).

Related Measures. The following sections discuss related measures that led to the development of RBAP by leveraging lessons learned and techniques of interest. These measures include the status index of Katz (1953) and the closely related centrality measures for asymmetric relations developed by Bonacich (1987), Bonacich and Lloyd (2001), as well as the radiality and integration measures proposed by Valente and Foreman (1998). For a more comprehensive comparison between the related measures, the reader is referred to Wasserman and Faust (1994, pg. 198–219) and Borgatti and Everett (2006).

a. Katz Status. Katz developed a measure for the status of individuals, in the context of a popularity contest, based not only upon how many people choose the 'most popular' individual but also accounting for who is doing the choosing. Katz suggested that this recursive measure may also be "used to study influence,

transmission of information, etc" (Katz, 1953, pg. 39).

Katz notes that the column sums of the sociomatrix, **X**, referred to by him as the choice matrix, pertains to the number of people that 'choose' that individual. Further, noting that the elements of the powers of the sociomatrix, given by $\mathbf{X}^p$, provides the number of directed walks of length p from node i to node j, he posited that this equates to the indirect p-step ($p \geq 1$) choices of a given individual by the group (Katz, 1953; Wasserman and Faust, 1994).

All possible walks may be accounted for by raising the sociomatrix to the power of infinity. An additional assumption that longer walks were less effective or influential than shorter ones required an attenuation factor $\alpha \in [0,1]$. Accepting these constructs, Katz's objective was to find the column sums of the matrix **T,** defined by

$$\mathbf{T} = \alpha\mathbf{X} + \alpha^2\mathbf{X}^2 + \alpha^3\mathbf{X}^3 + \cdots + \alpha^k\mathbf{X}^k + \cdots. \quad (1)$$

Given the computational limitations of the early 1950s, Katz cleverly sought to take advantage of the geometric series, defined as

$$\sum_{k=1}^{\infty} r^k = \frac{r}{1-r},\ r < 1, \quad (2)$$

which negated the need to explicitly calculate the infinite powers of matrices (Katz, 1953, pg. 41–42). Consequently, assuming the replacement of r with $\alpha\mathbf{X}$ results in convergence, Katz's status index, **s,** for a network of N individuals is easily calculated by

$$\mathbf{s}^{(1\times N)} = \mathbf{1}^{(1\times N)}\mathbf{T} = \mathbf{1}^{(1\times N)}\left[\frac{\alpha\mathbf{X}}{\mathbf{I}-\alpha\mathbf{X}}\right] = \mathbf{1}^{(1\times N)}\left[(\mathbf{I}-\alpha\mathbf{X})^{-1} - \mathbf{I}\right]. \quad (3)$$

The conventional status index, simply the column sums of **X,** was divided by $(N-1)$ as a normalization procedure. Katz took a similar approach, dividing the elements of **s** by the value, $m \cong (N-1)!\alpha^{(N-1)}e^{1/\alpha}$, which accommodated the underlying construct of Katz's new technique (Katz, 1953, pg. 42). Figure 1 illustrates the network analyzed by Katz, consisting of actors A through F.

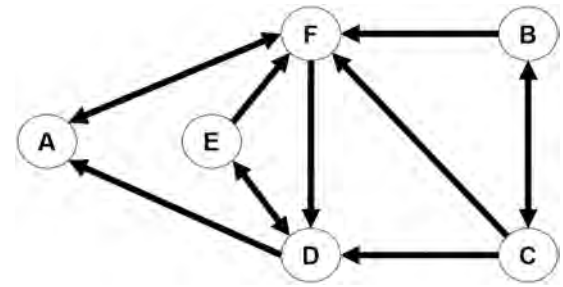

Figure 1. Choice Matrix (Katz, 1953, pg. 40).

The original status vector, one element for each of the six actors, is **s** = [0.4 0.2 0.2 0.6 0.2 0.8]; essentially, actors with higher numbers of ties received, or high in-degree, such as F, D, and A in descending order, dominate with regards to earlier measures ascertaining actor status. Alternatively, the status vector using Katz's measure, with an attenuation factor of $\alpha = 0.5$, is **s** = [0.47 0.04 0.04 0.41 0.22 0.45].

Using Katz's measure, actor A now scores higher than actor F, albeit slightly. Despite the relatively low in-degree of actor A, his status is highest because both of the actors with the highest in-degree, actors F and D, choose actor A. The change in status for actors B, C and E from being equivalent using the original status measure to E differing from B and C using Katz's measure is accounted for in a similar fashion (Katz, 1953, pg. 42). Note that Katz's measure may also be used to estimate actor power simply by analyzing the transpose of the sociomatrix, $\mathbf{X}^t$.

b. Eigenvector-Based Measures. The underlying premise of eigenvector centrality is summarized as,

> Being chosen by a popular individual should add more to one's popularity. Being nominated as powerful by someone seen by others as powerful should contribute more to one's perceived power. Having power over someone who in turn has power over others makes one more powerful (Bonacich and Lloyd, 2001, pg. 192).

Conceptually, this approach is closely related to Katz's measure; mathematically, this measure is described by

$$s_i = x_{1i}s_1 + x_{2i}s_2 + \cdots + x_{ni}s_n, \qquad (4)$$

where an entry in the adjacency matrix, $x_{ij} \in \mathbf{X}$, implies that actor i contributes to j's status and $s_i \in \mathbf{S}$ denotes the status of individual i (Bonacich and Lloyd, 2001, pg. 192–3). In order to determine solutions to this system, the generalized form

$$\lambda s_i = x_{1i}s_1 + x_{2i}s_2 + + x_{ni}s_n \qquad (5)$$

is commonly known and solved as the eigenvalue problem (Bonacich and Lloyd, 2001, pg. 193).

As expected, network structure plays an important role in the results of this analysis method. However, there are unique cases where the numerical results may not capture the more intuitive understanding of centrality. For example, all actors within each of the hypothetical, directed networks shown in Figure 2 have zero status due to ". . . positions that receive no choices have no status and contribute nothing to any other position's status" (Bonacich and Lloyd, 2001, pg. 139).

To deal with this conceptual and mathematical issue, Bonacich and Lloyd proposed 'α-centrality' that allows "every individual some status that does not depend on his or her connection to others" (Bonacich and Lloyd, 2001, pg. 193). With the vector of exogenous sources of status, **e,** and a parameter reflecting the ". . . relative importance of endogenous versus exogenous factors in the determination of centrality" (i.e., α), the matrix solution for status is given by

$$\mathbf{s} = \alpha\mathbf{X}^t\mathbf{s} + \mathbf{e} \Rightarrow \mathbf{s} = (\mathbf{I} - \alpha\mathbf{X}^t)^{-1}\mathbf{e}. \qquad (6)$$

Note that Katz's model (Equation 3) is similar to this approach and simply differs by a constant of one when **e** is a vector of ones (Bonacich and Lloyd, 2001, pg. 194). Although the theoretical development, range, and magnitude of **e** are not discussed, other than being a vector of ones in their example, the new approach both permits analysis of asymmetric relationships and is equivalent to the original formulation in Equation 4 as α approaches the inverse of the largest eigenvalue, $\lambda_{\max}^{-1}$ (Bonacich and Lloyd, 2001, pg. 196–7).

c. Geodesic Measures. Valente and Foreman (1998) developed a dual-purpose measure based upon a reverse geodesic distance approach. Given that the measures of interest are integration ("can be reached by many others rapidly") and radiality ("the degree to which an individual's relations reach out into the network"), the measure is dual-purpose in that the input is either the adjacency matrix or its transpose, respectively (Valente and Foreman, 1998, pg. 90). The integration measure for a given actor k is formally defined as

$$I(k) = \frac{\sum_{j \neq k} RD_{jk}}{N - 1}, \qquad (7)$$

where RD_{jk} is the reverse geodesic distance, computed by subtracting the geodesic distance between j to k from 1 plus the network diameter, D, which is defined as the longest, shortest-path distance between nodes in a network (Wasserman and Faust, 1994, pg. 112; Valente and Foreman, 1998, pg. 92). To calculate radiality, the same measure is simply applied to the transpose of the adjacency matrix (Valente and Foreman, 1998, pg. 93).

Note that there is no attenuation factor associated with longer geodesic paths. In addition, RD_{jk} cannot accommodate multiple instances of geodesic paths between any two given actors. Consequently, radiality may not truly capture "the degree to which an individual's relations reach out into the network" if multiple shortest paths implies more potential

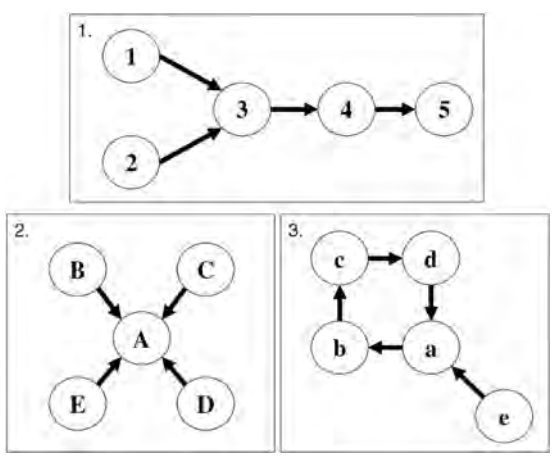

Figure 2. Notional Networks (Bonacich and Lloyd, 2001, pg. 192).

for the exertion of influence or power (Valente and Foreman, 1998, pg. 90).

Lastly, a reach-based measure of centrality that "counts the number of nodes each node can reach in k or less steps" is offered by Borgatti et al. (2002). This too can be applied to directed and undirected networks. However, this particular measure does not accommodate multiple shortest paths and, from the documentation available, the method of attenuation, if any, is not immediately apparent.

A couple points of contention exist: the characteristics of the flow captured or assumed by the related measures and, the somewhat arbitrary choice and allowed range of the 'attenuation' factor. What differentiates RBAP from previously developed measures of power is (1) the use of multiple instances of shortest paths; (2) the process of accounting for any available options to the actors regarding alternative shortest paths; and, (3) uncoupling the concept of 'attenuation' from conditions necessary for a system solution.

Flow Characteristics. The length of a walk, trail, path, or chain is determined simply by summing the lengths of each of its arcs. A trail is "a walk in which all of the lines are distinct, though some nodes may be included more than once" (Wasserman and Faust, 1994, pg. 107). A path is defined as "a walk in which all nodes and (consequently) all lines are distinct" (Wasserman and Faust, 1994, pg. 105). Additionally, given a digraph, the term path implies that the direction of the arcs within the path also follow the direction of the path, otherwise it is a chain (Bazaraa et al., 1990, pg. 422). Paths within a dichotomous network are of primary interest in RBAP. The goal of RBAP is to characterize the relative position of members within a clandestine network by assessing possible transmission paths that may promulgate information or influence via interpersonal communication.

Recall the use of powers of the sociomatrix by Katz and the related eigenvalue-based measures developed by Bonacich (1987) and Bonacich and Lloyd (2001). The elements within the powers of the sociomatrix capture a variety of directed edge sequences that may not necessarily be conducive to operations security. In addition, the length of these sequences within these measures extends to infinity. This 'infinite' amount of communication exchanges between individuals is also counter to a clandestine network's operational security objectives. This suggests that a more direct, geodesic-based approach is more appropriate. Note that for any connected network the longest geodesic with edge lengths of one is bounded above by $(N - 1)$.

Leenders previously pointed out that the information contained within the powers of the sociomatrix is often misunderstood or misperceived, depending upon the operative definition of 'walk' (Leenders, 2002, pg. 32). Using the network definitions discussed above, Deo offers a more precise definition of the content of the powers of the sociomatrix, which is summarized in the following theorem.

Theorem 2.1. The (i, j)th entry in $\mathbf{X}^p$ equals the number of different, directed edge sequences of p edges from the ith vertex to the jth. (Deo, 1974, pg. 222)

Deo also noted that these sequences fall into three categories:

(1) Directed paths from i to j: those directed edge sequences in which no vertex is traversed more than once;
(2) Directed walks from i to j: those directed edge sequences in which a vertex may be traversed more than once, but no edge is traversed more than once; and,
(3) Those directed edge sequences in which an edge may also be traversed more than once (Deo, 1974, pg. 222).

Observe that the second and third categories of information flow are likely detrimental to the security goals of a clandestine network. For example, using Figure 1 with $p = 4$, a possible directed edge sequence of length four between A to D could include A-F-A-F-D. If the network of interest is trying to maintain secretive communications, the banter between A and F may be unlikely or at least unwise (Post, 2005, Chapter 2). Although the measure developed by Katz is easy to implement, this measure captures network behavior that goes beyond the circuitous paths posited by Stephenson and Zelen (1989).

Therefore, RBAP seeks to assess actor position, in any of the analytical categories, by measuring the propagation of network phenomena

via efficient information channels offered by geodesic paths. Further, unlike the radiality, integration, and previous reach-based measures, RBAP also accounts for multiple options, if any, available to actors in a given context (i.e., multiple shortest paths).

Attenuation. The somewhat arbitrary nature of the 'attenuation' factor, particularly in Katz's measure, has been noted in previous works (Clark, 2005; Borgatti and Everett, 2006). Conceptually, the value of α is likened to the 'attenuation' of a signal or influence as a function of distance traveled, which presumably ranges between complete attenuation, $\alpha = 0$, and no attenuation, $\alpha = 1$.

For his status measure, Katz suggests $\alpha \in [(2\lambda_{max})^{-1}, \lambda_{max}{}^{-1}]$ (Katz, 1953, pg. 42). Using the same example discussed by Katz, the largest eigenvalue of **X** for the network corresponding to Figure 1 is 1.68; this implies $\alpha \in [0.298, 0.595]$. Hence, this approach and its underlying assumptions restrict the 'attenuation' space within which the analyst may work from the onset. Depending upon network structure, analysis may be accomplished outside this recommended range. However, acceptable values of α are ultimately required for the underlying geometric series to converge, thereby providing a meaningful result. Otherwise, nonsensical results may occur, despite the fact that α still falls within the reasonable analytical range between 0 and 1.

Interestingly, even within the recommended range the most 'central' actor is dependent upon the value of α. Figure 3 depicts the results of Katz's measure as applied to the notional network in Figure 1, with α varied across the range recommended by Katz. The graph captures the rank order of the status for each of the six actors, with the values 6 and 1 indicating the highest-and lowest-ranking status scores respectively. Interestingly, two crossover points exist, resulting in actors A and D exchanging status rankings at $\alpha \approx 0.36$ and actors A and F exchanging status rankings at $\alpha \approx 0.48$. Such phenomena are clearly of interest; however, unless the network-dependent ranges impinged upon α are normalized, interpretation of such effects is untenable. Similar observations may be made for the eigenvalue-based measures.

From these previous efforts, an opportunity clearly exists to (1) enhance the measurement concepts developed by Katz (1953), Valente and Foreman (1998), Bonacich (1987) and Bonacich and Lloyd (2001); and, (2) separate the concept of 'attenuation' from the con-

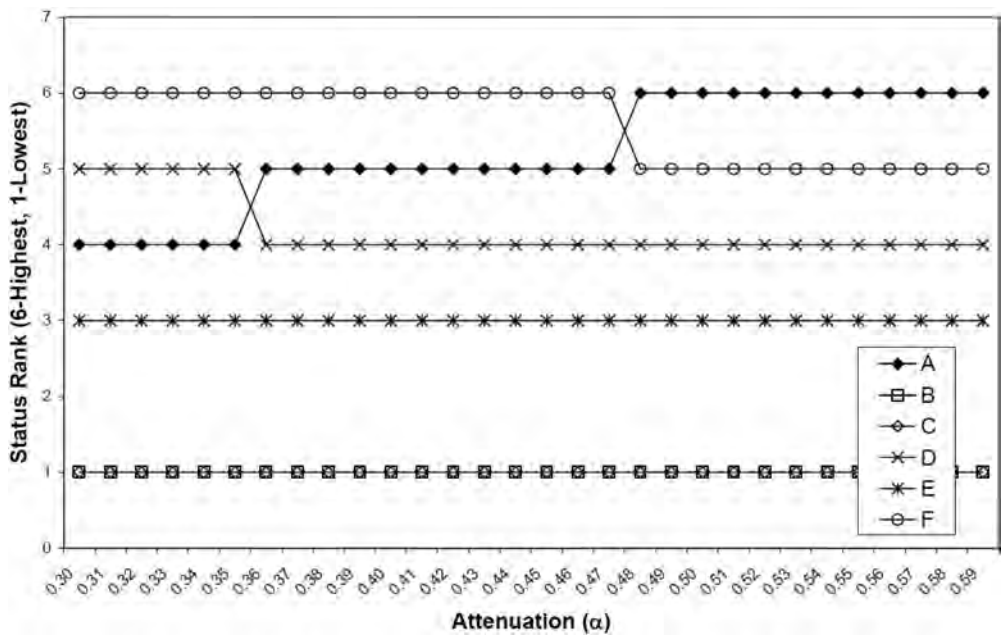

Figure 3. Change in status with attenuation.

ditions required for system solution. These changes offer a means to more closely capture the network phenomena inherent to clandestine networks, as well as improve upon the analytical basis for assumptions dealing with its attenuation as it flows through the network; hence, these comprise the underlying motivations for the RBAP measure.

ASSUMPTIONS AND DEVELOPMENT

Recall that the shortest path between any two individuals of a connected network with N individuals ranges between 1 and $(N - 1)$. Deo's theorem is extended via Corollary 3.1 to enumerate the number of all pair-wise shortest paths, simply by raising the adjacency matrix to powers ranging from 1 to $(N - 1)$. Note also that this RBAP approach assumes a dichotomous representation of the network, thereby relegating the distance of existing interpersonal relationships to 1. Although this clearly cannot accurately represent the heterogeneous nature of interpersonal relationships, the screening nature of RBAP also assumes that very limited information about the actors and their interactions is available.

Corollary 3.1. Given an adjacency matrix $\mathbf{X}$, as it is raised to the pth power, $p = 1, \ldots, (N - 1)$, the first non-zero, (i,j)th element in $\mathbf{X}^p$, $i \neq j$, yields the number of shortest paths of length p from i to j.

Proof. Let x_{ij}^p denote the (i,j)th element in $\mathbf{X}^p$. For each (i,j)th element in $\mathbf{X}^p$, $i \neq j$, if $x_{ij}^p > 0$, and $x_{ij}^k = 0$ for $k = 1,\ldots,(p - 1)$, this implies that no directed edge sequences of length $1, \ldots, (p - 1)$ exist. Therefore, the shortest path between i and j must be of length p. This further implies that the value x_{ij}^p must also fall in the first category stated by Deo, which is the number of directed, shortest paths from i to j.

Use of Corollary 3.1 facilitates the enumeration of shortest paths and their lengths between all actors. The definitions in Table 1 serve as the basis for RBAP. One other underlying assumption of this measure is that the highest level of power is obtained when an actor can reach all other actors within one step. Similar arguments are made for status (e.g., can be reached by all others) and centrality (e.g., is adjacent to all others). Consequently, the numbers provided in the matrices ($R_{\mathbf{x}^l}$) must be normalized to avoid actors with numerous but indirect paths to all others from scoring higher than actors that, in the case of power, can reach all other $(N - 1)$ actors within one step. This is accomplished with the variable $r(k)_l$.

For example, consider the network in Figure 4. Actor i, reaches three other actors via a shortest path of length 1. Therefore, in order to reach any other actor, j, the maximum number of shortest paths of length 2 is bounded above by $r(i)_1 = 3$. If the three dashed paths existed in the network, the value of $R_{\mathbf{x}^2}(i,j)$, the number of shortest paths from i to j of length 2, would be 3. This value and all other values in the ith row of $R_{\mathbf{x}^2}$ are normalized by dividing by $r(i)_1 = 3$.

Suppose further that from node i, two new nodes were reached via a shortest path of length 2 (nodes d and e in Figure 5). Therefore, to reach any node j via a shortest path of length 3, there are at most $3 \times 2 = 6$ possibilities, given by the paths $(i{-}a{-}d{-}j)$, $(i{-}a{-}e{-}j)$, $(i{-}b{-}d{-}j)$, $(i{-}b{-}e{-}j)$, $(i{-}c{-}d{-}j)$, and $(i{-}c{-}e{-}j)$. Consequently, this requires that the value $R_{\mathbf{x}^3}(i,j)$, as well as all other values in the ith row of $R_{\mathbf{x}^3}$, be divided by $r(i)_1 \times r(i)_2 = 6$. To facilitate this calculation, the matrix satisfying the conditions

Table 1. RBAP Definitions

α	An attenuation factor, with a similar, penalizing purpose to that used in Katz (1953); however, for RBAP there is no restriction other than $\alpha \in [0, 1]$
$R_{\mathbf{x}^l}$	$(N \times N)$ matrix that stores the number of shortest paths of length l from any two given actors where the criteria of Corollary 3.1 are satisfied
$r(k)_l$	The number of other actors reached by actor k via a shortest path of length l
$\mathbf{r}_l$	$(N \times N)$ diagonal matrix where, $\forall r(m)_l > 0$, $\mathbf{r}_l(m,m) = r(m)_l^{-1}$ for $m = 1,\ldots,N$; zero otherwise

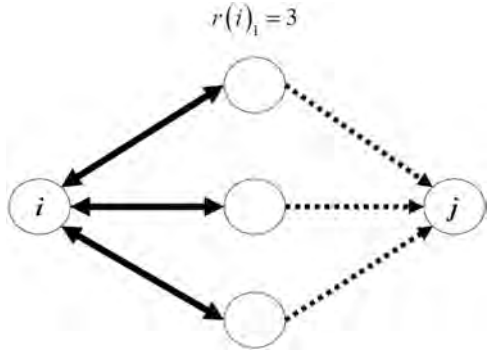

Figure 4. Paths to j given $r(i)_1 = 3$.

of Corollary 3.1 is pre-multiplied by the matrix $\mathbf{r}_l$, defined in Table 1.

An 'attenuation' factor, $\alpha \in [0, 1]$, not unlike those seen in related works, represents the diminishing effectiveness of communication or influence as a function of path length. However, unlike the works of Katz (1953) or Bonacich and Lloyd (2001), calculating the RBAP measure is not predicated upon finding a specific value for α. This offers some analytical freedom, as α can take on any value within its range without negating the measure's results. Of course, deciding the specific value to assign to α is a clear opportunity for further sociological research. However, as demonstrated in the example analysis, there are potential benefits to performing what is essentially sensitivity analysis on actor RBAP scores across the range of α.

In addition, the attenuation is assumed to only apply to indirect contact (i.e., when the shortest path between two individuals is greater than 1). Therefore, RBAP simply reduces to degree centrality (simple, in-, or out-degree depending upon the data and application) when $\alpha = 0$ and is bounded above by the total number of other nodes that can be reached from any given node when $\alpha = 1$. Consequently, the endpoints of this new range now provide a consistent and meaningful interpretation, with $\alpha = 0$ and $\alpha = 1$ measuring actor position from a local and global perspective, respectively. Letting $\mathbf{1}$ be an $(N \times 1)$ vector of ones, the $(N \times 1)$ RBAP measure is defined as

$$\mathbf{RBAP} = [R_{\mathbf{x}} + \alpha \mathbf{r}_1 R_{\mathbf{x}^2} + \alpha^2 \mathbf{r}_1 \mathbf{r}_2 R_{\mathbf{x}^3} + \cdots + \alpha^{N-2} (\prod_{l=1}^{N-2} \mathbf{r}_l) R_{\mathbf{x}^{N-1}}]\mathbf{1}. \quad (8)$$

A MATLAB program provides a sensitivity analysis procedure that applies RBAP to the full range of α. Given the known conditions or interpretations at either end, the varying assumptions regarding the losses (or lack thereof) of communication or influence as a function of path length are easily investigated.

To address the logical question of determining a specific setting for α, a proxy is proposed. Such a proxy for α could include the clustering coefficient of the network, denoted $\gamma(G)$, which is the average of the clustering coefficients for each of the actors within a network. The clustering coefficient for a given actor i is denoted $\gamma_i(i)$. Given the number of neighbors of i (denoted b_i), the individual-specific clustering coefficient is the "ratio of actually existing connections between the b_i neighbors and the maximal number of such connections possible $(b_i^2 - b_i)$" (Sporns, 2002, pg. 178) (cf. Watts, 1999, pg. 32–3). Consequently, higher clustering coefficients may imply more cohesive and interactive groups and therefore lower communication or influences losses, corresponding to higher values of α.

Although not a necessary condition to perform the calculations, application of this procedure assumes that the network of interest is connected. Considering that this measure is reach-based, the centrality calculated for isolates-actors with no or very few links to other individuals-is zero, as expected (Brass, 1995,

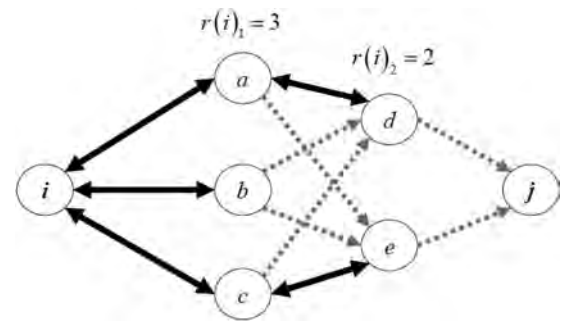

Figure 5. Paths to j given $r(i)_1 = 3$ and $r(i)_2 = 2$.

pg. 46). If the graph is comprised of more than one component, all output will be relative to the specific components and not to the network in total. Subsequently, caution must be taken to avoid misinterpretation of the output by unknowingly comparing results among two or more components rather than across all actors, particularly if the values are normalized. If the graph is comprised of several components, analysis should be accomplished on the specific components of interest, rather than applying this measure to a number of disconnected components simultaneously.

Finally, since the RBAP value for any given actor is bounded above by $(N - 1)$ regardless of α, actor RBAP scores may be normalized, yielding a number between 0 and 1. This version of the RBAP is given by

$$\mathbf{RBAP}' = \frac{\mathbf{RBAP}}{N-1}. \tag{9}$$

Without normalization, the interpretation of RBAP is the number of other actors that can be effectively communicated with, persuaded, influenced, and so forth, ranging in value between the number of an actor's immediate contacts to all other individuals within the network $(N - 1)$. With normalization, the interpretation is similar, but is in the context of percent of the other $(N - 1)$ actors. Consequently, the output can be tailored to meet the needs of decision makers, who may prefer percentages versus having to compare RBAP scores against the value $(N - 1)$. Some examples are now offered to discuss the resulting nature of RBAP, using the non-normalized version.

DISCUSSION

As an initial comparison, RBAP was applied to the transpose of the choice matrix specified by Katz (1953). This permits a comparison between Katz's status results and the prestige (as opposed to power) use of RBAP. With $\alpha = 0.5$, the Katz and RBAP status rankings are shown in the first two rows of Table 2.

Observing that there are similarities, and differences, between the two approaches, a more equitable comparison between the two methods was attempted. Recall that in Equation 1, Katz allowed infinite path lengths. As aforementioned, considering that the context of interest is the measurement of influence or communication among clandestine networks, this may be an unrealistic assumption. Instead, suppose the path length limit of $(N - 1)$ used by RBAP was imposed upon Katz's measure while still normalizing by the original definition of *m.* Let this adjusted measure be defined as

$$\mathbf{s}_{adj} = \mathbf{1}^{(1\times N)} \sum_{l=1}^{N-1} \alpha^l \mathbf{X}^l. \tag{10}$$

The results, shown in the third row of in Table 2, offer an improved comparison to RBAP. The differences are essentially due to the Katz's inclusion of directed edge sequences other than shortest paths. However, given the underlying differences between the measures, perfect correlation between RBAP and any other existing, path-based measure was not one of the research objectives.

Applying the sensitivity analysis procedure for RBAP to all three hypothetical networks discussed by Bonacich and Lloyd (2001) (see Figure 2) yields the results shown in Figure 6 which are as expected. For example, from the perspective of radiality or power, actors 1 and 2 in Network 1 are more effective than all others in reaching out to the remaining actors. Whereas, actor 5 has no outward connections

Table 2. Katz and RBAP Comparison ($\alpha = 0.5$).

Method	Rank (Value)				
	High	←		→	Low
Katz	A (0.47)	F (0.45)	D (0.41)	E (0.22)	B and C tie (0.04)
RBAP	F (4.25)	D (3.50)	A (3.25)	E (2.37)	B and C tie (1.00)
Katz (adj)	F (0.25)	A (0.24)	D (0.22)	E (0.11)	B and C tie (0.04)

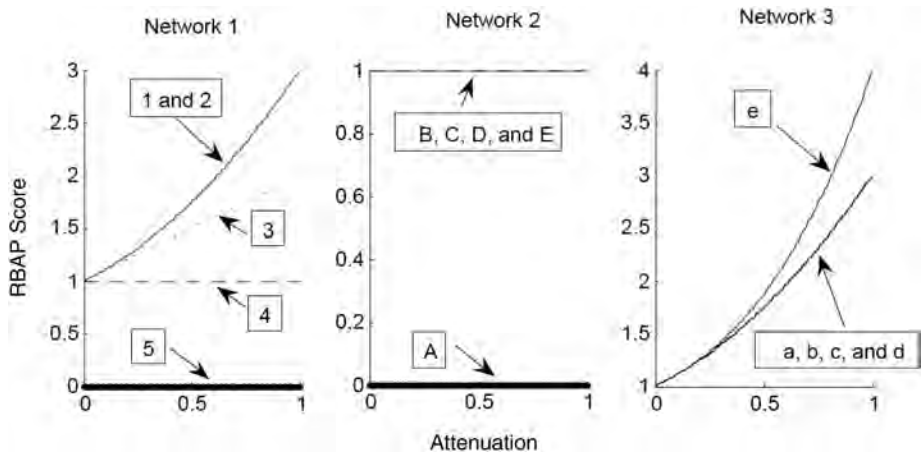

Figure 6. RBAP applied to Bonacich and Lloyd (2001) networks.

and therefore has no capacity to influence others.

Note that the original purpose for RBAP was to determine the individuals who are potentially the most influential or powerful, given limited information about the network's members and their interconnections. However, the flexibility of RBAP permits the analysis of any analytical category (power, status, and centrality) which is determined simply by the nature of the data input.

A logical concern regarding the implementation of RBAP is that of computational efficiency. From Equation 8, the time required for calculation is dominated by the term, R_{X^1}, which is worst-case $O(N^3)$. Since the measure calculations are complete when all actors have been reached, the worst-case times required for evaluating a given network are predicated upon the network's diameter. To quantify this effect, RBAP was applied to a number of line graphs (as shown in Figure 7), ranging in size from $N = 10 \ldots 1330$, so that the measure must continue to the largest diameter possible, $(N - 1)$. (Note that the limit of 1300 nodes was due to memory limitations; all tests were performed on a 3.4 GHz Pentium 4 with 1GB RAM running Windows XP Pro).

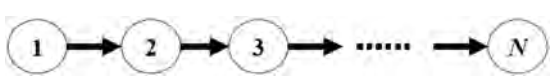

Figure 7. Line graph of size *N*.

The performance in seconds is compared to *N* in Figure 8. The solid line in Figure 8 represents a polynomial of degree three fit to the data; this can be used as a rough estimate of the worst-case time required to compute the RBAP measure given a network of size *N*. Noting that (1) the polynomial is increasing substantially with *N* and (2) that the size of clandestine networks, particularly terrorist networks, can be much larger than 1330 individuals, worst-case run times may be prohibitive. Unfortunately, this limitation is also shared by other social network analysis approaches, which use $O(N^3)$ algorithms to determine related measures such as all-pairs shortest paths and reachability (e.g., Cyram, 2004).

However, the line graph represents an extreme, and unlikely, topology of a social network, even if the members are engaged in clandestine activity. As an example, the trusted prior contacts of the 9–11 hijacker network analyzed by Krebs had 19 (known) individuals; the diameter of this network, based upon the relationships ascertained from open source data, was 9 (Krebs, 2002, pg. 46). The relationship between population size and network diameter has been of interest since Milgram traced correspondence paths, wherein the famous 'six degrees of separation' between ostensibly distant and unconnected actors was observed (Milgram, 1967). Such a 'six-degree' graph would yield a variation of the polynomial in Figure 8 but would result in dramati-

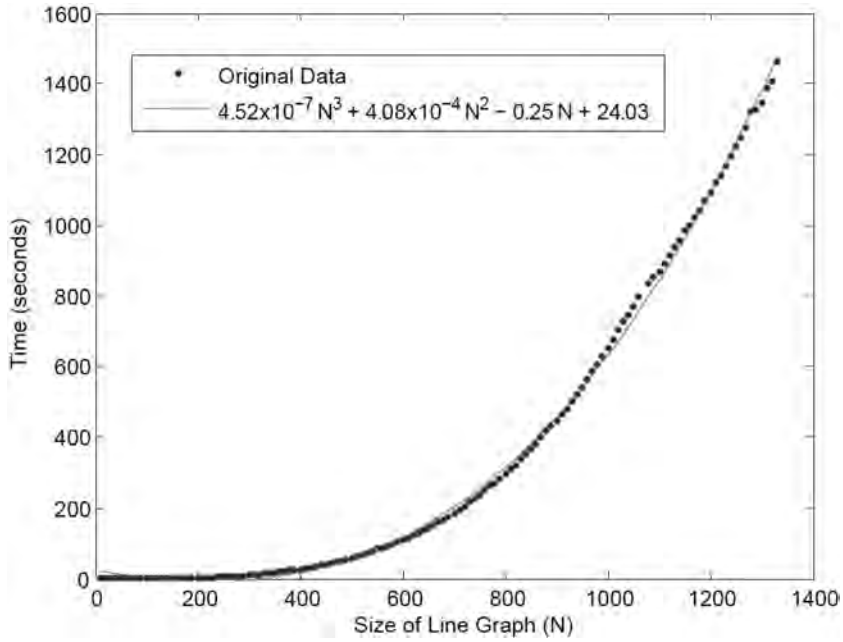

Figure 8. Network Size (N) versus RBAP runtime.

cally reduced computational requirements. Several works have popularized this phenomenon, termed the 'small-world' property (Barabási, 2002; Buchanan, 2002; Watts, 1999).

Numerous connections between real-world, emergent networks and small-world network behavior have been made. Examples include Internet connections, cellular metabolism, Hollywood movie-stars, protein regulatory networks within cells, research collaborations, social networks, and sexual relationships (Buchanan, 2002; Barabási and Bonabeau, 2003, pg. 54). As a result of the small-world property, "their diameter is $O(\log N)$ instead of $O(N)$" (Eppstein and Wang, 2004, pg. 40). Similar findings have been made in analyzing networks evolving via preferential-attachment mechanisms described by Barabási and Albert (1999) (Liben-Nowell, 2005, pg. 16–8). In addition, more recent research by Leskovec et al. (2005) has shown diameter to actually decrease with increased network size. These observations translate directly to corresponding savings in RBAP computational performance. Figure 9 summarizes the run time required to perform the RBAP measure for networks ranging from 100 to 1400 nodes with varying diameters as opposed to the worst-case diameter of $(N - 1)$. As expected, if D is much less than N, then the computation time required is significantly reduced. For example, the 1300-actor network with $D = 1299$ required 1344.9 seconds to complete. A comparable 1300-actor network with $D = 30$ required 38.8 seconds. Therefore, in lieu of real-world, large, terrorist network data sets, initial experimental results suggest that this is a promising approach with regards to computational efficiency.

Of course, the equivalence between social networks and network data gathered to characterize actors and relationships enmeshed within clandestine activity remains an open question. The object of study is still comprised of people with links indicating some form of interaction. Fortunately, previous authors have addressed some of the issues that often plague the application of social network analysis techniques to clandestine networks. For example, several efforts have studied the implications of network sampling upon classic centrality measures using social network data (Costenbader and Valente, 2003) and random networks (Borgatti et al., 2006); the former concluding that the stability of measures is dependent upon network topology, and the latter indicating stability, using random graphs, is somewhat predict-

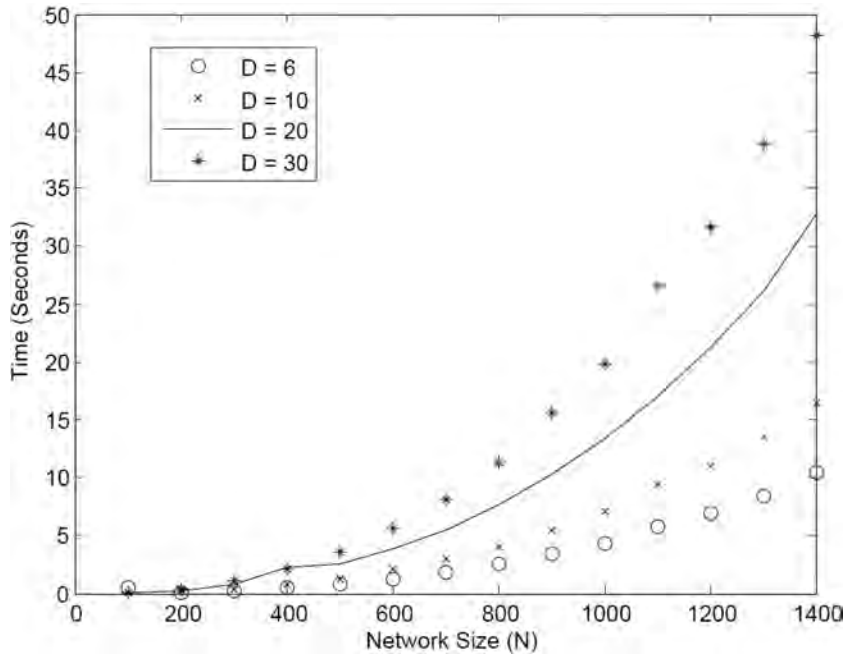

Figure 9. Impact of Diameter upon RBAP runtime.

able, particularly for denser networks. In addition, there is increasing interest in applying social network analysis techniques to terrorist organizations (cf. Krebs, 2002; Carpenter et al., 2002; Carley et al., 2002; Fellman and Wright, 2003; Thomason et al., 2004, to name a few) Consequently, for this paper we assume that clandestine networks are indeed social in nature and will ultimately exhibit the small-world property such that the network diameter will

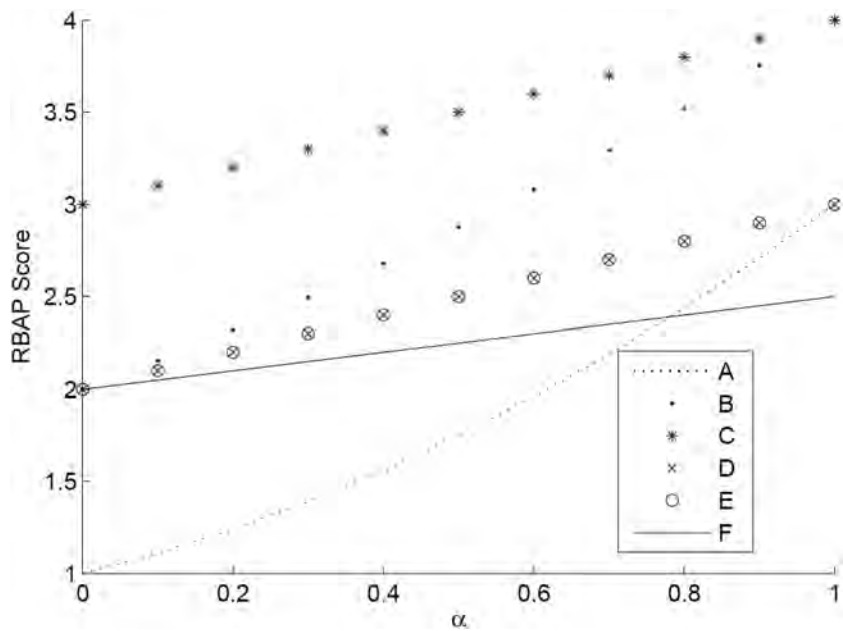

Figure 10. RBAP and Katz Network.

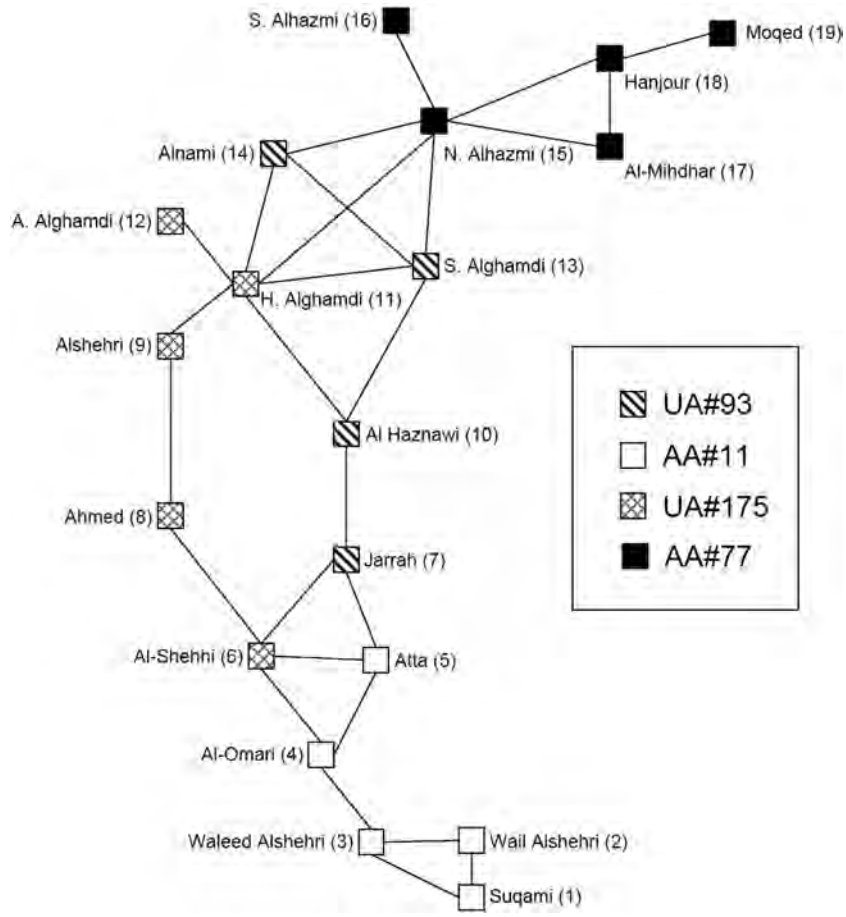

Figure 11. Trusted Prior Contacts (Krebs, 2002, pg. 46).

indeed be much less than the number of actors within it. Since the practical computational bounds of RBAP are heavily dependent upon network diameter, this property alone will contribute to improved performance, given 'reasonably-sized' networks.

Given that the underlying motivation for RBAP is to provide a means to identify potential actors within an adversarial network that exhibit greater amounts of power or influence among the others (i.e., leaders, potential leaders, coordinators, liaisons, etc.), an analysis of the hijacker network presented by Krebs (2002) is of interest. Of course, one must always consider that the adversarial network is constantly trying to either avoid detection or steer our resources in their favor (cf. Sparrow, 1991; Xu and Hsinchun, 2004; Baumes et al., 2004).

EXAMPLES

We first highlight the usefulness of the sensitivity analysis code developed for this research by applying it to the Katz network shown in Figure 1. The results in Figure 10 show similar behavior to Katz's measure in that the most influential individual is dependent upon the level of attenuation selected.

However, for RBAP all values of α provide valid results, given that the attenuation level is justified by careful analysis of the network as a whole. Note that for this application, $\alpha = 0$ reverts RBAP to simple out-degree centrality; whereas, at $\alpha = 1$, the RBAP scores are bounded above by the number of reachable actors. The most influential actors, B and C, are able to reach all other actors but have limited

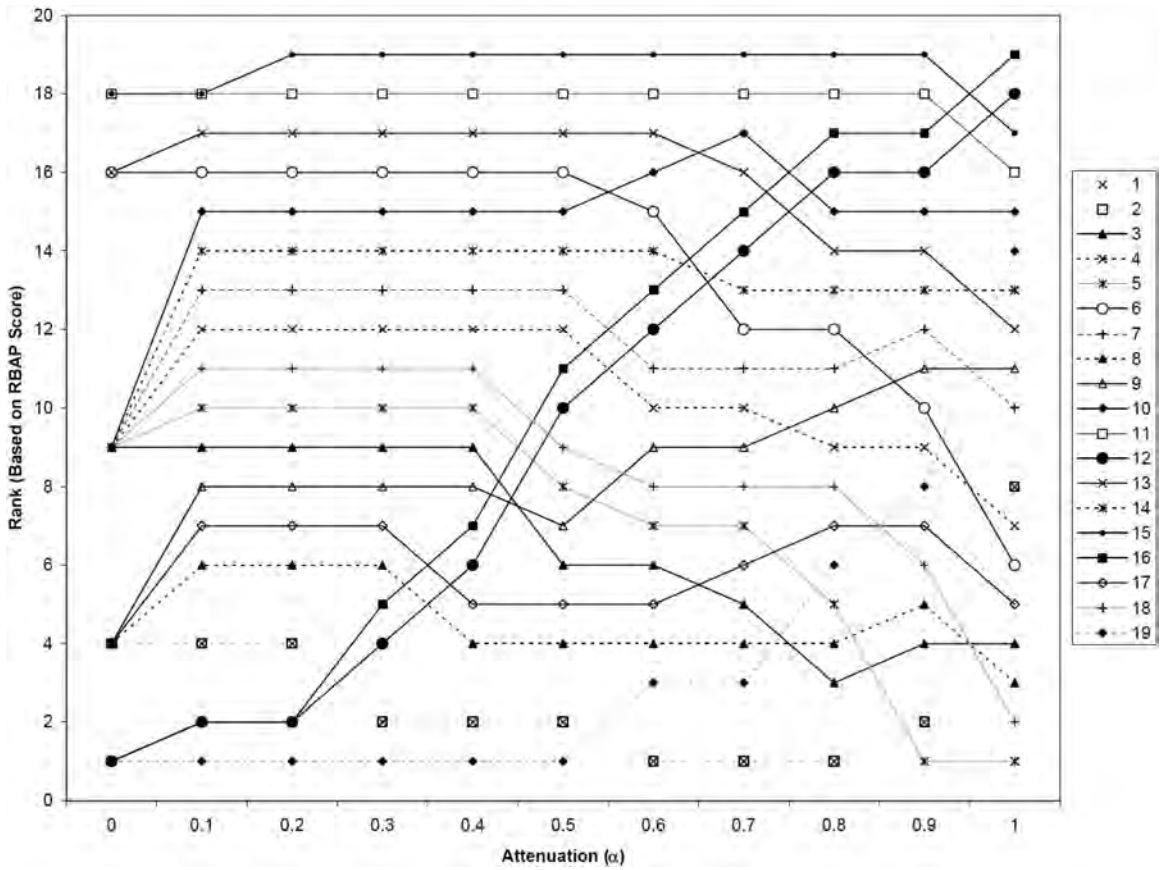

Figure 12. RBAP and Hijacker Network.

options or alternative geodesic paths to do so. Such topological consequences are captured by $\mathbf{r}_I$ which explains the resulting scores being less than $(N - 1 = 5)$ for these actors.

In addition to providing the traditional type of information to a decision maker, the sensitivity analysis approach also offers a means to quickly screen actors, identifying those that are potentially of interest, based upon their network positions within a local and global perspective.

Turning now to the 9–11 hijacker network studied by Krebs (2002), this screening technique is demonstrated. This data comprises the network of trusted, prior contacts among the hijackers, extracted from open source information by Krebs; the network is shown in Figure 11. To facilitate analysis an identifying number, shown in parenthesis by each hijacker's name, was added. Note also that the resulting graph is undirected; therefore, applying RBAP to this data is in the context of centrality or radiality rather than power or influence.

An initial look at the rank orderings based upon RBAP scores and varying levels of α are provided in Figure 12; higher RBAP scores result in higher rankings, ranging from low (1) to high (19) for this network.

As observed with the Katz data, determining the most central individuals according to the RBAP measure is predicated upon the amount of attenuation assumed. Mohammed Atta (actor 5), the purported ring leader, is initially tied with seven other individuals, all having a degree of 3, for rank 9. However, as α is increased to 1, meaning less attenuation with longer paths, Atta's rank goes down substantially; this effect is illustrated in Figure 13. Crossovers such as these may reveal individu-

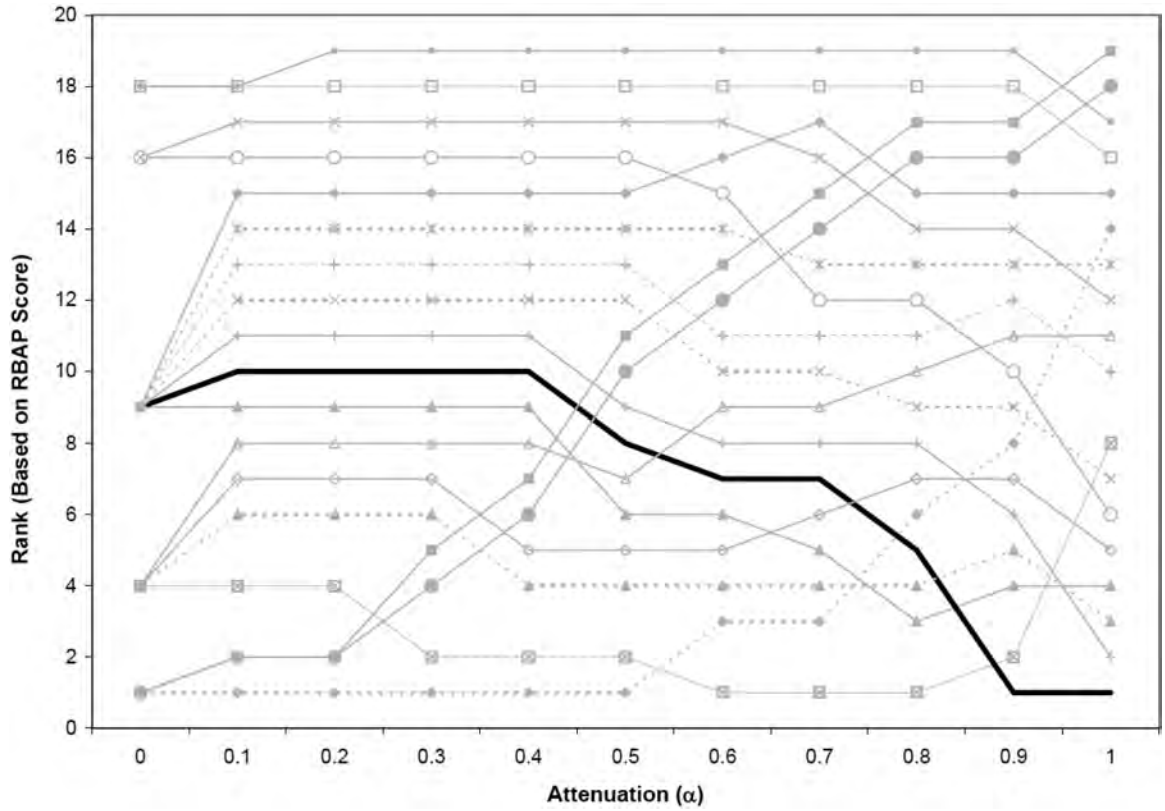

Figure 13. Mohammed Atta (5) RBAP Rankings.

als that are strategically connected from a local perspective, but effectively cut off from the remainder of the network from a global perspective. Such an individual could serve as a cell or team coordinator practicing good operational security techniques.

Other interesting results from Figure 12 are those actors who remain low (or high) regardless of the level of α as well as those who start low at $\alpha = 0$ and move up with increasing α. Consider actors 1 and 2—Suqami and Wail Alshehri—whose RBAP measures tend to stay relatively low over the range of α; this effect is highlighted in Figure 14. These terrorists are not only in the periphery of the network, but they are both connected via actor 3, Waleed Alshehri, who is also somewhat isolated from the network and whose RBAP measure exhibits similar behavior to that of Atta (decreasing with α).

In contrast, two other apparently isolated actors, A. Alghamdi (12) and S. Alhazmi (16), are connected directly to two of the most central actors (11 and 15) from a betweenness, information centrality, eigenvector, and Katz perspective (Verified by Cyram, 2004). Consequently, despite the low degree of actors (12) and (16), they are connected directly to the core of the network which significantly improves their corresponding RBAP scores as the impact of attenuation is diminished (see Figure 15).

Note also that the most central actors, H. Alghamdi (11) and N. Alhazmi (15) not only begin with a high rank (due to their high degree) but maintain their relatively high ranking throughout the range of α, as seen in Figure 16.

Clearly, the size of the sample network lends itself to inferences of actor roles by inspection. As networks grow in size and connectedness, potential targets or individuals of interest are frequently not intuitively obvious. Considering this, RBAP may be used to search for actors with similar patterns as those highlighted within this analysis. Thus, a new mea-

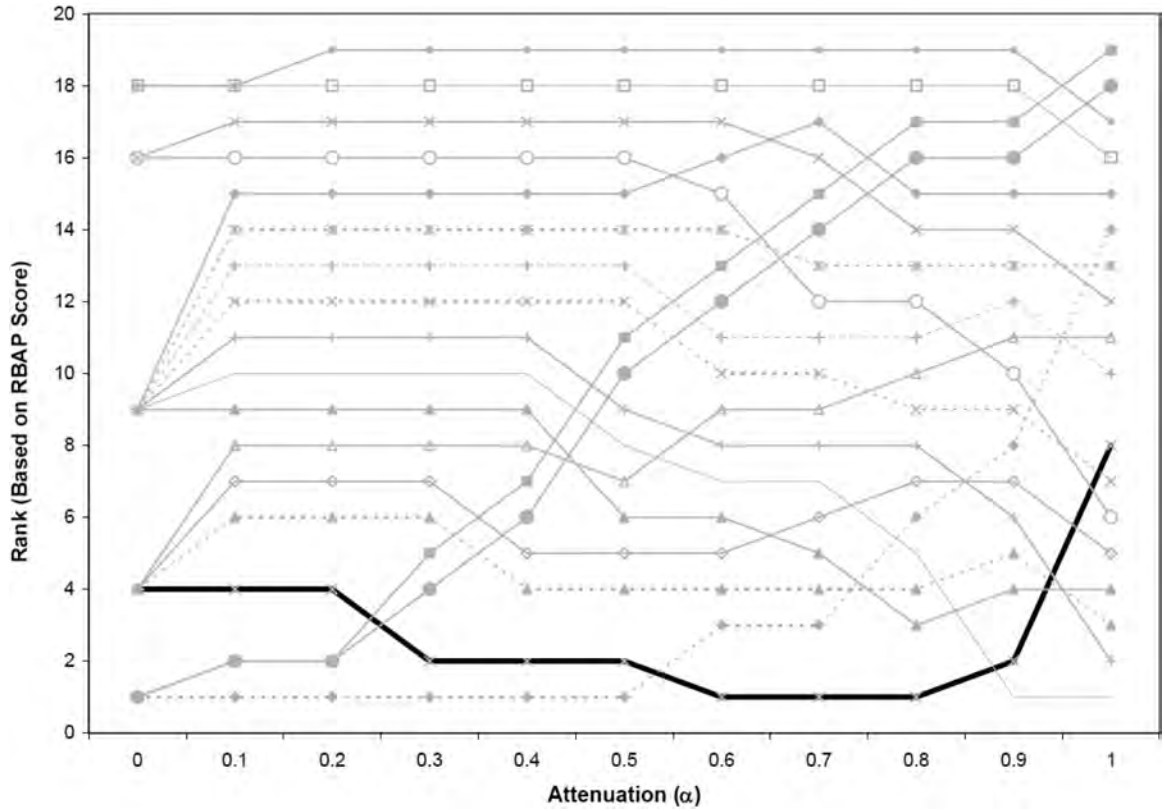

Figure 14. Low-Scoring Peripheral Actors, Suqami (1) & Wail (2).

sure is added to the SNA arsenal, assisting the warfighter in finding the proverbial needle in a haystack, and supporting the prosecution of adversarial networks.

CONCLUSIONS AND RECOMMENDATIONS

The measure presented shares some aspects of other walk-and geodesic-based approaches to gaining insight into an actor's potential for influence or power based upon their position within a given network. However, RBAP provides more analytic freedom regarding the common assumption of attenuation as a function of distance between individuals. The small-world property often inherent to social networks provides a degree of computational efficiency to the measure. Consequently, assuming that the network of interest is reasonably sized (e.g., 3000 actors or fewer) this measure should be responsive to changing information (an estimate solely based upon memory limitations of the hardware used for this research).

The intended purpose for RBAP is to facilitate the investigation of adversarial non-cooperative networks, particularly if the network consists of large number of actors. Actors of interest may be identified by consistently high (or low) RBAP scores as well as those that improve or decrease significantly with a corresponding change in α. Those individuals that are identified through this process can then be subject to increased intelligence scrutiny, either to improve the accuracy of the network data, or to set the stage to affect the organization for political purposes.

Such political endeavors often involve persuading an organization to change their position on a given issue, to modify their inherent approach used to achieve their goals, or to even

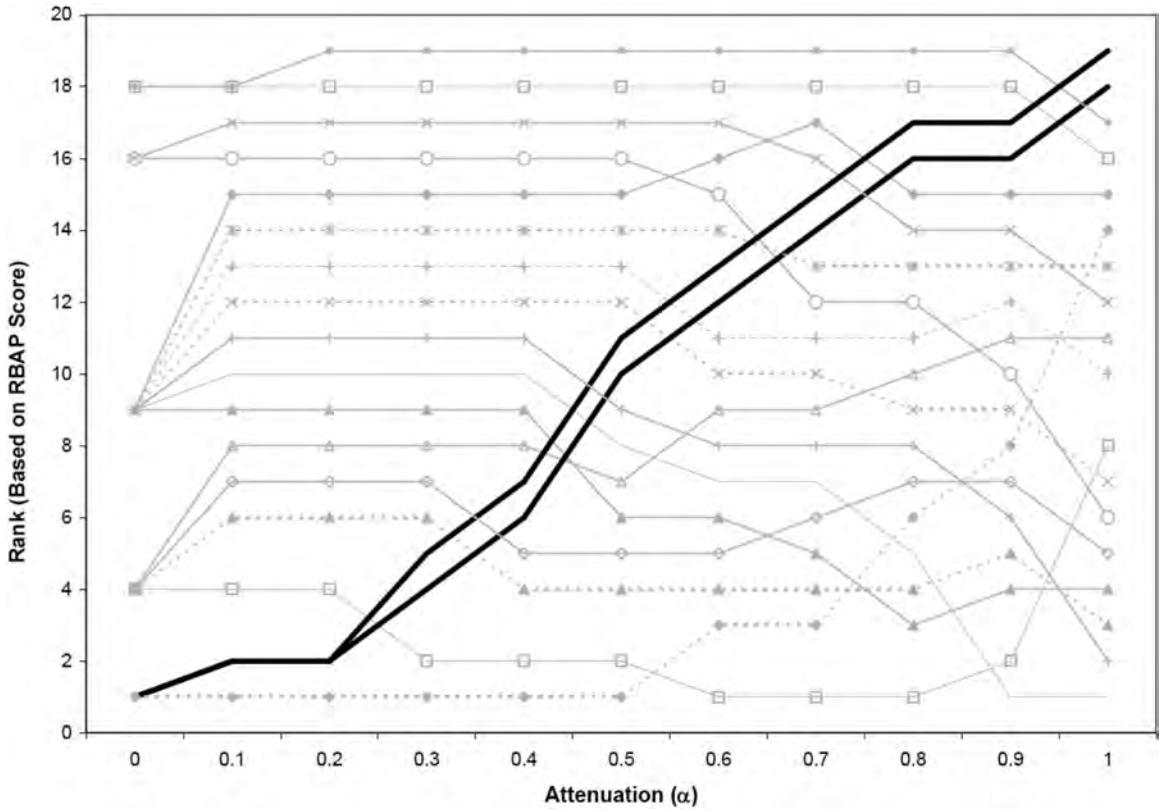

Figure 15. Well-Connected Peripheral Actors, Alghamdi (12) & Alhamzi (16).

disband entirely. Given that an adequate amount of information regarding the individuals and their associated relationships has been obtained, courses of action to achieve these political endeavors could include persuading the entire organization from within. For example, assuming the clandestine network is adversarial, one must first determine those individuals that are accessible. Among this set, those with higher RBAP scores, and who are consequently more effective at reaching or influencing others, would make attractive participants of collusion.

Although α has been specified as a scalar to this point, a possible extension of this measure could incorporate a matrix of individual-specific attenuation factors. Therefore each individual i would be assigned an attenuation factor, α_i. The scalar α in Equation 8 would simply be replaced by the $(N \times N)$ diagonal matrix where $A(i,i) = \alpha_i$, zero otherwise. A possible means to estimate these values could be derived via a decision analytic model using the five bases for power—attraction, expert, reward, coercive, and legitimate—specified by (French, 1956, pg. 183–5) or individual characteristics such as charisma, appearance, and so forth. This data could then be used to gain insight into the effort required to discredit (or support) a specific individual, thereby diminishing (or increasing) their relative power or influence within the network. Holding all other individual's attenuation factors constant, sensitivity analysis of the attenuation factor of the individual of interest would yield the concomitant change in power structure based upon the RBAP scores.

In addition, it is important to note that RBAP is not solely relegated to the study of networks comprised of individuals and relations. Relationships and interactions among individuals, among organizations, among resources, or among a mix of such 'networked' elements may also be analyzed using RBAP, as

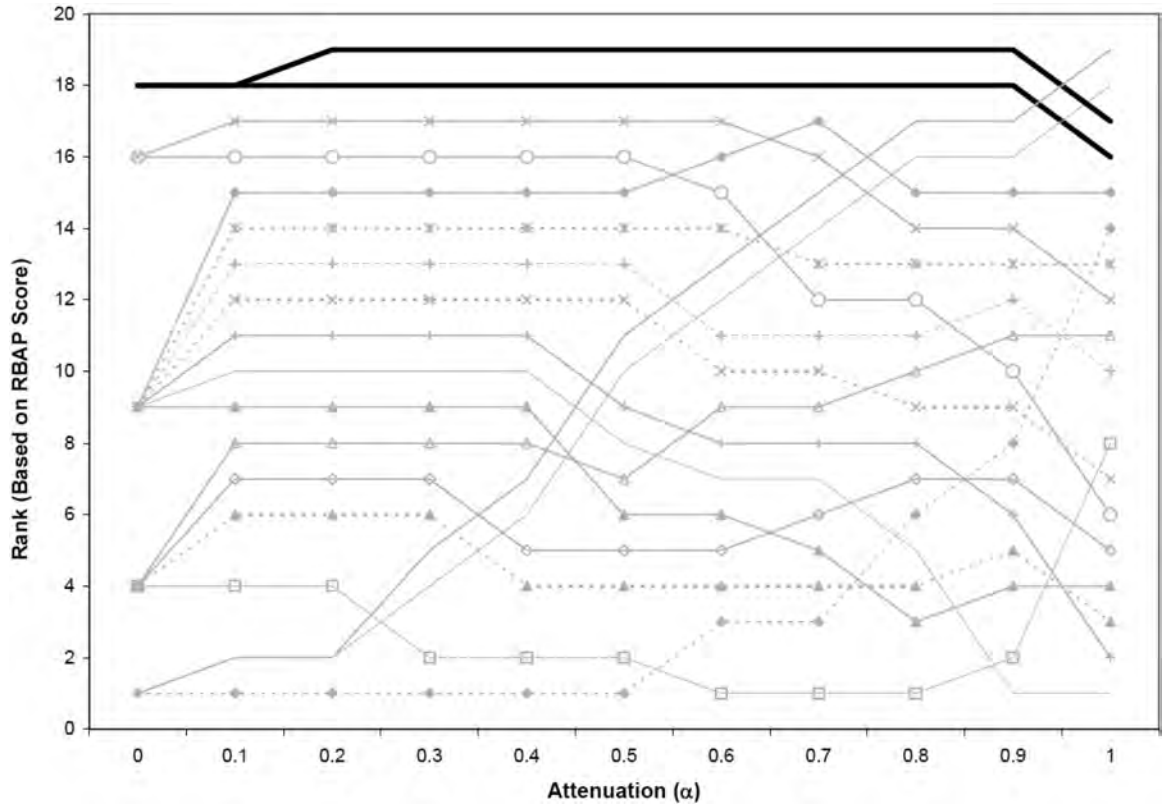

Figure 16. Central Actors, H. Alghamdi (11) & Alhamzi (15).

long as meaningful constructs quantifying their relationships can be identified (cf. Carley et al., 2002).

Other future research activities should investigate improvements in quantifying the robustness of RBAP to missing or incorrect data. The use of random networks (i.e., random graphs) to study other social network phenomena (and their representative measures) has been attempted, but with mixed results (cf. Newman et al., 2002; Newman, 2003; Borgatti et al., 2006; Watts, 1999). Robustness of classical network centrality measures given data errors such as "... edge deletion, node deletion, edge addition, and node addition" has been explored by (Borgatti et al., 2006). Unfortunately, the underlying graphs used in their experiments were random in nature, as opposed to a more representative small-world network topology (Newman et al., 2002, pg. 2571). Nonetheless, the authors concluded that responses to error were ultimately a function of error type and network density (Borgatti et al., 2006). Although similar findings using 'real' and experimental network data are provided in Bolland (1988), the redundant nature of the data may have biased the experimental results. The sensitivity of other, more general, network measures such as global efficiency, critical path length, density, diameter, and radius of scale-free graphs has also been explored by Barabási and Bonabeau (2003) and Thomason et al. (2004). Whether these conclusions map to more appropriate network topologies remains to be seen and, based upon the analysis of network disruption as seen in Albert et al. (2000), is likely heavily dependent upon where the missing data lies within the network. Fortunately, RBAP may be quickly recalculated in the event that changes to existing data or new inputs occur.

From a counter-terrorism perspective, the RBAP measure offers another means to gain insight into adversarial, clandestine networks

such as al Qaeda, Ansar al Islam, and the many others that threaten the existence of peace throughout the globe. Due to the secretive nature inherent to these organizations, methods that provide useful information despite limited or uncertain data are of interest. From a social networks perspective, this measure is not intended to be a direct competitor to the numerous, classical measures in existence, but a complement to enhance the study of network data.

DISCLAIMER

The views expressed in this article are those of the authors and do not reflect the official policy or position of the United States Air Force, the Department of Defense, or the United States government.

ACKNOWLEDGEMENT

This article is based upon research jointly sponsored by the National Air and Space Intelligence Center (GTRB) and the Air Force Research Laboratory (RHAL), whom the authors wish to thank for their support. The authors would also like to thank the anonymous reviewers and Dr. Ed Pohl, who acted as Editor for this article, for their insightful comments and support.

REFERENCES

Albert, R., Jeong, H., Barabási, A., 2000. Error and attack tolerance of complex networks. Nature 406, 378–382.

Baker, W. E., Faulkner, R. R., 1993. The social organization of conspiracy: Illegal networks in the heavy electrical equipment industry. American Sociological Review 58 (6), 837–860.

Barabási, A., 2002. Linked: The New Science of Networks. Perseus Publishing, Cambridge.

Barabási, A., Albert, R., Oct 1999. Emergence of scaling in random networks. Science 286 (5439), 509–512.

Barabási, A., Bonabeau, E., 2003. Scale-free networks. Scientific American, 50–59.

Baumes, J., Goldberg, M., Magdon-Ismail, M., Wallace, W. A., Jan 2004. Discovering hidden groups in communication networks. Lecture Notes in Computer Science 3073, 378–389.

Bazaraa, M. S., Jarvis, J. J., Sherali, H. D., 1990. Linear Programming and Network Flows. John Wiley and Sons, Hoboken.

Bolland, J. M., 1988. Sorting out centrality: An analysis of the performance of four centrality models in real and simulated networks. Social Networks 10, 233–253.

Bonacich, P., Lloyd, P., 2001. Eigenvector-like measures of centrality for asymmetric relations. Social Networks 23, 191–201.

Bonacich, P. B., 1987. Power and centrality: A family of measures. American Journal of Sociology 92, 1170–1182.

Borgatti, S. P., Carley, K. M., Krackhardt, D., 2006. On the robustness of centrality measures under conditions of imperfect data. Social Networks 28, 124–136.

Borgatti, S. P., Everett, M. G., 2006. A graph-theoretic perspective on centrality. Social Networks 28, 466–484.

Borgatti, S. P., Everett, M. G., Freeman, L. C., 2002. UCINET for Windows: Software for Social Network Analysis. Analytic Technologies, Harvard, MA.

Brass, D. J., 1995. A social network perspective on human resources management. Research in Personnel and Human Resources Management 13, 39–79.

Buchanan, M., 2002. Nexus: Small Worlds and the Groundbreaking Theory of Networks. W. W. Norton & Co., New York.

Carley, K. M., Lee, J., Krackhardt, D., 2002. Destabilizing networks. Connections 24 (3), 79–92.

Carpenter, T., Karakostas, G., Shallcross, D., 2002. Practical issues and algorithms for analyzing terrorist networks. Invited Paper, WMC 02.

Clark, C., 2005. Modeling and analysis of clandestine networks. Master's thesis, Air Force Institute of Technology.

Costenbader, E., Valente, T., 2003. The stability of centrality measures when networks are sampled. Social Networks 25, 283–307.

Cyram, 2004. NetMiner II. Vol. Ver. 2.5.0. Cyram Co., Ltd, Seoul.

Deo, N., 1974. Graph Theory with Applications to Engineering and Computer Science. Prentice-Hall, Inc., Englewood Cliffs.

Eppstein, D., Wang, J., 2004. Fast approximation of centrality. Journal of Graph Algorithms and Applications 8 (1), 39–45.

Fellman, P. V., Wright, R., 18 Sep 2003. Modeling terrorist networks: Complex systems at the mid-range. In: Complexity, ethics and creativity conference, London school of economics Edition.

French, J. R., 1956. A formal theory of social power. Psychological Review 63, 181–184.

Goldhamer, H., Shils, E. A., 1939. Types of power and status. The American Journal of Sociology 45 (2), 171–182.

Katz, L., 1953. A new status index derived from sociometric data analysis. Psychometrika 18, 34–43.

Krebs, V. E., 2002. Mapping networks of terrorist cells. Connections 24 (3), 43–52.

Leenders, R., 2002. Modeling social influence through network autocorrelation: Constructing the weight matrix. Social Networks 24, 21–47.

Leskovec, J., Kleinberg, J., Faloutsos, C., 2005. Graphs over time: Densification laws, shrinking diameters and possible explanations. In: Proceedings of the 11th ACM SIGKDD Intl. Conf. on Knowledge Discovery and Data Mining. ACM Press, New York, NY, pp. 177–187.

Liben-Nowell, D., 2005. An algorithmic approach to social networks. Ph.D. thesis, Massachusetts Institute of Technology.

Milgram, S., 1967. The small world problem. Psychology Today 22, 61–67.

MIPT, 2006. Memorial Institute for the Prevention of Terrorism (MIPT) Terrorism Knowledge Base. URL http://www.tkb.org/Home.jsp

Newman, M., Watts, D., Strogatz, S., 2002. Random graph models of social networks. PNAS 99 (90001), 2566–2572.

Newman, M. E. J., 2003. The structure and function of complex networks. SIAM Review 45, 167–256.

Post, J. M., 2005. Military Studies in the Jihad Against the Tyrants: The Al-Qaeda Training Manual. USAF Counterproliferation Center, Maxwell Air Force Base.

Sparrow, M. K., 1991. The application of network analysis to criminal intelligence: An assessment of the prospects. Social Networks 13, 251–274.

Sporns, O., 2002. Neuroscience Databases: A Practical Guide. Klüwer, Ch. 12: Graph theory methods for the analysis of neural connectivity patterns, pp. 169–183.

Stephenson, K., Zelen, M., 1989. Rethinking centrality: Methods and examples. Social Networks 11, 1–37.

Thomason, B. E., Coffman, T. R., Marcus, S. E., 2004. Sensitivity of social network analysis metrics to observation noise. In: 2004 Aerospace Conference. Vol. 5. IEEE, pp. 3206– 3215.

Valente, T. W., Foreman, R. K., 1998. Integration and radiality: Measuring the extent of an individual's connectedness and reachability in a network. Social Networks 20, 89–105.

Wasserman, S., Faust, K., 1994. Social Network Analysis: Methods and Applications. Cambridge University Press, Cambridge.

Watts, D., 1999. Small Worlds: The Dynamics of Networks between Order and Randomness. Princeton University Press, Princeton.

Xu, J. J., Hsinchun, C., 2004. Fighting organized crimes: Using shortest-path algorithms to identify associates in criminal networks. Decision Support Systems 38 (3), 473–487.

Modelling The Rural Infantry Battle: The Effects Of Live Combat On Military Skills And Behaviour During The Approach Phase

L. R. Speight

Cranfield University, Defence College of Management & Technology
rspeight.pheon@ukgateway.net

D. Rowland

Newman & Spurr Consultancy Ltd.

APPLICATION AREA
Land and Expeditionary Warfare

OR METHODOLOGY
Simulation

INTRODUCTION

The previous paper in this series (Speight & Rowland, 2006) set out a generic structure for the rural infantry battle and provided details of a model of the approach phase of such a battle. This model was closely aligned to the evidence gathered from a series of instrumented peacetime exercises featuring trained troops in realistic tactical settings. However, there is a world of difference between the physical and psychological environments encountered in even the most realistic of trials and those that are normally faced by soldiers in genuine combat. The paper just referred to lists several factors that will have a major impact on infantry battle outcomes. Yet Rowland (1987, 2006) and Rowland & Richardson (1997) have shown that, even when the effects of such influences are allowed for, historical analysis reveals a severe degradation in combat performance compared to that achieved in trials. For infantrymen in the defence the kill rates achieved in trials are something like a factor of ten down on those that could be expected based on range results. Further analysis by these two authors suggests that the rates actually achieved in wartime conditions are a further factor of ten down on those achieved in trials. Doubtless combat models based on idealised portrayals of peacetime performance may serve many useful purposes. However, if military OR is to list combat validity as one of its aspirations, in the sense of describing or predicting the probable outcome of genuine battles, then analysts will need to understand the way in which performance degrades in these conditions. They must then face up to the challenge of representing this degradation in a convincing fashion in their battle models. It is often suggested that, if models are used only as a means of comparison, then such comparisons will not be affected if these effects are excluded in all the alternatives being compared. Obviously, though, the indicated ranking of alternatives may be changed if the latter are differentially affected by combat degradation.

This paper is concerned with military effectiveness in dismounted infantry combat, both at the level of individual behaviour and also at that of collective performance. It does not attempt to take account of additional factors, such as surprise, artillery bombardment, the influence of any armour/anti-armour weapons that may be present, or the like. These could, of course, have a major impact on any engagements where they are involved. However, until the major behavioural issues are properly dealt with in modelling terms it will not be possible to allow for these additional factors in a satisfactory manner.

This, then, is a modest first step towards the goal of incorporating behavioural factors in a resolved infantry combat model. Even so, the amount of effort involved has been considerable, relying on several interacting and contributing work packages, substantial in their own right, that cannot be incorporated in a single article. First among these has been the construction of a model that represents a stylised infantry assault in peacetime conditions, based on instrumented tactical trials using trained soldiers (Speight & Rowland, 2006). This has included the means used to produce statistically accurate representations of key terrain features noted at the trial sites. Next has come the evidence that, compared to such peacetime performance, that actually achieved in historical battles has been appreciably degraded (Rowland, 1987, 2006 and Rowland & Richardson, 1997). Lastly, we have relied on accounts, by some directly involved in live combat, of the behavioural patterns that may be responsible for this degradation (Rowland & Speight, 2007). The ideal for the present article would be to make it completely self-contained. However, these sources and the details they contain are all integral parts of one continuing effort, and will be mentioned as required.

Figure 1 sets out in schematic form the approach taken in this paper. The boxes above the dotted line are concerned with the basic model just referred to. The boxes below the dotted line are concerned with its adaptation to combat conditions. Targets are acquired according to the logic previously devised and, other things being equal, these acquisitions will lead to firings against the targets concerned. The level of what we have termed Combat Anxiety in each soldier will be a function of this incoming fire. The response to this level of anxiety will depend on each soldier's Warfighting Resolve. This postulated individual attribute is based on the results of historical analysis (Rowland & Speight, 2007) and is

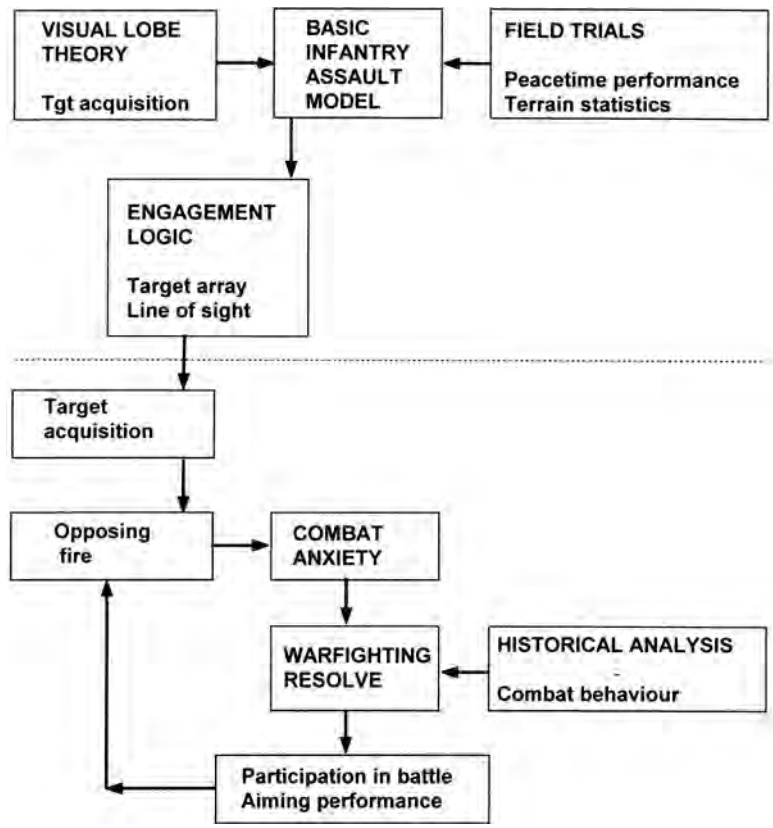

Figure 1. The figure sets out in schematic form the research approach that that will described in greater detail in this paper.

associated with military rank. The degree to which a soldier participates in the battle, and the extent to which his firing accuracy is affected, is thus a function both of the incoming fire and the Warfighting Resolve of the individual concerned. This in turn will affect the characteristics of each side's returned fire.

We are not aware of other initiatives closely akin to the present one in the area of infantry combat modelling so, as already mentioned, we see this very much as a first step. As such we believe that the actual results that we quote are not so important as the patterns that they follow, as compared to those revealed by historical analysis. It also seems important to provide descriptions of the underlying modelling approaches used, as well as some that were rejected, and the considerations that led to these choices. Once this has been accomplished we should stand a better chance of improving on the present logic and, most importantly, of directing our search for such additional data as might put all future initiatives on a sounder footing.

MODELLING CONCEPTS AND THEIR LINKS TO HISTORICAL ANALYSIS

A central thread running through the present initiative is that of individual differences in the willingness or ability of soldiers to make an active contribution in battle. Many experienced military observers have noted these differences. Early attempts at description and quantification were produced by such as Wigram (1943) and, especially, by the pioneering efforts of Marshall (1967). The work described by Rowland & Speight (2007) was based on the records of many other military personnel who, during their direct participation in front line combat, made a careful note of

the patterns of behaviour that they had observed. The main thrust of the paper just mentioned was to collate this evidence and then devise a descriptive framework, linking the latter to the work of others that had also stressed individual differences in various aspects of battlefield behaviour. Clearly, a set of descriptive data, even if it is complete, does not in itself constitute a model. The task for the model builder is to fill in the gaps, assemble the additional information needed for the task, and then construct a coherent dynamic logic that will explain the observed phenomena and, hopefully, predict future combat performance. In the present case it will be apparent that, despite the efforts of those who have striven to collect suitable data, many gaps remain. In what follows we shall indicate clearly where, in the absence of hard evidence, we have had to resort to subjective judgement.

Warfighting Resolve

It is postulated that the tendency to exhibit a pattern of conduct in combat that can best be described as heroic is a stable individual attribute which, for present purposes, we shall describe as a combatant's Warfighting Resolve. Where a given individual stands on the dimension of Warfighting Resolve will determine in large part his likely behaviour pattern during front line action. Table 1 is a copy of Table 4 in Rowland & Speight (2007), but with the final column, headed Estimation method, replaced by one listing shorthand descriptors suitable for modelling purposes. (The three letter code following each descriptor will be used in table headings.) The groups A to F can be thought of as occupying different intervals along the Warfighting Resolve dimension, with Group A having the least resolve and Group F the most. The demarcation points between successive groups may be considered as fixed thresholds. However, the expected number of soldiers lying between these successive increasing thresholds (shown as mean proportions in column 2) should be taken as estimates for British troops only. In other armies the distribution of values along the Warfighting Resolve dimension may be different.

At this stage three points should be emphasised. Firstly, the described level of combat performance strictly applies to attacking troops during the advance and to the same troops in hasty defence. However, having assumed that the Warfighting Resolve of an individual is an enduring characteristic, it is also assumed that somewhat analogous patterns of behaviour (yet to be described) will be exhibited by that same individual in defence. Secondly, the Groups D and E weapon use descriptors (not effective and sub-optimal) have not been defined by the original military observers, let alone quantified. Thirdly, for all groups, especially in the case just highlighted, behavioural patterns must both be specified in unequivocal terms and linked to objective model features.

Table 1. Behaviour patterns for groups with different levels of Warfighting Resolve

Group	Mean Proportion	Level of combat performance	Shorthand descriptor
A	7%	Avoids combat	Non participant (NPA)
B	7%	Present until contact, but then does not fire or advance	Freeze at contact (FAC)
C	36%	Continues after contact, but subsequently halts and does not fire	Freeze during advance (FDA)
D	37%	Maintains advance to end, but no effective weapon use	Passive follower (PAF)
E	11%	Full participation in combat, but weapon use sub-optimal	Active follower (ACF)
F	2%	Heroic performance - full participation & fully effective weapon use	Hero (HRO)

The distribution of Warfighting Resolve as a function of rank and of force effectiveness

From the material summarised in Rowland & Speight (2007) it seems quite clear that the incidence of the different categories of combat behaviour listed in Table 1 may be expected to vary as a function of rank. Those highest up the promotion ladder are the most likely to exhibit a pattern of behaviour that can be classed as heroic. At the same time, although incidences of non participant behaviour at these elevated levels might be very rare, cases of this kind certainly have occurred in the past. These considerations have led us to suppose that Warfighting Resolve must be a distributed attribute, but that the parameters of this unknown distribution (or family of distributions) must differ as a function of rank. It seems reasonable to suppose that the location of any distribution will shift towards the heroic end of the scale as rank increases. We have investigated two main approaches to account for, or describe, this shift. As will be made clear in the section that follows, any approach of this kind will have to accommodate the possibility that different forces may differ in terms of their Warfighting Resolve.

We first examined a version of the approach used in the all-too-familiar statistical routines for the analysis of variance. A model of this kind supposes that the distribution of Warfighting Resolve can be described by an overall mean; a set of constants, one for each rank; plus a random within-rank component. The model may be extended to different armies by replacing the overall mean with a set of means, one for each army being considered. A first assumption would be that the within-rank components would all be normally distributed with equal variance. Fitting this model to the historical data is reasonably straightforward. Assuming the within-rank distributions to be standard normal, one must first find a set of rank means, plus a heroic behaviour threshold, that will yield the historically observed proportion of heroes as a function of rank. One must then find the set of threshold values for the other categories of combat behaviour that yield the overall (all ranks included) proportions of such behaviour.

It will be obvious to readers of the present article that, with this many parameters being adjusted, the fit to the historically observed gallantry award and behaviour category figures will be perfect. A model of this kind will only have value if one is prepared to rely on it for extrapolation. One use might be to give an indication of the likely proportions of behaviour other than heroic within each rank. Another might be to predict the likely behaviour patterns for forces that may be more, or less, effective than those that have yielded one's historical data. In this latter context the behavioural model as just described did not give wholly convincing answers. It was felt that an exceptionally effective force, with a greatly raised proportion of heroes, would also be expected to have a greatly increased proportion of active followers, the category next in line for combat resolve. However, to obtain an increase that was anything like commensurate with the increase in heroes seemed to require the normal distribution to be replaced with one that was positively skewed. Although different alternative distributions were examined, they brought in their train such complications as a requirement for extra arbitrary parameters, as well as, in many cases, a somewhat strained logic to account for the different rank means. These considerations led us eventually to an alternative approach.

The different rank constants mentioned above presumably evolve due to the various effects of initial screening, selection and subsequent promotion. The second approach we devised in order to account for the observed phenomena made overt reference to these processes. Informal evidence, and objective investigations such as the FIGHTER 1 study conducted by the US (Egbert *et al,* 1957 and Harrison *et al,* 1953), suggest that before the event one cannot forecast with absolute confidence how an individual will perform in live combat. Initial screening, selection and promotion cannot therefore be conducted directly in terms of Warfighting Resolve, but rather in terms of some complex military attribute that is imperfectly correlated with it. We shall call this dimension judged military aptitude. We also as-

sume that the different ranks will occupy different contiguous intervals along this aptitude dimension, the highest ranks occupying the top end of the spectrum and the lowest the bottom end. We have assumed that Warfighting Resolve is normally distributed. Just as in the last approach, the evidence suggests to us that, within the serving military population, the distribution of judged military aptitude is not normal, but rather is positively skewed. To a first approximation though, since judged military aptitude only has an indirect bearing on Warfighting Resolve, the upper tail of the normal distribution may be regarded as an adequate representation of this skewed distribution. Using this approximation we can compute the proportion of those with a given rank likely to exhibit a given combat behaviour pattern via one of the many published algorithms dealing with the bivariate normal distribution surface. Although not yielding a perfect fit to the observed data, this second approach calls for many fewer fitted parameters. Once having selected an arbitrary truncation point along the judged military aptitude dimension, plus a correlation coefficient between this dimension and that of Warfighting Resolve that results in the observed overall proportion of heroes, the rest falls automatically into place. The effectiveness of the force as a whole can be altered by adjusting the assumed truncation point on the normally distributed judged military aptitude dimension.

The effectiveness of different forces

It has been emphasised above that the expected proportion of soldiers exhibiting the different behaviour patterns set out in Table 1 are estimates for British forces only. If we accept these as a set of benchmark values it seems reasonable to suppose that there may be forces which habitually show a higher, or some a lower, degree of collective resolve. For instance, Gurkha infantry have achieved an enviable reputation for skill and prowess in infantry battle, and Rowland (2006) has given a summary of some objective historical analysis that supports this view. He mentions that some, such as Caplan (1995), believe that this reputation has been based on Western perception rather than reality. However, Rowland has shown that the Gurkhas outscore their British counterparts by a factor of 1.6:1, not only in terms of gallantry awards but also in terms of the number of casualties inflicted on attackers per defending soldier. This comparison is important, because British Army Gurkhas are organised and trained in similar fashion to their UK comrades-in-arms. Gallantry awards are also made on the same basis in both native British and British Gurkha regiments.

Because of their importance in this context the comparison between the performance of British and Gurkha troops is set out in more detail in Table 2. The first line sets out the number of gallantry awards per soldier killed in action (KIA) in World War II. The results are quoted on a per KIA basis in order to control for exposure to danger. The next line gives the number of casualties inflicted on the enemy per defending soldier, as established by historical analysis. The figures quoted are for World War I and World War II defensive engagements in open country at an attacker:defender force ratio of 1:1. They are

Table 2. Comparisons of British and Gurkha infantry gallantry awards and combat performance

Criterion	British infantry	Gurkha infantry	British:Gurkha ratio
Gallantry awards in WW2			
Awards per KIA	0.105	0.168	1:1.60
Defensive infantry combat WW1 & WW2			
Attack casualties per defender	0.23	0.37	1:1.62
Insurgent guerrilla warfare, Malaya 1953 & 1954			
Kills per platoon contact	0.63 & 0.65	1.00 & 1.02	1:1.59
Guerrillas eliminated per battalion	22.8 & 16.0	32.7 & 28.1	1:1.59

based on 19 Gurkha battles and very many more British battles. The significance of the quoted difference is less than 1%. The last two lines give results, quoted by Coates (1992), pertaining to the insurgent guerilla campaign in Malaya in 1953/54. The first of these lists the number of kills per patrol contact for the two years in question. The last shows the number of insurgents eliminated (both killed and captured, mainly the former) per battalion per year. It will be seen that the ratio both of awards and of all three performance measures hover around the 1:1.6 mark.

Just as the Gurkhas appear to have achieved higher levels of performance than the British benchmark, so it seems reasonable to suppose that yet other forces may be expected to perform at a lower standard. Hartley & Helmbold (1995) in their analysis of the Inchon-Seoul campaign showed that, although the average numbers of US and N Korean personnel involved were very similar, US forces inflicted more than five times as many casualties on their opponents than did the N Koreans. These figures are garnered from US sources only and there may also be biases due to weapon lethality. Nevertheless, the size of this ratio does suggest that N Korean forces were significantly less effective than their US opponents.

The present authors would not claim to be expert on matters pertaining to military morale, sociology or performance. In this paper, therefore, we shall not speculate on the causal factors that might underlie such apparent differences in force effectiveness. We simply observe that they seem to be an inescapable feature of military history. In what follows we assume that differences in force effectiveness can be represented in rough-and-ready fashion by differences in the assumed selection:rejection ratio on the judged military aptitude dimension, resulting in turn in a higher or lower average on the dimension of Warfighting Resolve. Since the rank structure is determined principally by the organisational demands of the force, these differences will be reflected to a greater or lesser extent at each rank level.

The effects of armour/anti-armour weapons and of artillery bombardment

The results of historical analysis concerning the suppressive effects on infantry of armour/anti-armour weapons and following artillery bombardment have been given in Rowland (1987). Although there is no need to repeat the details here a point worth making is that both of these types of suppression seem to be multiplicative in their effects. They appear to affect both hero and non-hero alike. In the case of artillery bombardment we should emphasise that the neutralising effect is transitory, although the rate of decay is not well established. So long as an infantry attack follows the preparatory artillery bombardment with minimal delay its effectiveness will be maintained. The fact that the temporal density of bombardment is as important as the spatial density suggests a process of intermittent renewal and decay. Anecdotal evidence suggests that this same feature may also hold true for direct fire artillery bombardment by, for example, opposing armour. In what follows we shall likewise assume a renewal and decay process for the suppressive effects of small arms fire.

MODEL REALISATION

Speight & Rowland (2006) have given an account of a model of the attack phase of an infantry engagement, together with details of its construction and validation against the outcomes of instrumented tactical trials using trained soldiers. The present investigation made use of this model, embellished with additional features, as detailed below, required for the representation of live combat effects. One such feature was an additional attribute for each soldier represented in it: his heroic status, as detailed in the last column of Table 1. The model is a time stepped stochastic simulation in which attacking infantrymen, grouped in sections, advance towards a defence consisting of riflemen placed in trenches. Although the terrain is flat, it can be represented as open, mixed or close. For close terrain the spatial density of obscuring objects, as well as the cu-

mulative distribution of their cross-sections, approximates statistically those of the engagement sites classified as close after detailed analysis of the terrain plots obtained during the KINGS RIDE series of field trials. In mixed terrain the spatial density of objects is only half that of close terrain. Open terrain has a complete absence of obscuring objects During each time step the model cycles through the following calculations: attacker movement; line of sight determinations; troop actions and status changes; followed by a test for completion. Each soldier included in the model can be in one of the following states: searching for targets; checking a target that has been detected; aiming; firing; reloading; recovering from a non-lethal wounding; or killed. Defenders are all considered to be immobile, but attackers can either be mobile or, if they have been rendered so by non-lethal wounding, immobile. The model keeps track of every round fired, its trajectory and, if the latter should encounter an opposing rifleman, the consequences of that impact. This study employed only eight man sections in attack, each with a standard 30 metre frontage and a 12 metre gap between adjacent sections. The advance commenced approximately 350 metres from a line of 4 man trenches and ended some 50 metres from this line. The trenches were spread evenly (but with some random variation) across the attack frontage, with a minimum spacing of 10 metres between adjacent trenches. In order to align the results with those stemming from historical analysis the output measure in this investigation is casualties (defined as those killed plus those immobilised by wounding). This is in contrast to Speight & Rowland (2006), where the results were quoted in terms of kills only.

The sampling of Warfighting Resolve

In the "benchmark case, with effectiveness similar to that of British troops, a force composition along the judged military aptitude dimension was assumed equivalent to a truncation point of 0.935 on the standard normal distribution. (It will be recalled, though, that this is an arbitrary figure. It should not be taken to imply that 82.5% of the population as a whole has been rejected, or that the distribution to the left of the truncation point is approximately normal.) The correlation coefficient between judged military aptitude and Warfighting Resolve (*i.e.,* the slope of the linear regression line) was assumed to be 0.725. (These truncation point and correlation values were arrived at via an informal procedure optimising the fit demonstrated in Table 3 below. However, so long as these two parameters were varied jointly, it was found that the obtained fit was not very sensitive to the exact values chosen.) Figure 2 illustrates graphically the effects of the assumed selection and promotion scheme in terms of expected combat behaviour. The vertical scale corresponds to numbers within the force. The different ranks are strung out along the judged military aptitude dimension. Since the latter is positively correlated with that of Warfighting Resolve, increasing proportions of each rank fall above the thresholds of more effective combat behaviour as one progresses towards the higher ranks. This simple scheme yields the expected incidence of heroic behaviour as a function of rank shown in the first column of Table 3. Also shown in the second column is the actual pattern of gallantry awards uncovered by the research reported in Rowland & Speight (2007). Here the last line represents the expected proportion in an infantry company overall, given the manning complement usually found during World War II (but it should be noted that this number has been rounded up to two figure accuracy in Table 3 of the just-quoted paper). The overall fit

Table 3. Model predictions of proportions attaining heroic status compared to proportions yielded by historical research (Rowland & Speight, 2006)

Rank	Model	Historical analysis
Captain, Major	0.226	0.300
Lieutenant	0.099	0.060
Senior NCO	0.059	0.060
Junior NCO	0.026	0.025
Private	0.005	0.005
Overall	0.0177	0.0177

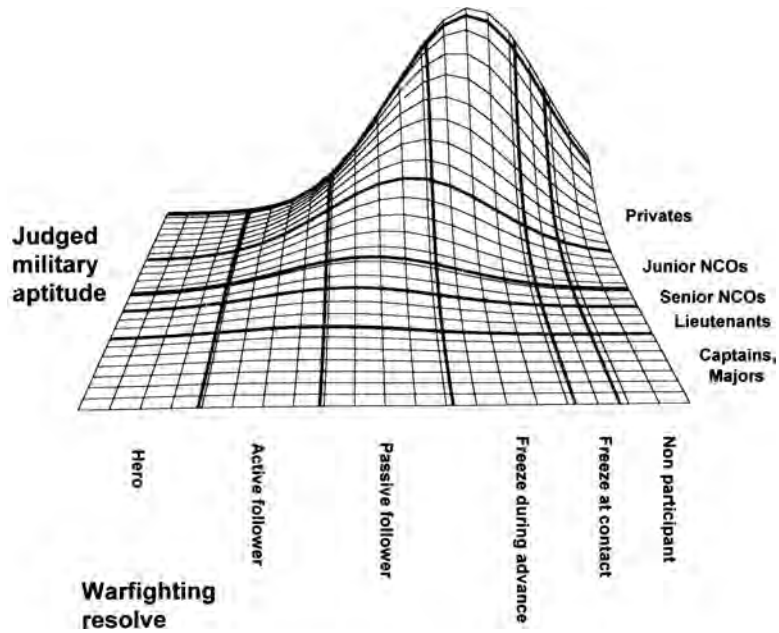

Figure 2. The figure illustrates the effects of initial screening, selection and promotion along the judged military aptitude dimension and its impact on standing along the Warfighting Resolve dimension (and hence on the likelihood of different combat behaviour patterns as a function of rank).

and that for non-commissioned ranks is remarkably good. However, this scheme does overestimate the figure for Lieutenants and underestimate that for Captains and Majors. It will be noted, though, that in practice the commissioned officers are subject to a different induction process than that for other ranks, almost certainly with more complex criteria for selection and promotion. As will shortly be made clear, for the purposes of the present investigation this anomaly is not thought to be crucial.

For completeness the values implicit in Figure 2 are given in numerical form in Table 4. It will be recalled that both the judged military aptitude and Warfighting Resolve dimensions are assumed to be distributed in standard normal fashion. The first line of figures (in italics) shows the threshold values pertaining to the listed behaviour category on the normally distributed Warfighting Resolve scale. Thus, if an individual has a personal Warfighting Resolve rating of more than 1.928 but less than 2.725 he would be held to be an Active Follower. Simi-

Table 4. Model predictions of the proportions within each behaviour category as a function of rank (Behaviour category codes are given in Table 1.)

	HRO	ACF	PAF	FDA	FAC	NPA	
Behaviour Threshold	*2.725*	*1.928*	*1.055*	*0.249*	*−0.042*	*−9.999*	Rank Threshold
Captain, Major	0.226	0.414	0.304	0.053	0.002	0.001	*2.713*
Lieutenant	0.099	0.348	0.423	0.119	0.008	0.003	*2.394*
Senior NCO	0.059	0.283	0.462	0.174	0.014	0.007	*2.170*
Junior NCO	0.026	0.187	0.464	0.278	0.032	0.021	*1.695*
Private	0.005	0.069	0.343	0.408	0.086	0.089	*0.934*
Overall	0.018	0.112	0.370	0.360	0.070	0.070	

larly, the last column (also in italics) gives the assumed rank thresholds (for a British force) on the judged military aptitude scale. Assuming a half-company size force in attack and defence, we did compute a handful of results in which the different ranks were distributed within sections and trenches in a pattern that might be expected based on command and control considerations. Command and control functions are not, of course, properly represented in this simple model. Nevertheless, the judicious placement of those more likely to exhibit effective combat behaviour patterns did increase attrition on both sides to a certain extent. However, the effect compared to those linked to all the other factors under investigation was very small. In order to investigate effects due to force ratio and sizes, therefore, we have ignored rank and have simply sampled the Warfighting Resolve status of each individual according to the proportions to be expected within an infantry company.

Force effectiveness variations

In addition to our benchmark set of results we have computed model predictions for two different force compositions, one with higher average Warfighting Resolve levels and one with lesser. For ease of reference we have labelled our benchmark force *B*, the more effective force *B*+ and the less effective one *B*−. The assumed truncation points on the judged military aptitude dimension were held to be 0.934 for *B*, 1.275 for *B*+ and 0.523 for *B*−. Each force followed the same Warfighting Resolve sampling logic, given the different initial selection rates. The *B*+ selection assumptions yielded an assumed proportion of heroes a factor of 1.6 larger than that in the *B* force, as determined historically for Gurkha gallantry awards (see above). The *B*− assumptions were entirely arbitrary, chosen simply to yield a nominally less effective force than the other two. The expected overall proportions within the different behaviour categories for these three forces are shown in Table 5. There is a slight complication: if the *B*+ forces are meant to represent Gurkhas. British Gurkhas have normally had British, rather than Gurkha, officers. The logic employed to generate the figures shown in Table 5 implies that these, too, were a selected elite, with greater average Warfighting Resolve than the officers in purely British regiments. In what follows we shall use the figures shown in Table 5 whenever we simulate engagements involving *B*+ troops. However, we shall return to this point in our final discussion.

Model response as a function of Warfighting Resolve categories

Table 1 gives a description of various levels of behaviour that we have categorised under the heading of Warfighting Resolve. We have already pointed out that some of the descriptors, such as sub-optimal or ineffective weapon use, are somewhat lacking in precision. In all cases, though, these forms of behaviour will have to be tied down as definite rules of model response. As a prelude we shall need to consider the phenomenon of anxiety, both prior to and during an engagement.

It is natural to suppose that all troops will feel some measure of anxiety at the imminent prospect of battle. In a sense it is meaningless to ask whether one soldier feels more anxious than another: anxiety is an internal state that cannot be measured directly. What can be as-

Table 5. Overall proportions within each behaviour category for forces of different combat effectiveness

	HRO	ACF	PAF	FDA	FAC	NPA
Force effectiveness						
B−	0.011	0.074	0.299	0.390	0.099	0.128
B	0.018	0.112	0.370	0.360	0.070	0.070
B+	0.029	0.159	0.420	0.308	0.047	0.038

sessed, and what does matter, are differences in the overt behavioural signs of anxiety including, especially, whether that man will perform his military functions in a satisfactory manner. Historical analysis suggests that, for the small minority of non participants, what we shall call pre-combat anxiety will be sufficient cause for them to absent themselves from the forward area of the battlefield. In what follows we shall first consider the combat features (as portrayed in the model) that may raise the level of anxiety during an engagement, and then the individual responses of those with different degrees of Warfighting Resolve to these raised anxiety levels.

Once the force makes contact with the enemy and comes under fire we shall assume that the underlying layer of pre-combat anxiety is overlaid with an additional layer of what we shall call combat anxiety. There is little available evidence to suggest just what characteristics of the opposing fire do most to affect the internal state of the troops on the ground. One extreme position would be to suppose that it is not the fire as such that gives rise to anxiety, but simply the knowledge that the enemy is present and active. However, this would be to fly in the face of the evidence collected by Cawkill (1997) from veterans with front line infantry experience. They suggested that they were indeed suppressed by enemy fire, although there were marked individual differences in the length of time that they were suppressed. If in this model one were to give no weight at all to the density of the opposing fire then this would produce an effect that many will regard as anomalous. The model assumes that a greater proportion of effective troops will place themselves in positions where they are vulnerable and fully exposed to opposing fire than will be the case with less effective troops. If the density of attacking fire has absolutely no psychological effect then the result will be that defenders will kill more of the attackers, on average, if they are faced with effective troops than with ineffective ones.

A common approach to the modelling of artillery suppression is to start by assessing by some means or other a region of effect for individual shells or rounds. These assessments can then be linked to estimates of the probability that an individual will be within such a region at a given instant, plus rates of fire and rules for the aggregation and decay of individual effects. Although much of the emphasis in the past has been placed on indirect artillery, some attempt has been made in peacetime conditions to assess the likely region of effect for both artillery and small arms fire. In the 1970's the US Army Combat Development Experimentation Command (CDEC) conducted a lengthy series of suppression trials under the headings of SASE (Small Arms Suppression Evaluation) and SUPEX (Suppression Experiment, Phases I to III). The method was basically to place volunteer observers using periscopes in protected positions, asking them to lower these periscopes when they felt the latter were vulnerable to rounds or charges at various offsets. Without attempting to summarise the considerable data that resulted, the indications were that small arms rounds (either singly or in bursts) would have to pass very close indeed (appreciably less than 10 metres) to have a significant effect. We found that a Combat Anxiety sub-model based on the single or cumulative effects of individual rounds, each with a very small region of effect, had a very one-sided impact. Attackers, with their difficulties of target acquisition, have low firing rates. In comparison the weight of fire from the defenders is very considerable. Thus, while attacker anxiety levels are raised to high levels, those of the defenders are relatively little affected. If there is indeed a relationship between anxiety and firing rates, then there will be a feedback loop that will increase this disparity. This does not square with the observations of Marshall (1967), who found that non-participation was affected relatively little by the strength of the opposing fire.

We have come to the conclusion that, when soldiers are facing aimed small arms fire in live combat, the region of effect may well be larger than that yielded by such peacetime experiments. The peacetime experiments may still yield valuable information concerning the relative detectability of rounds of different calibre. However, when a soldier detects an individual passing round what is important is that it pro-

vides evidence that the enemy is firing at the group of which he is a member. The shot in question may not have hit the man who has detected it, but his anxiety will be raised by the thought that he may be the intended target for the next round fired. With these thoughts in mind we have adopted the following approach for our modelling representation of combat anxiety.

> *Combat Anxiety (CA).* At the start of an engagement it is assumed that all involved have CA values of zero. Immediately following an opposing round passing within a miss distance of 50 metres or less the CA of the individual concerned will increase to a value of unity. Thereafter, assuming that no other round passes within this miss distance, this level of CA will decay exponentially at a rate corresponding to a half life of 60 seconds. However, if a further round should pass within 50 metres the level of CA will revert to unity.

This definition of Combat Anxiety still favours the defence, in the sense that average CA levels will be less at all stages than those of the attackers. However, the effect is more balanced than would have been the case if we adopted a measure that relied on the number of rounds passing within a short distance of any individual.

We come now to the question of response to a given CA level so measured. We assume that this will be affected by an individual's standing in terms of Warfighting Resolve. Those with the three lowest categories of Warfighting Resolve are assumed either to avoid raised CA levels, or to respond to raised CA levels, by failing to participate in the battle in any meaningful way. Heroes are assumed to be immune to the effects of raised CA levels. For the two remaining categories, those dubbed active or passive followers, we have tried to strike a balance between the evidence produced by Marshall (1967) of failure to fire, and the evidence seen on film or TV of hasty and inaccurate fire in dangerous situations. We have postulated increasingly inaccurate fire as CA levels are raised, until a point is reached, in the case of passive followers, where firing ceases all together. In order to incorporate these ideas in our model we have added an additional pair of attributes for each soldier: his Warfighting Resolve category and his current CA level. Each soldier can also be in an additional state: if he is concealed he is invisible to the enemy, cannot be killed and does not participate in the battle in any meaningful way. Table 6 now sets out the model representations of the more qualitative behaviour patterns listed in Table 1.

We should highlight two of the behavioural characteristics that the suppression model just outlined brings in its wake. Firstly, it should be noted that the CA thresholds for a given behavioural effect are all lower for passive followers than they are for active followers. The exponential decay characteristics assumed for Combat Anxiety thus ensure that for a given battle event this level of suppression will last longer for the passive followers than it will for active followers. This is in line with the individual differences in length of suppression noted by Cawkill (1997). A second feature worth emphasising stems from the heavier weight of fire produced by the defenders (who have a comparatively simple target acquisition task) compared to that produced by the attackers. As a result, with this depiction of small arms suppression, very few of the attacking passive followers are firing at all by the time they are within a hundred metres or so from the line of defensive trenches. At the same point in the engagement a goodly proportion of the defending counterparts are still firing, albeit with reduced accuracy.

RESULTS

For the main experiment the model was run for $B+$, B and $B-$ troops in defence and attack; in open, mixed and close terrain; and for all combinations of $n = 3, 4, \ldots 12$ eight-man attacking sections and $m = 4, 5, \ldots 15$ four-man defending trenches. For each of these $3\times3\times3\times10\times12$ conditions the model was run 1,000 times. The Warfighting Resolve of each attacker and defender was sampled at the start of each battle run. For purposes of comparison the model was also run with all heroes in both attack and defence in the

Table 6. Model behaviour as a function of Warfighting Resolve category

Warfighting Resolve category	Model behaviour
NPA Non participant	Moves to status concealed at commencement of each battle run.
FAC Freeze at contact	Neither attackers or defenders search for targets. Attackers assume prone position on first occasion CA > 0, defenders move to status concealed.
FDA Freeze during advance	Neither attackers or defenders search for targets. Attackers assume prone position on first occasion round passes within 5m, defenders move to status concealed. (*Note that this is substantially less than the 50m miss distance that will cause FAC troops to freeze.)*
PAF Passive follower	Both attackers and defenders cease searching for targets when CA > 0.25 and recommence if CA < 0.25. Aiming accuracy is a function of CA: 0.25 > CA > 0.1 dispersion is × 2.5 normal; 0.1 > CA > 0.05 dispersion is × 2.0 normal; 0.05 > CA > 0.01 dispersion is × 1.5 normal; CA < 0.01 dispersion is normal.
ACF Active follower	Both attackers and defenders search for targets at all times. Aiming accuracy is a function of CA: CA > 0.9 dispersion is × 2.5 normal. 0.9 > CA > 0.8 dispersion is × 2.0 normal. 0.8 > CA > 0.6 dispersion is × 1.5 normal. CA < 0.6 dispersion is normal.
HRO Hero	Both attackers and defenders search for targets at all times. Aiming accuracy is unaffected by CA

3×10×12 terrain/sections/trenches conditions. Rather than attempting to summarise all this material we shall quote selectively from it in order to illustrate the main themes that have emerged.

Comparison with the major empirical results yielded by historical analysis

A major result stemming from historical analysis, not mentioned in either Rowland, 1987 or Rowland, 2006, is that defender losses during the approach phase are normally very small indeed - so small that they can, for purposes of establishing comparisons, be ignored. This finding is certainly mirrored in our model outputs. For the largest battles that we have studied (12 sections attacking 15 trenches) in open terrain and with *B+* troops in attack and defence, there is on average less than one defender killed per engagement. We should also mention that, now that the effects of individual differences are represented, results are very variable indeed. Despite the large sample sizes some random variability is still visible superimposed on the overall trends of the results.

Figure 3 presents the output of our model in the terms that the historical analysis findings are presented in Rowland (1987 & 2006): the number of attacker casualties per defender as a function of the number of attackers present per defender. Each data point stems from 1,000 independent model runs. The results labelled as

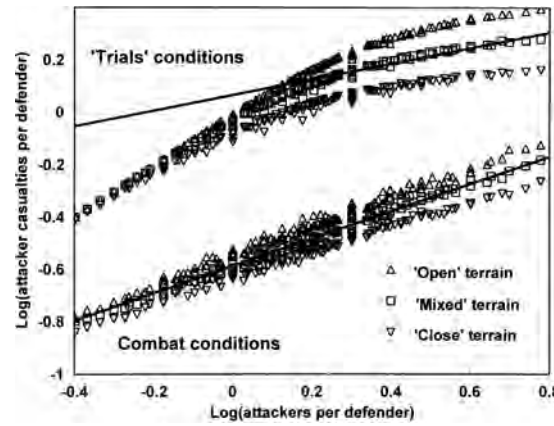

Figure 3. Number of attacker casualties per defender as a function of the attacker:defender ratio. Each point is the average of 1,000 model runs.

trials conditions are with both attacking and defending troops all treated as heroes. The bottom set of results, labelled as combat conditions, samples both attackers and defenders according to the Warfighting Resolve proportions to be expected of benchmark (*B*) troops. The two lines are those yielded by linear regression analyses, with all the data being used for the combat condition but only those corresponding to attacker: defender ratios of 1:1 and higher for the trials condition. Figure 4 shows, in bold, the regression lines yielded by historical analysis (Figure 4 in Rowland, 1987, and Figure 3.2 in Rowland, 2006). Here the top line is derived from the results obtained during the Kings Ride II series of field trials. The bottom heavy line (with 95% confidence limits also drawn in) is obtained from the analysis of actions in the Boer War and the US Civil War. Like our own simulation results, machine guns did not feature in these two conflicts. Also included, once again, are the regression lines (dotted) derived from our own model outputs.

Dealing with the bottom set of results first, little comment seems necessary. Given all the factors likely to influence genuine historical data, not represented in systematic fashion in our model, the degree of correspondence is remarkably close. While both the KINGS RIDE relationship and that fitted to model trials outputs show similar trends, the defensive performance indicated by the former is appreciably superior to that shown by the latter. However, the live trial results are based on relatively few actual exercises, employing laser simulators intended to represent self-loading rifles, as well as devices representing a very much more potent light support weapon. Parallel trials in which the troops fired genuine rifles with live ammunition showed significant differences compared to both these trials equipments, especially the light support weapon. The model results incorporate our best estimate of the inferior performance to be expected with genuine rifles.

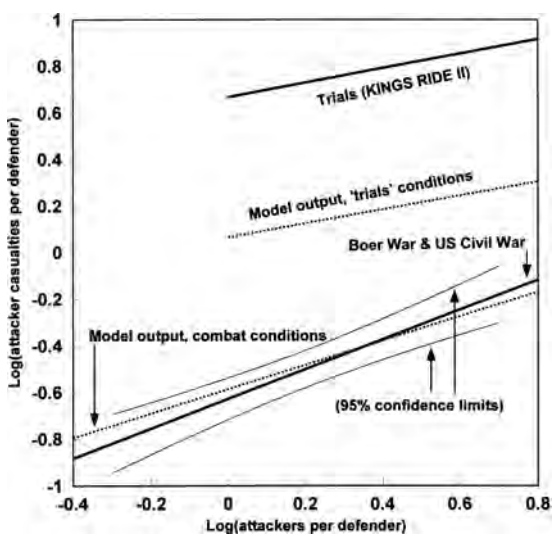

Figure 4. Comparison of model outputs with the results stemming from historical analysis.

Rowland (1987 & 2006) quotes an historical average of 0.23 attacker casualties per defender when force ratios are at parity. The value predicted by the model is 0.26. Obviously, the latter figure reflects the set of parameter values chosen on this occasion. If, for example, we had increased the critical miss distance for inducing Combat Anxiety from 50m to 100m, then the figure produced by our model would have been precisely 0.23. Although this close correspondence is comforting, it is our opinion that any feelings of triumph should remain muted until useful progress has been made on the various issues set out in our concluding remarks.

There is one further feature of Figure 3 that is worth noting. If we look at the top set of data points we may note that there is a clear separation of kill rates in the three different terrain conditions at higher attacker:defender force ratios. This difference is far less marked in the pseudo-combat conditions shown at the bottom of the figure. It is possible that in real warfare differences due to terrain clutter may shrink into insignificance.

Effects due to force effectiveness

Taken over all the conditions covered by the main experiment, attacker casualties imposed by *B*+ defenders exceeded those inflicted by *B* defenders by a factor of 1.21. The equivalent figure for *B*− defenders was 0.77. Given the modelling logic that has been devised a failure to show a difference would cer-

tainly have been extraordinary. However, it will be recalled that the $B+$ logic was based on the combat achievements of Gurkha troops reported by Rowland (2006), and their kills per defender were a factor of 1.6 greater than those inflicted by their British comrades in arms. The issues raised by this difference will be picked up in the final discussion.

Figure 5 shows the casualties inflicted by $B-$ defenders when faced by $B-$ and by $B+$ attackers, and Figure 6 shows the same thing when $B+$ troops are in the defence. Since the results that they produced were little different we have averaged the outputs yielded by the three different terrains. Taking the results in Figure 3 first, it will be seen that the $B-$ defenders inflict greater casualties on less effective opponents than they do on the more effective attackers. Differences are small when there are small numbers in opposition, but rise as the number of attackers increases. $B+$ forces are less affected in this way than are $B-$ forces, and it should be noted that for small numbers in opposition the defence actually inflicts more casualties when it is faced by more effective troops. The trend is reversed as the number of attackers increases. As has been noted previously, there are two competing mechanisms at work here. $B+$ soldiers are more willing to present themselves as vulnerable targets than are their $B-$ counterparts, and so target acquisition is easier. However, they also produce more suppressive fire. As the number of attackers increases this last effect more than compensates for the greater target availability. Whether or not this suppression effect is regarded as realistic is a matter for personal judgement. It will be added as another topic for debate in the final discussion.

The effects of terrain relief (advance speed)

Rowland (2006) has reported some historical analysis concerning the effects of terrain relief in infantry combat. It was noted that in mountainous areas the defence appeared in general to be more effective in terms of inflicting casualties on the enemy. The present model, with its representation of the battlefield as a plane surface, is hardly suited to the examination of relief effects. However, the analysts surmised that the major causal mechanism at work here could be the effect that terrain relief has in slowing the advance. As a rule of thumb, casualties appeared to increase in inverse proportion to the square root of the advance rate. Thus an advance rate of, say, one tenth normal would be expected to increase casualties by a factor of $\sqrt{10} = 3.16$. Figure 4.17 in Rowland (2006) illustrates the observed relationship. Readers should note the appreciable scatter of

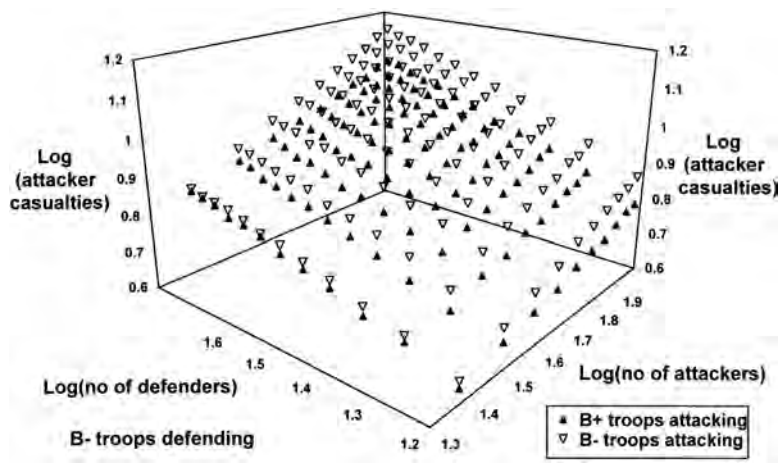

Figure 5. Number of $B-$ and $B+$ attacker casualties inflicted by $B-$ defenders as a function of numbers in attack and in defence.

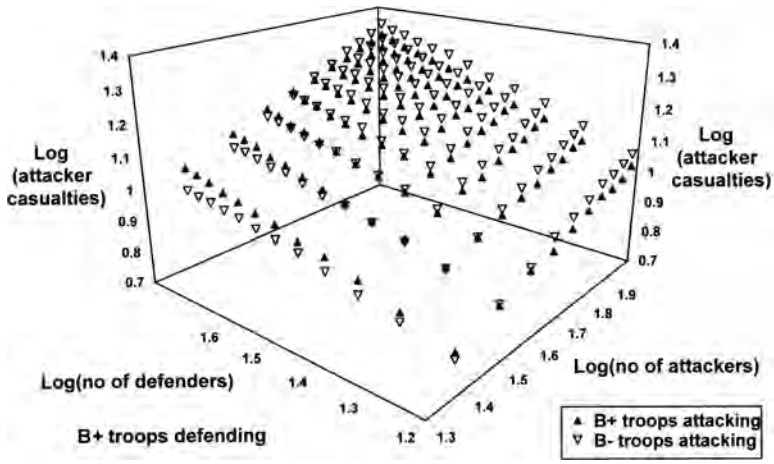

Figure 6. Number of $B-$ and $B+$ attacker casualties inflicted by $B+$ defenders as a function of numbers in attack and in defence.

empirical results around the line defining the theoretical relationship.

We decided to conduct a very limited investigation (if only because of the greatly increased running times involved) to see whether our model produced outputs roughly in line with the relationship yielded by historical analysis. The model was run with advance speeds reduced by a factor of 10, with B and $B-$ troops both in the attack and in the defence, in all three types of terrain, with the numbers of attackers and defenders given in Table 7. The table shows the ratio of attackers killed at the reduced advance speed compared to that noted at normal speed (in each case taking the geometric mean of the ratios achieved in the three different terrains). The average ratio in Table 7 is 3.13, compared to the figure of 3.16 yielded by historical analysis. This correspondence is comforting in confidence-building terms. However, there is another useful by-product.

Table 7. Ratio of attacker casualties at one-tenth advance speed, compared to numbers of casualties when advance speed is normal

Attacking Defending	Number of attackers	Number of defenders	Ratio of number of attackers killed (reduced speed / normal speed)			
			B B	B $B-$	$B-$ B	$B-$ $B-$
	24	24	2.49	3.14	2.02	2.44
	24	36	2.10	2.64	1.66	2.09
	24	48	1.88	2.43	1.54	1.86
	24	60	1.76	2.24	1.46	1.74
	40	60	2.87	3.29	2.22	2.54
	48	16	5.10	5.86	3.77	4.66
	48	24	4.14	5.32	3.13	4.02
	48	32	3.72	4.81	2.77	3.45
	48	48	3.25	4.19	2.40	3.05
	48	60	2.98	3.88	2.19	2.76
	96	24	6.27	7.16	4.77	5.73
	96	32	5.83	6.99	4.34	5.30
	96	48	5.18	6.45	3.78	4.76

The modelling exercise indicates that there may be systematic effects that could help to explain some of the variance in the empirical results. The results suggest that the advantage due to an elevated defence position may increase as the attacker: defender odds lengthen; as the size of the battle increases; and/or if the combat effectiveness of the attacker exceeds that of the defender. Conversely, the advantage may be less as the numerical odds shorten; the size of the battle decreases; and/or the attacker has relatively low combat effectiveness. Thus, if a denotes the ratio just defined, and r and b are the numbers of attackers and defenders respectively, then the following linear equation accounts for 99.1% of the variance of the data shown in Table 6:

$$y = 0.0854 + 0.455x_1 + 0.108x_2 + 0.217x_3$$

where

$y = \ln(a)$

$x_1 = \ln(r/b)$

$x_2 = \ln(rb)$

$x_3 = 1$ for $B/B-$ attackers/defenders,
0 for B/B or $B-/B-$ attackers/defenders,
-1 for $B-/B$ attackers/defenders.

This illustrates the role that modelling may play in suggesting hypotheses for further test, and in indicating research avenues that could be productive. We should also mention that, for a small number of cases, we examined the effect of reducing the distance at which attackers were first exposed, and also of splitting the battle into two equal halves. The idea was to introduce in an artificial manner some of the effects that might have resulted if undulating relief had been included in the model. These changes made little difference to the results.

CONCLUDING DISCUSSION

It will be obvious to all that, in the interests of producing a reasonably compact model, we have included within it some fairly artificial constructs. By reducing the screening, selection and promotion process to a single dimension with sharp demarcation points we undoubtedly fail to do justice to its complexities. Few would believe that the momentary level of anxiety in combat will depend solely on the miss distance of the last enemy bullet. We could go on. But the key question we must ask is whether or not these artificial abstractions do a sufficiently good job of recreating the major interactions that occur in genuine combat. We have been encouraged by what we have seen when we have compared our outputs with the patterns revealed by historical analysis. It should be noted that the model construction process was not driven by the historical casualty findings. The guiding principle in design was to include in our model representations of the key mechanisms noted in realistic peacetime trials, and then to overlay these with the individual behavioural differences noted in actual combat. The comparison with the effects of advance speed was added as an extra test of model robustness. We have not attempted to adjust any parameters in order to obtain a better fit. There are some informal features of the model which give us reason to hope that it is in tune with combat realities. For instance, in Cawkill's (1997) questionnaire study experienced soldiers were of the opinion that it is at the start of the engagement where the desire to take cover is normally at its height. With our suppression logic the majority of those taking cover do so at this point. We feel comfortable, therefore, with what has been achieved at this stage. We believe that the model as it stands provides a better description than most of what we may expect in war, and that it can provide useful pointers to topics that now merit more research and investigation.

For the two authors the development and testing of this model has been a stimulating and educational experience. We believe that the processes of modelling and of historical analysis have been complementary to one another. While historical analysis has provided the basic facts, modelling has suggested possible causal mechanisms, directing attention to areas where further knowledge or analysis is either needed or could be profitable. On the basis of what we have learned we would put forward the follow-

ing topics for serious consideration by the military OR community:

Limitations of the available historical combat data. Even in the clinical environment of the computing laboratory engagement results are extremely variable. Each of the data points in this paper is based on 1,000 model runs but, even so, trends are not completely regular. Recorded historical cases are far less numerous, and each will be affected by real life influences that, with the best will in the world, will never be fully captured in our artificial models. We also know from our attempts to reconcile different historical accounts of the same battle that estimates of the number of casualties and the number of assailants on each side can be subject to error. These errors will be distinct from any variability arising from the various processes of war, and their effect will be to distort the values of any functional relationships that we attempt to derive from the historical data (see, e.g., Cheng & Van Ness, 1999). Errors in estimating attacker casualties and starting numbers will tend to reduce the value of the slope parameter in any fitted linear regression equation. However, because both casualties per defender and attacker:defender ratio are divided by the same estimated number of defenders, any errors in the latter will have an opposite effect: the value of the slope parameter will be inflated. If the number of defenders is overestimated then this means that both the calculated casualties per defender and the attacker:defender ratio will both be too low, and *vice versa.* High quality combat data will collectively form the foundation for any worthwhile advance in this area of combat modelling, and the search for good historical examples must be one of the most important tasks that members of the military OR community can undertake.

Providing a fuller account of the total infantry engagement. We have concentrated here on the approach phase of the rural infantry battle, and have not moved on to the close quarter battle that may follow. But, even if we were to exclude the latter, our description of the approach phase is incomplete. Our simulation stops short at a point 50 metres from the defending trenches, where the nature of the battle will change and both sides start to prepare for the hand-to-hand fighting that may ensue. However, a proportion of all attacks will be doomed to fail at or before this point, and when failure occurs some sort of withdrawal will follow. During this phase it will be difficult for the would-be attacker to supply suppressive fire. To the casualties incurred during the advance must be added those that are incurred during any withdrawal. How all these effects will balance out will not be certain until the simulation is complete.

The way that the target acquisition process is portrayed. As mentioned when the construction of this model was first described (Speight & Rowland, 2006), there is no unified theory of visual search and detection. There is also a dearth of good empirical data concerning search tasks truly representative of that faced by an infantryman in combat. It is certainly possible that the visual lobe theory used in the present model, and described briefly in the reference above, wrongly estimates the reduction of acquisition times as the number of potential targets increases.

Suppression by small arms fire. Although Cawkill's (1997) veterans have provided valuable evidence concerning the suppressive effects of small arms fire, the details of our portrayal have relied heavily on our own subjective judgement. It is highly likely that the judgement of different experts will vary in this area. Obviously, any firm evidence concerning the suppressive effects of enemy small arms fire would be of great value in improving the way that it is represented in our models. Here we have restricted ourselves exclusively to aimed fire. However, in most of the infantry exchanges we see portrayed on our television screens purely suppressive fire is much in evidence. A better understanding of suppression seems essential if we are to achieve a better understanding of real battles.

The effect of infantry weapons other than the rifle. In this investigation we have deliberately restricted ourselves to one weapon type: the self-loading rifle. The process of model development seemed challenging enough without adding any extra complexity. However, ever since they assumed prominence in World War I machine guns have had a major influence in infantry battles. In the present version of our model we have traced the trajectory of every rifle round fired: if we attempted to do this for machine guns we would obviously have a more difficult task on our hands. But there are also other, more important, challenges to be faced with this weapon. Both Marshall (1967) and Rowland (1987 & 2006) report that the

combat degradation with these more potent weapons appears to be less than with rifles. It may be that commanders select for these responsibilities those that they judge will show most resolve in battle, or that the firer is emboldened by the greater chance of inflicting damage on the enemy. Many machine guns are served by a two man crew, and this also could be a factor in making them less susceptible to combat degradation. There may be other causes or combinations of causes. However, as things stand no model of infantry combat will be complete until all the most important infantry weapons are represented in a convincing manner.

The synergistic effects of heroism and of supervision. In our survey of behaviour in combat (Rowland & Speight, 2007) we acknowledged the importance of direct supervision in maintaining military effectiveness in battle, and yet we acknowledged that we did not have the evidence to construct an accurate portrayal of its effects. Time and again, though, one sees cases where meaningful participation in the battle evaporates once the guiding influence of supervision is removed. However, we have come to suspect that informal example has just as important role to play as has formal leadership. Man is a social animal, and depends greatly on the support and encouragement of others. Witness the well-known tendency, reported by Holmes (2003) and many others, for troops to bunch when exposed to fire, even though in rational terms this may make them more vulnerable. In our simulation we found that, compared to the performance of benchmark troops, the increase in casualties inflicted by *B+* defenders was only a factor of 1.21, although the number of heroes was increased by a factor of 1.6. Historical records suggest that for Gurkhas, who were the template for *B+* troops, the factor for both casualties inflicted and gallantry awards was 1.6. All the evidence that we have reviewed leads us to the belief that the propensity for courage and resolve in battle is distributed in some way among the serving population. Our own efforts in the current investigation also convince us that, if we simply stick to a linear model, it would take a much more exotic distribution mechanism than any that we have investigated in order to yield a result in line with the historical research. A better explanation would be to accept that the effects of supervision and heroism are non-linear. The example and direction provided by effective leaders and by heroes serves to stiffen the resolve of others in the force, and helps them to execute their military tasks more effectively. Provisional research results suggest that, both in infantry and armoured warfare, the effectiveness of a crew-served weapon reflects the performance of the crew member with the highest level of Warfighting Resolve (see, e.g., Rowland, 2006, page 164). By the same token, a relative dearth of heroes or of effective leadership may weaken the whole force, and lead to a lower level of overall effectiveness than would be predicted by a simple linear model.

None of the issues set out above are easy to tackle, but they do seem crucial if we are to model combat performance in a truly convincing fashion. There are difficult modelling choices to be made, and many of these may affect the relationships found here. Finally, though, we should emphasise that thus far we have concentrated almost entirely on attrition and the infliction of casualties. What matters in the wider picture is not just what the casualty level may be, but whether either side succeeds or fails in its attacking or defensive mission. We hope to turn to the issue of the likely outcome of infantry engagements in a future paper.

Author Statement

David Rowland formerly worked with the Defence Operational Analysis Establishment/ Defence Operational Analysis Centre/Centre for Defence Analysis, DERA, under whose aegis the trials activities and historical data collection were conducted.

REFERENCES

Caplan, L. *Warrior Gentlemen.* Oxford: Berghan Books, 1995.

Cawkill, P. The Suppressive Effects of Infantry Direct Fire: A Survey of 71 British Army Soldiers and Royal Marines. Defence Evaluation & Research Agency, Centre for Human Sciences. *Unpublished Report.* 1997.

Cheng, C-L. and Van Ness, J.W. *Statistical Regression with Measurement Error.* London: Arnold, 1999.

Coates, J. *Suppressing Insurgency.* Colorado: Westview Press, 1992.

Egbert, R. *et al.* FIGHTER 1: An Analysis of Combat Fighters and Non Fighters. US HumRRO, *Task Report 14,* 1957.

Harrison, M.D. *et al.* An Interview Study of Human Relations in Effective Infantry Squads. US PRBS, *Report No 983,* 1953.

Hartley, D.S. and Helmbold, R.L. Validating Lanchester's Square Law and Other Attrition Models. *Naval Res. Logist.,* **42,** 609–633, 1995.

Holmes, R. *Acts of War.* London: Weidenfeld & Nicolson, 2003.

Marshall, S.L.A. *Men Against Fire: The Problem of Battle Command in Future War.* New York: Morrow & Co., 1967.

Rowland, D. The Use of Historical Data in the Assessment of Combat Degradation. *J. Opnl. Res. Soc.,* **38,** 149–162, 1987.

Rowland, D. *The Stress of Battle: Quantifying Human Performance in Combat.* London: The Stationary Office. 2006.

Rowland, D. and Richardson, A. Assessing the Effect of Heroism on Combat Effectiveness. 14th Intl. Symposium on Military Operational Research, Royal Military College of Science, Shrivenham. 1997.

Rowland, D. and Speight, L.R. Surveying the Spectrum of Human Behaviour in Front Line Combat. *Mil. Opns. Res.,* **12**(4), 47–60, 2007.

Speight, L.R. and Rowland, D. Modelling the Rural Infantry Battle: Overall Structure and a Basic Representation of the Approach Battle. *Mil. Opns. Res.,* **11**(1), 5–26, 2006.

Wigram, Lt Col L. Report to Directorate of Military Training, British North African Forces, 1943. (*Appendix II in* Forman, D. *To Reason Why.* London: Deutsch, 1991.)

About our Authors

DEFINING EFFECTS FOR PROBABILISTIC MODELING

by Dr. Mark A. Gallagher, Gregory J. Ehlers, Wesley D. True, and Marc R. Warburton

Mark Gallagher leads modeling, simulation, and analyses of resource issues. He has degrees from the Air Force Academy and the Air Force Institute of Technology. His current research interests include effects based operations, information operations, and cost analysis. He is serving a second term on the MORS Board of Directors.

Greg Ehlers currently works as an analyst at USSTRATCOM. He has degrees from Creighton University and the Air Force Institute of Technology. Greg has been a Co-Chair for the Effects Based Operations Special Sessions at the last two MORS Symposia. His research interest are linear programming and network modeling.

Wes True's current position as Future Capabilities Deputy Division Chief allows him access to USSTRATCOM efforts in Science and Technology, Experimentation, and Capability Integration across the command. He has degrees from St Michael's College in Winooski Vermont and the Air Force Institute of Technology. His current research interests include effects based operations and reliability estimation techniques. He has chaired the MORS Effects Based Operations Focus Group for two consecutive years.

Mr. Warburton is a senior analyst with SAIC's Strategic Deterrence Assessment Lab, which supports USSTRATCOM HQ/J525 Deterrence Assessment Branch. He has degrees in physics from the University of California at Davis and San Diego. He is currently researching applying game theory and Bayesian inference to effects-based deterrence decision making.

A COMPARISON OF MULTIVARIATE OUTLIER DETECTION METHODS FOR FINDING HYPERSPECTRAL ANOMALIES

Timothy E. Smetek and Kenneth W. Bauer Jr.

Timothy E. Smetek is an officer in the United States Air Force currently assigned to the Office of the Secretary of Defense, Program Analysis & Evaluation Directorate. He conducted is doctoral research at the Air Force Institute of Technology in the area of applied statistics and hyperspectral image analysis. He holds the following degrees: BS, Aerospace Engineering, University of Virginia, 1991; MAS, Embry-Riddle Aeronautical University, 1994; MS, Operations Research, Air Force Institute of Technology, 1998; PhD, Operations Research, Air Force Institute of Technology, 2007.

Kenneth W. Bauer is a Professor of Operations Research at the Air Force Institute of Technology where he teaches classes in applied statistics and pattern recognition. His research interests lie in the areas of automatic target recognition and multivariate statistics. Dr. Bauer holds the following degrees: BS, Miami University (Ohio), 1976; MEA, University of Utah, 1980, MS, Air Force Institute of Technology, 1981; PhD, Purdue University, 1987.

EFFICIENT EMPLOYMENT OF NON-REACTIVE SENSORS

by Moshe Kress, Roberto Szechtman, and Jason S. Jones

Moshe Kress teaches advanced combat modeling and probability models at the Naval Postgraduate School (NPS). Prior to his position at the NPS, he had been with the Center for Military Analyses in Israel. His research interests are combat modelling, homeland security problems, biodefense and logistics. He is the Military and Homeland Security Area Editor for *Operations Research* and an associate editor for NRL and MOR. His home page is http://www.nps.navy.mil/orfacpag/resumePages/Kress.htm.

Roberto Szechtman teaches simulation and probability models at the Naval Postgraduate School (NPS). His research interests are Simulation Theory, Applied Probability, and Military Operations Research. His homepage is http://www.nps.navy.mil/orfacpag/resumePages/Szechtman.htm

Lieutenant Commander Jason Jones is an Information Professional Officer in the US Navy. He earned a MS in Modeling, Virtual Environments, and Simulation (MOVES) from the Naval Postgraduate School in September 2006. He is currently

assigned to Carrier Strike Group SEVEN, serving as the Communications Officer for the RONALD REAGAN Strike Group.

REACH-BASED ASSESSMENT OF POSITION

by Lt Col J. Todd Hamill, USAF, Dr. Richard F. Deckro, Dr. Robert F. Mills, and Dr. James W. Chrissis

Lt Col Hamill is an operations research analyst at US Strategic Command. He holds a BSOR from the United States Air Force Academy, an MSIE from New Mexico State University, an MSOR and operations research PhD from Air Force Institute of Technology. His interests include mathematical programming, social network analysis, simulation, and fishing.

Richard F. Deckro is a Professor of Operations Research at the Air Force Institute of Technology. He holds a BSIE from the University of Buffalo, and an MBA and a Doctorate of Business Administration in Decision Sciences from Kent State University. He is particularly interested in the application of operations research to information operations and information assurance, reconstruction and stabilization, measures of effectiveness and assessment, behavioral modeling including social networks, and modeling fourth generation operations, counter insurgency, and irregular warfare.

Robert F. Mills is an Assistant Professor of Electrical Engineering at the Air Force Institute of Technology, Wright-Patterson Air Force Base, Ohio. He received his Ph.D. in electrical engineering from the University of Kansas in 1994, his MSEE from the Air Force Institute of Technology in 1987, and BSEE from Montana State University in 1983. His research interests are in communication systems, network management and security, information operations, and systems engineering. Dr. Mills is a member of Eta Kappa Nu and Tau Beta Pi and is a senior member of the IEEE.

James W. Chrissis is an Associate Professor of Operations Research at the Air Force Institute of Technology. He holds a B.S. in Mathematics from the University of Pittsburgh, and a M.S. and Ph.D. in Industrial Engineering and Operations Research from Virginia Tech. His research interests include engineering optimization, mathematical programming, simulation, and stochastic systems. Dr. Chrissis is a member of the Institute for Operations Research and Management Sciences (INFORMS), The Society for Industrial and Applied Mathematics (SIAM), the Military Operations Research Society (MORS), the American Institute for Aeronautics and Astronautics (AIAA), serving on the multidisciplinary optimization and analysis technical committee (MDOA-TC), and Sigma Xi.

MODELLING THE RURAL INFANTRY BATTLE: THE EFFECTS OF LIVE COMBAT ON MILITARY SKILLS AND BEHAVIOUR DURING THE APPROACH PHASE

by L. R. Speight and D. Rowland

After completing his UK national service in the army Ron Speight obtained a first degree in psychology. He then joined what was the precursor to the Army Personnel Research Establishment, working on a variety of human factor topics over a number of years. During a spell in Whitehall he undertook a project modelling anti-tank helicopter engagements. This prompted him to obtain a doctorate in Operational Research. Apart from a short spell in scientific and technical intelligence this led in turn to a career in military OR, culminating in seven and a half years as Chief of the OR Division of what was then the SHAPE Technical Centre in the Hague. Napoleon's much-quoted maxim has it that "In war, moral considerations account for three-quarters, the balance of forces only for the other quarter". Ron's career experiences led him to the belief that the community of would-be model users was relatively poorly served as regards representation of the effects of the most important three-quarters of battle. It therefore seemed natural on retirement to join forces with David Rowland in order to make a modest contribution in this area. This collaboration has led to a number of papers being published in *Military OR.*

Following engineering roles in London Transport bus chassis development and cutting tools for Wilkinson Sword, David Rowland moved on to model the wet-shaving market from historical data. He started his professional career in defence by modelling the off-road mobility of military vehicles, but soon moved into the field of OR. He had a long association with infantry and armoured warfare field trials, producing basic data needed to underpin many UK combat models. He then became a founding member of the Historical Analysis group in what was the Defence Operational Analysis Establishment. He has recently authored a book, *The Stress of Battle,* describing in brief a great deal of the work undertaken by himself and his colleagues from the 1970's onwards. His professional experiences have made him particularly well suited to undertake the investigation reported in this paper. Apart from the engagements that have yielded the findings that have been used to develop the combat behavior descriptors, he has studied the accounts of literally thousands of battles, involving different armies in different eras and in many theatres of war. Wherever possible the emphasis has been on the production of quantitative results. He has now finally retired, but he maintains his life long interest in military matters and the realities of actual combat.

A JOURNAL OF THE MILITARY
OPERATIONS RESEARCH SOCIETY
A JOURNAL OF THE MILITARY
OPERATIONS RESEARCH SOCIETY
A JOURNAL OF THE MILITARY
OPERATIONS RESEARCH SOCIETY
A JOURNAL OF THE MILITARY
OPERATIONS RESEARCH SOCIETY

Military Operations Research

Volume 13 Number 4
2008

A JOURNAL OF THE MILITARY
OPERATIONS RESEARCH SOCIETY
A JOURNAL OF THE MILITARY
OPERATIONS RESEARCH SOCIETY
A JOURNAL OF THE MILITARY
OPERATIONS RESEARCH SOCIETY
A JOURNAL OF THE MILITARY
OPERATIONS RESEARCH SOCIETY
A JOURNAL OF THE MILITARY
OPERATIONS RESEARCH SOCIETY
A JOURNAL OF THE MILITARY
OPERATIONS RESEARCH SOCIETY
A JOURNAL OF THE MILITARY
OPERATIONS RESEARCH SOCIETY
A JOURNAL OF THE MILITARY
OPERATIONS RESEARCH SOCIETY
A JOURNAL OF THE MILITARY
OPERATIONS RESEARCH SOCIETY
A JOURNAL OF THE MILITARY
OPERATIONS RESEARCH SOCIETY

MORS
Military Operations Research Society

Military Operations Research

A publication of the Military Operations Research Society

The Military Operations Research Society is a professional society incorporated under the laws of Virginia. The Society conducts a classified symposium and several other meetings annually. It publishes proceedings, monographs, a quarterly bulletin, *PHALANX*, and a quarterly journal, *Military Operations Research*, for professional exchange and peer criticism among students, theoreticians, practitioners and users of military operations research. The Society does not make or advocate official policy nor does it attempt to influence the formation of policy. Matters discussed or statements made in the course of MORS symposia or printed in its publications represent the opinions of the authors and not the Society.

Editor
Dr. Richard F. Deckro
Department of Operational Sciences
Air Force Institute of Technology
AFIT/ENS, Bldg 641
2950 Hobson Way
Wright-Patterson AFB, OH 45433-7765

Publisher
Corrina Ross-Witkowski, Communications Manager
Military Operations Research Society
1703 N. Beauregard Street, Suite 450
Alexandria, VA 22311

Editors Emeritus
Dr. Peter Purdue
Naval Postgraduate School

Dr. Gregory S. Parnell, FS
United States Military Academy

Editorial Board
Dr. Patrick D. Allen
General Dynamics Advanced Information Systems
Dr. Paul F. Auclair
LexisNexis
Dr. Stephen J. Balut
Institute for Defense Analyses
Dr. Jerome Bracken
IDA, RAND & Department of State
Dr. Alfred G. Brandstein
The MITRE Corporation
Dr. Gerald G. Brown
Naval Postgraduate School
Dr. Marion R. Bryson, FS
Consultant
Maj Stephen P. Chambal
AFRL
Dr. Charles A. Correia
Virginia Commonwealth University
Dr. Paul K. Davis
RAND and Pardee RAND Graduate School
Dr. Jerry Diaz
Homeland Security Institute
Prof. Patrick J. Driscoll
United States Military Academy
LTC Barry C. Ezell, Ph.D.
Army Capabilities Integration Center
Dr. Bruce W. Fowler
U. S. Army Research and Development Command
Dr. Mark A. Gallagher
HQ USAF/A9R
Dr. Dean S. Hartley, III
Hartley Consulting
Dr. Raymond R. Hill, Jr.
Wright State University
Prof. Wayne P. Hughes, Jr., FS
Naval Postgraduate School
Dr. Jack A. Jackson
Science Applications International Corporation
Dr. James L. Kays
Naval Postgraduate School
Dr. Moshe Kress
Center for Military Analyses, Israel & Naval Postgraduate School
LTC Mike Kwinn
US Military Academy
Dr. Andrew G. Loerch
George Mason University
Dr. James K. Lowe
US Air Force Academy
Mr. Brian R. McEnany, FS
Consultant
Dr. Michael L. McGinnis
VMASC
Dr. James T. Moore
Air Force Institute of Technology
Prof. W. Charles Mylander III
United States Naval Academy
Dr. Kevin Ng
Department of National Defence, Canada
Dr. Daniel A. Nussbaum
Naval Postgraduate School
Dr. Edward A. Pohl
University of Arkansas
Dr. Kevin J. Saeger
Los Alamos National Laboratory
Dr. Robert S. Sheldon, FS
Group W Inc.
Dr. Joseph A. Tatman
Innovative Decisions, Inc.
Mr. Eugene P. Visco, FS
Visco Consulting
Dr. Keith Womer
University of Missouri – St. Louis

Military Operations Research, the journal of the Military Operations Research Society (ISSN 0275-5823) is published quarterly by the Military Operations Research Society, 1703 N. Beauregard Street, Suite 450, Alexandria, VA 22311-1745. The domestic subscription price is for one year and for two years; international rates are for one year and for two years. Periodicals Postage Paid at Alexandria, VA. and additional mailing offices.

POSTMASTER: Send address changes to *Military Operations Research*, the journal of the Military Operations Research Society, 1703 N. Beauregard St, Suite 450, Alexandria, VA 22311. Please allow 4–6 weeks for address change activation.

Military Operations Research is indexed in the following Thomson ISI® services: Science Citation Index Expanded (SciSearch®); ISI Alerting Services®; CompuMath Citation Index®.

© Copyright 2008, Military Operations Research Society, 1703 N. Beauregard St, Suite 450, Alexandria, VA 22311. ISSN Number 0275-5823

Printed on Recycled Paper

MILITARY OPERATIONS RESEARCH SOCIETY
BOARD OF DIRECTORS

OFFICERS

President
John F. Keane*
Johns Hopkins University/APL
jack.keane@jhuapl.edu

President Elect
LTC Michael J. Kwinn* Jr
United States Military Academy
michael.kwinn@usma.edu

Vice President Finance and Management (F&M)
Joseph C. Bonnet*
The Joint Staff
joseph.bonnet@js.pentagon.mil

Vice President Meeting Operations (MO)
Kirk A. Michealson*
Lockheed Martin
kirk.a.michealson@lmco.com

Vice President Professional Affairs (PA)
Dr. Niki C. (Deliman) Goerger
United States Military Academy
nicolette.goerger@usma.edu

Secretary
William M. Kroshl
Johns Hopkins University/APL
william.kroshl@jhuapl.edu

Immediate Past President
Patrick J. McKenna*
USSTRATCOM/J821
mckennap@stratcom.mil

Executive Vice President
Brian D. Engler*
Military Operations Research Society
brian@mors.org

*** - Denotes Executive Council Member**

OTHER DIRECTORS

Dr. Patrick D. Allen, General Dynamics, UK Ltd.
Col John M. Andrew, USAFA/DFMS
Lt Col Andrew P. Armacost, HQ AFSPACE/A9
Dr. Michael P. Bailey, MCCDC
Col Suzanne M. Beers, AFOTEC Det 4/CC
Dr. Theodore Bennett, Jr., Naval Oceanographic Office
Mary T. Bonnet, AD, Innovative Management Concepts, Inc.
Dr. W. Forrest Crain, AD, Director Capabilities Integration, Prioritization and Analysis
William H. Dunn, AD, Alion Science and Technology
Helaine G. Elderkin, FS, AD, Computer Sciences Corporation
Dr. Karsten Engelmann, Center for Army Analysis
John R. Ferguson, AD, SAIC
Dr. Mark A. Gallagher, HQUSAF/A9R
Michael W. Garrambone, General Dynamics
Debra R. Hall, General Dynamics
LTC Clark H. Heidelbaugh, Joint Staff, J7
Robert C. Holcomb, Institute for Defense Analyses
Timothy W. Hope, AD, Whitney Bradley & Brown, Inc.
Dr. John R. Hummel, Argonne National Lab
Gregory T. Hutto, 53rd TMG/OA
Maj. KiraBeth Therrien, SAF/US
Gregory A. Keethler, AD, Lockheed Martin
Jane Gatewood Krolewski, USAMSAA
Dr. Lee Lehmkuhl, MITRE
Trena Covington Lilly, Johns Hopkins University/APL
Dr. Andrew G. Loerch, AD, George Mason University
Dr. Daniel T. Maxwell, Innovative Decisions
Dr. Willie J. McFadden II, AD, Booz Allen Hamilton
Lana E. McGlynn, AD, McGlynn Consulting Group
Dr. Gregory A. McIntyre, Applied Research Associates, Inc.
Terrence McKearney, The Ranger Group
Anne M. Patenaude, DUSD(R)/RTPP
Dr. Steven E. Pilnick, Naval Postgraduate School
Mark D. Reid, AD, MITRE
Cortez D. (Steve) Stephens, AD, MCCDC
Donald H. Timian, Army Test and Evaluation Command
Eugene P. Visco, FS, AD, Consultant
Corinne C. Wallshein, OSD/PA&E
COL Richard I. Wiles, USA-ret, FS, AD

MILITARY OPERATIONS RESEARCH SOCIETY
1703 N. BEAUREGARD STREET
SUITE 450
ALEXANDRIA, VA 22311
Telephone: (703) 933-9070

Editorial Policy and Submission of Papers

EDITORIAL POLICY

The title of our journal is *Military Operations Research*. We are interested in publishing articles that describe *operations research* (OR) methodologies used in important *military* applications. We specifically invite papers that are significant military OR applications. Of particular interest are papers that present case studies showing innovative OR applications, apply OR to major policy issues, introduce interesting new problem areas, highlight educational issues, and document the history of military OR. Papers should be readable with a level of mathematics appropriate for a master's program in OR.

All submissions must include a statement of the major contribution. For applications articles, authors are requested to submit a **letter** to the editor—exerpts to be published with the paper—from a **senior decision-maker** (government or industry) stating the benefits received from the analysis described in the paper.

To facilitate the review process, authors are requested to categorize their articles by **application area** and **OR method**, as described in Table 1. Additional categories may be added. (We use the MORS working groups as our applications areas and our list of methodologies are those typically taught in most OR graduate programs.)

INSTRUCTIONS TO *MILITARY OPERATIONS RESEARCH* AUTHORS

The purpose of the "instructions to *Military Operations Research* authors" is to expedite the review and publication process. If you have any questions, please contact Ms. Corrina Witkowski, MORS Communications Manager (email: morsoffice@aol.com).

General

Authors should submit their manuscripts (3 copies) to:

Dr. Richard F. Deckro
Military Operations Research Society
1703 N. Beauregard St, Suite 450
Alexandria, VA 22311-1717

Alternatively, manuscripts may be submitted electronically in Microsoft Word or Adobe Acrobat by emailing the manuscript and associated materials to richard.deckro@afit.edu *AND* to corrina@mors.org.

Per the editorial policy, please provide:

- authors statement of contribution (briefly describe the major contribution of the article)
- letter from senior decision-maker (application articles only)
- military OR application area(s)
- OR methodology (ies)

Approval of Release

All submissions must be unclassified and be accompanied by release statements where appropriate. By submitting a paper for review, an author certifies that the manuscript has been cleared for publication, is not copyrighted, has not been accepted for publication in any other publication, and is not under review elsewhere. All authors will be required to sign a copyright agreement with MORS.

Abbreviations and Acronyms

Abbreviations and acronyms (A&A) must be identified at their first appearance in the text. The abbreviation or acronym should follow in parentheses the first appearance of the full name. To help the general reader, authors should minimize their use of acronyms. If required, a list of acronyms can be included as an appendix.

TABLE 1: APPLICATION AREAS & OR METHODS

Composite Group	APPLICATION AREA	OR METHODOLOGY
I. STRATEGIC & DEFENSE	Strategic Operations Nuclear Biological Chemical Defense Arms Control & Proliferation Air & Missile Defense	**Deterministic Operations Research** Dynamic Programming Inventory Linear Programming Multiobjective Optimization Network Methods Nonlinear Programming
II. SPACE/C41SR	Operational Contribution of Space Systems Battle Management/Command and Control ISR and Intelligence Analysis Information Operations/Information Warfare Countermeasures Military Environmental Factors	**Probabilistic Operations Research** Decision Analysis Markov Processes Reliability Simulation Stochastic Processes Queuing Theory
III. JOINT WARFARE	Unmanned Systems Land & Expeditionary Warfare Littoral Warfare/Regional Sea Control Strike Warfare Air Combat Analysis & Combat ID Special Operations and Irregular Warfare Joint Campaign Analysis	
IV. RESOURCES	Mobility & Transport of Forces Logistics, Reliability, & Maintainability Manpower & Personnel	**Applied Statistics** Categorical Data Analysis Forecasting/Time Series Multivariate Analysis Neural Networks Nonparametric Statistics Pattern Recognition Response Surface Methodology
V. READINESS & TRAINING	Readiness Analytical Support to Training Casualty Estimation and Force Health Protection	
VI. ACQUISITION	Measures of Merit Test & Evaluation Analysis of Alternatives Cost Analysis Decision Analysis	
VI. ADVANCES IN MILITARY OR	Modeling, Simulation, & Wargaming Homeland Defense and Civil Support Computing Advances in Military OR Warfighter Performance and Social Science Methods Warfighting Experimentation	**Others** Advanced Computing Advanced Distributed Systems (DIS) Cost Analysis Wargaming

Length of Papers

Submissions will normally range from 10–30 pages (double spaced, 12 pitch, including illustrations). Exceptions will be made for applications articles submitted with a senior decision-maker letter signed by the Secretary of Defense.

Format

The following format will be used for dividing the paper into sections and subsections:

TITLE OF SECTIONS

The major sections of the paper will be capitalized and be in bold type.

Title of Subsections

If required major sections may be divided into subsections. Each subsection title will be bold type and be Title Case.

Title Subsection of a Subsection

If required subsections sections may be divided into subsections. Each subsection title will be Title Case. Bold type will not be used.

Paper Electronic Submission with Figures, Graphs and Charts

After the article is accepted for publication, an electronic version of the manuscript must be submitted in Microsoft Word or Acrobat. For each figure, graph, and chart, please include a camera-ready copy on a separate page. The figures, graphs, and tables should be of sufficient size for the reproduced letters and numbers to be legible. Each illustration must have a caption and a number.

Mathematical and Symbolic Expressions

Authors should put mathematical and symbolic expressions in Microsoft Word or Acrobat equations. Lengthy expressions should be avoided.

Footnotes

We do not use footnotes. Parenthetical material may be incorporated into a notes section at the end of the text, before the acknowledgment and references sections. Notes are designated by a superscript letter at the end of the sentence.

Acknowledgments

If used, this section will appear before the references.

References

References should be cited with the authors and year. For example, one of the first operations research texts published with several good military examples (Morse & Kimball, 1951).

References should appear at the end of the paper. The references should be unnumbered and listed in alphabetical order by the name of the first author. Please use the following format:

For journal references, give the author, year of publication, title, journal name, volume, number, and pages—for example:

Harvey, R.G., Bauer, K.W., and Litko, J.R. 1996. Constrained System Optimization and Capability Based Analysis, Military Operations Research, Vol 2, No 4, 5–19.

For book references, give the author, year of publication, title, publisher, and pages—for example:

Morse, P.M., and G.E. Kimball. 1951. Methods of Operations Research. John Wiley, 44–65.

For references to working papers or dissertations cite the author, title, type of document, department, university, and location, for example:

Rosenwein, M. 1986. Design and Application of Solution Methodologies to Optimize Problems in Transportation Logistics. Ph.D. Dissertation. Department of Decision Sciences, University of Pennsylvania, Philadelphia.

Appendices

If used, this section will appear after the reference.

United States Postal Service Statement of Ownership, Management and Circulation (PS Form 3526)

1. Publication Title: Military Operations Research
2. Publication Number: 0275-5823
3. Filing Date: 12 November 2008
4. Issue Frequency: Quarterly
5. Number of Issues Published Annually: 4
6. Annual Subscription Price: Domestic-$70.00/year and Foreign-$135.00/year
7. Complete Mailing Address of Known Office of Publication: Military Operations Research Society, 1703 N. Beauregard Street, Suite 450, Alexandria, VA 22311-1745
8. Complete Mailing Address of Headquarters or General Business Office of Publisher: Military Operations Research Society, 1703 N. Beauregard Street, Suite 450, Alexandria, VA 22311-1745
9. Full Names and Complete Mailing Addresses of Publisher, Editor, and Managing Editor
Publisher: Military Operations Research Society, 1703 N. Beauregard Street, Suite 450, Alexandria, VA 22311-1745
Editor and Managing Editor: Dr Richard Deckro, Military Operations Research Society, 1703 N. Beauregard Street, Suite 450, Alexandria, VA 22311-1745
10. Owner: Military Operations Research Society, 1703 N. Beauregard Street, Suite 450, Alexandria, VA 22311-1745 and Institute for Operations Research and the Management Sciences, 7240 Parkway Drive, Suite 310, Hanover, MD 21076
11.Known Bondholders, Mortgagees, and Other Security Holders Owning or Holding 1 Percent or More of Total Amount of Bonds, Mortgages, or Other Securities. None
12. Tax Status (For completion by nonprofit organizations authorized to mail at nonprofit rates). The purpose, function, and nonprofit status of this organization and the exempt status for federal income tax purposes: Has Not Changed During Preceding 12 Months
13. Publication Title: Military Operations Research
14.Issue Date for Circulation Data Below: August 4, 2008

15. Extent and Nature of Circulation	*Average No. Copies Each Issue During Preceding 12 Months*	*No. Copies of Single Issue Published Nearest to Filing Date*
a. Total Number of Copies	2,263	1,638
b. Paid and/or Requested Circulation		
(1) Paid/Requested Outside-County Mail Subscriptions Stated on Form 3541. (Include advertiser's proof and exchange copies)	2,021	1,394
(2) Paid In-County Subscriptions Stated on Form 3541(Include advertiser's proof and exchange copies)	0	0
(3) Sales Through Dealers and Carriers, Street Vendors, Counter Sales, and Other Non-USPS Paid Distribution	27	34
(4) Other Classes Mailed Through the USPS	0	0
c. Total Paid and/or Requested Circulation	2,048	1,428
d. Free Distribution by Mail Samples, complimentary, and other free)		
(1) Outside-County as Stated on Form 3541	0	0
(2) In-County as Stated on Form 3541	0	0
(3) Other Classes Mailed Through the USPS	0	0
e. Free Distribution Outside the Mail (Carriers or other means)	7	7
f. Total Free Distribution	7	7
g. Total Distribution	2,055	1,435
h. Copies not Distributed	208	203
i. Total	2,263	1,638
j. Percent Paid and/or Requested Circulation	99%	99%

16. This statement of Ownership will be printed in the December 2008 *issue of this publication.*