Judge's Commentary: The Outstanding Contest Judging Papers

Donald E. Miller
Department of Mathematics
Saint Mary's College
Notre Dame, IN 46556
dmiller@saintmarys.edu

This year's problem, while framed in the terminology of a competition such as the MCM, finds application in several other areas of decision-making. One such area, itself a competition of sorts, occurs when a school makes its decisions on the award of scholarships. Another occurs in the screening of applicants for a specified position. In both of these situations, the potential exists for more nominally qualified applicants than positions. Thus, those making the decision, the judges in our problem, must either rank-order or otherwise quantify the applicants in an attempt to decide which ones are "best." Further, these decisions must be made under time constraints that make it impossible for each "judge" to evaluate every applicant; even if they were to do so, it is doubtful that the evaluations would be in complete agreement. It is this element that complicated evaluation of the contest papers.

The assumption of an absolute rank-ordering made the problem seem deceptively easy, resulting in a broad range of papers, from quite simple to very elaborate. At one end of the spectrum, we found papers that assumed absolute rank-ordering and simply developed a heuristic solution to the basic problem; some even recognized the assumption was unrealistic but chose to model the problem as stated, since that is what was requested. Others who used this assumption clearly didn't believe it, since they rejected a simple merge-sort in favor of more complicated algorithms that denied existence of the assumption. Still others attempted to refine the problem using theories ranging from topics in graph theory to concepts of fuzzy sets. Some of the better bias-elimination refinements included matrix reduction, regression with error terms, and scoring normalization with specified probability distributions.

The judges felt that the best models were those that solved the basic problem for 100 papers, 8 judges, and 3 winners, then produced a successful refinement with adequate complexity to accurately model the process but with enough simplicity for the model to be useful. Further, each model would be clearly stated and its use demonstrated with a simple example. Thus, the ideal paper would solve the basic problem and demonstrate that the solution was optimal, or at least close to optimal. It would then generalize the solution to accommodate different numbers of papers, judges, and winners. Having completed this,

it would address judge bias and measure the success rate of the algorithm as a function of some quantitative measure of judge bias. It might then examine alternative algorithms, finding a relation between levels of judges bias and the success rate of these algorithms. It would continue by addressing the strengths and weaknesses of these methods, while being sure to clearly address all the points requested in the contest rules.

While many papers showed much insight into the problem and its complexities, the Outstanding papers were distinguished by the way that they addressed the problem of judge bias effectively, and with appropriate documentation, to allow for implementation of the recommended model. The team from the University of Science and Technology of China used Bayesian statistics, with a normal prior, to adjust for judge bias, then ran simulations to test for sensitivity of their method. Normal distributions, a common assumption, were also used by the Washington University team for both the intrinsic score of the paper and judge bias. A distinguishing feature of the paper from Fudan University was its stability analysis for different levels of judge bias.

Finally, here's a note on good practice in mathematical modeling as related to this problem. Common in requests for models are unrealistic assumptions that would make any model created with these assumptions of minimal use. Thus, it is necessary for the modeler to evaluate critically all assumptions and, if necessary, refine the problem to one that is realistic. Judges viewed the statement, "there is an absolute rank-ordering to which all judges would agree," to be such an assumption. Under this assumption, with 100 papers and 8 judges, there are methods of finding the top three papers, in order, with each judge reading at most 14 papers and with at most 109 papers read. In another situation and under different rules, the modeler would ask the author for elaboration on the statement before proceeding with the model. But in the absence of such consultation, the modeler should answer the question as stated and then refine it with realistic assumptions.

About the Author

Donald Miller is Associate Professor and of Chair of Mathematics at Saint Mary's College. He has served as an associate judge of the MCM for four years and prior to that mentored two Meritorious teams. He has done considerable consulting and research in the areas of modeling and applied statistics. He is currently a member of SIAM's Education Committee and past president of the Indiana Section of the Mathematical Association of America.