Judge's Commentary: The Outstanding Information Technology Security Papers

Frank Wattenberg
Dept. of Mathematical Sciences
United States Military Academy
West Point, NY 10996
Frank.Wattenberg@usma.edu

Introduction

The final judging for the 2004 Interdisciplinary Contest in Modeling took place at the United States Military Academy on Saturday, March 6, 2004. The judges spent an extensive and enjoyable day reading a very good and varied set of papers.

Bottom Line Up Front: There is Room at the Top

Although a number of submissions were very good and readers will recognize some well-known institutions among the Outstanding and Meritorious papers, there is room at the top. If this were a sporting event rated on a 10-point scale, it is quite likely that no one paper would have scored above 9.0. This IT Security Problem involved many complex issues, messy data, and several challenging tasks. Three points are crucial in addressing the requirements of this problem:

This is first and foremost a modeling competition. Modeling is often about ill-posed problems, in complex settings with uncertain data. Conclusions necessarily involve simplifications and uncertainties and confronting them is absolutely imperative. The papers were judged primarily on modeling.

The UMAP Journal 25 (2) (2004) 175–179.

The constraints of the contest are exactly the constraints in real life. In the real world, modelers always work with limited time and resources. Thus, real-world modeling requires making simplifications, justifying those simplifications, examining the impact of those simplifications and, above all, being intellectually honest about the shortcomings as well as the successes of the resulting models. Some submissions were marred by puffery.

Organization, clarity, and brevity are essential. The judges were surprised by the number of submissions that lacked a table of contents. Although a table of contents was not required, its omission usually reflected a general lack of organization. The summary too is particularly important for any report. Although many summaries were well-written, even the summaries in the Outstanding papers merited at best a grade of B; none talked about the potential shortcomings of their models due to modeling assumptions or uncertainties in the data.

The Problem

This year's problem dealt with information technology security for a new university campus. An undefended IT system is exposed to potential losses but, as usual, the costs of defense are considerable.

There are many possible approaches to this problem. The problem description focused on two categories of defenses—policies and technology.

- Policies include, for example, whether the network is wireless, as well as password policies—how complicated must passwords be and how often must they be changed.
- Technologies include things like firewalls and virus scanning.

The description also focused on risk in three areas—confidentiality, integrity, and availability. A breach of confidentiality can result in litigation or the costs associated with the release of proprietary or classified information. The integrity and availability of data and information are, of course, essential to their value.

In addition to a description of the structure of the situation, the problem included a 12-page enclosure with data about several alternatives in various categories—for example, it included data on eight different host-based firewalls. These data had two glaring features, and the judges looked specifically at how the submissions addressed the issues raised by these features:

- Alternative defensive measures were discussed individually with no information about how they might work in combination.

  The better submissions all addressed this issue at least briefly. In general, however, none of the submissions did an outstanding job. The fact is that there is a range of ways in which two alternatives might interact. For example, at one extreme, two different virus scanners might be completely redundant if they both picked up the same viruses; or, at the other extreme, they might protect from completely different viruses. The results are potentially very sensitive to whatever assumptions are made in this area.

This problem is compounded by the fact that assuming redundancy among measures in the same category reduces the computational complexity of the problem. Many submissions (including highly-rated ones) justified this assumption to make the problem computationally feasible. This is a reasonable assumption only if it is accompanied by a discussion of the sensitivity of the conclusions to this assumption.

- The data were based on multiple reviews of the measures and there was considerable variation in the conclusions of the various reviews.

  Here again, while most of the papers addressed this issue in some fashion, many of papers made simplifications without discussing the sensitivity of their conclusion to those simplifications.

Analysis

The different teams applied a variety of optimization techniques to their models. Some teams worked with models that were computationally infeasible and applied techniques—for example, simulated annealing—that led with relatively high probability to near optimal solutions; others made assumptions that led to computationally feasible optimization problems. Some teams used standard software and others wrote their own programs using C++ or other programming languages. Although some of the teams used sophisticated mathematics and algorithms (for example, simulated annealing) and others used sophisticated software effectively, neither was necessary for this problem. Many teams did first-rate work using straightforward implementations of their models with general-purpose tools.

The analytic part of this problem can be broken into two parts:

- evaluating the costs and the effectiveness of a mix of defensive measures, and
- searching the space of possible mixes of defensive measures to find an optimal or near optimal mix.

The first part rightfully drew the most attention in most of the submissions—this is the modeling part. This required considerable attention to details and to the extensive data provided. Most importantly, however, it required thoughtful analysis of the two difficulties mentioned earlier—the impact of combinations of defensive measures and how to handle the uncertainties in the data. This is also the first focus of the absolutely necessary sensitivity analysis. In its starkest form, an assumption that two measures are redundant leads to a recommendation that at most one measure should be employed, while an assumption that two measures cover disjoint sets of attacks may lead to a recommendation that both defensive measures should be employed.

The computational difficulty of the second part depended in part on the assumptions about how individual preventive measures interacted when used together. Other modeling assumptions also impacted this part of the problem. For example, some teams assumed that the same mix of defensive measures was used across the university, whereas others broke the problem up into different subnetworks. The new university's computing needs are diverse—ranging from student computers in dormitory rooms, to the commercial needs of a bookstore whose business skyrockets at the beginning of each semester, to the registrar's office and student health services that routinely deal with confidential data. In addition, the sophistication and professionalism of users is also very diverse—the registrar's office, bookstore, and student health services, for example, are more likely to accept stringent security measures than individual students, who might want to be able to install software of questionable origin.

We saw a wide variety of approaches to searching for an optimal or near-optimal solution and most had considerable merit. Here again, we focused on the implications of the underlying modeling assumptions and on an analysis of the sensitivity of the conclusion to the search procedure used in addition to the modeling assumptions.

Conclusions and Advice to Future Teams

This section is essentially an amplification of the same points made by Richard Cassady last year [2003, 188].

Assumptions Making simplifying assumptions is a critical part of modeling. In fact, good models are always the result of an iterative procedure beginning with fairly drastic simplifying assumptions to obtain some initial traction and then building progressively more sophisticated models based on sensitivity analysis and reality checks. Articulate your assumptions and their consequences. Your summary must identify clearly the assumptions made and their impact on your conclusions.

Analysis Analysis is not the last step. It is an integral part of the iterative modeling procedure. Do regular reality checks and above all use sensitivity analysis to guide your model development and to determine both the strengths and weaknesses of your conclusions.

Communication You must express and communicate your work well. Clarity of expression is a consequence of clarity of thought. If your summary and your paper are not clear then the modeling is almost certainly weak.

References As always, use proper citation and be careful about the provenance and worth of the work you use.

Congratulations are extended to all the participants on their accomplishments. Reading and judging the results of their weekend of interdisciplinary problem solving and modeling were enjoyable challenges for the judges.

Reference

Cassady, C. Richard. 2003. Judge's Commentary: The Outstanding Airport Screening Papers. *The UMAP Journal* 24 (2) 185–188.

About the Author

Frank Wattenberg is a professor in the Dept. of Mathematical Sciences at the United States Military Academy (USMA), West Point. He is particularly interested in modeling and simulation and in the use of technology for simulation and for education across the undergraduate curriculum. He is currently leading a team at the USMA that is developing on Online Book *Modeling in a Real and Complex World* to be published as part of the MAA Online Book Project. He is also working with colleagues at USMA and elsewhere to develop rich immersive environments for modeling and simulation. This project will produce environments with both virtual and hands-on components that students will revisit from middle school through college and from many different subject areas and levels. The architecture will support collaborative modeling and simulation based in part on the ideas of multiplayer games.