Wisdom of the Crowds in Minimum Spanning Tree Problems

Sheng Kung Michael Yi (skyi@uci.edu)
Mark Steyvers (steyver@uci.edu)
Michael D. Lee (mdlee@uci.edu)
Department of Cognitive Sciences,
University of California, Irvine
Irvine, CA 92697, USA

Matthew Dry (matthew.dry@adelaide.edu.au)
Discipline of Pharmacology
University of Adelaide
Adelaide, SA 5005, Australia

Abstract

The 'wisdom of the crowds' effect describes the finding that combining responses across a number of individuals in a group leads to aggregate performance that is as good as or better than the performance of the best individuals in the group. Here, we look at the wisdom of the crowds effect in the Minimum Spanning Tree Problem (MSTP). The MSTP is an optimization problem where observers must connect a set of nodes into a network with the shortest path length possible. A method is developed that creates aggregate solutions based only on the nodes connected in individuals' solutions, without access to spatial information about the nodes. Across the three problems analyzed, the solutions produced by the aggregation method perform better than even the best individual, leading to a strong wisdom of the crowds effect. We show this effect can be observed even with sample sizes as small as 6 individuals.

Keywords: Wisdom of the Crowds; Minimum Spanning Tree Problem; Decision Making; Problem Solving

Introduction

When a problem is posed to a group of individuals, a variety of answers or solutions may be returned. If the accuracy of the individual solutions is unknown, it would be useful to have the ability to extract the collective wisdom contained in the collection of individual responses by aggregating their solutions. The idea that an aggregate solution will perform better than the majority of individuals in the group is referred to as the 'wisdom of the crowds' effect (Surowiecki, 2004). Unlike most research in the topic of distributed cognition and collective intelligence (see Goldstone & Gureckis, 2009 for an overview), where individuals are able to interact in some fashion, individuals in a wisdom of the crowds environment tend to operate independently of one another. Despite this independence and the fact that group members may have widely varying levels of proficiency, aggregation can be found to be effectual in a number of scenarios.

The wisdom of the crowds effect has traditionally been demonstrated for simple questions for which there is a single answer. For example, Galton (1907) asked a large number of individuals to estimate the weight of an ox. He found that the median estimate for the weight of the ox was within 1% of the ox's actual weight. Similarly, Surowiecki (2004) reports that, when polled, the modal answer given by the audience in the US version of the game show "Who Wants To Be A Millionaire" for multiple choice questions is correct more than 90% of the time.

Recently, the wisdom of the crowds idea has also been applied to more complex problems. Steyvers, Lee, Miller, and Hemmer (2009) demonstrated the wisdom of the crowds effect for ordering problems, such as ordering a list of ten states from east to west, ordering the first ten amendments to the U.S. Constitution, or remembering the order of U.S. Presidents. For ordering data, simply taking the mode of individual answers can be problematic because, in many cases, all of the individual orderings are unique. Instead, Steyvers et al. (2009) developed several Bayesian aggregation models that looked at the underlying consistencies in the individuals' orderings to produce an aggregated solution.

A wisdom of the crowds effect has also been observed recently by Yi, Steyvers, Lee, and Dry (submitted), for a difficult combinational optimization problem known as the Traveling Salesman Problem (TSP: see Applegate, Bixby, Chvátal, & Cook, 2006 for a review). In the TSP, the goal is to connect a set of nodes to make the shortest path possible, with the constraints that each node can be visited only once, and the path must end at the same node as it started. The aggregation method developed by Yi et al. (submitted) did not require any spatial information about the locations of the nodes. Instead, the method took advantage of the knowledge of which nodes are connected in individual solutions and selected a solution that maximized the agreement across individuals as to the sequence of nodes visited.

Generating a wisdom of the crowds effect for TSP problems in this way provides an example of a potentially powerful and general approach to aggregating individual knowledge and abilities. The key feature is that all of the aggregation is based on the observed ordering of individuals and their patterns of agreement. No representation was needed of the complex multidimensional TSP stimuli, nor were evaluation measures for individual performance used. For these reasons, the results of Yi et al. (submitted) suggest

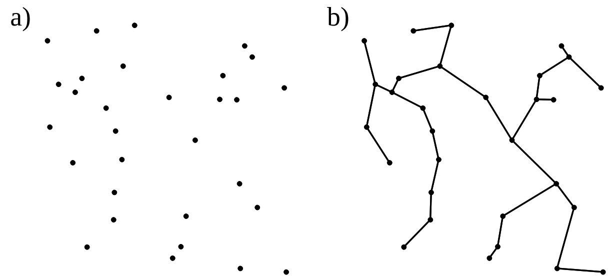

Figure 1: An example MSTP solution (a) and its optimal solution (b).

an approach to finding the wisdom of the crowd in challenging real-world situations where the problem space is too large or complicated to represent formally, and there is no clear way to quantify the merits of proposed solutions.

Of course, however, it may be that the TSP result is simply a special or isolated case. Accordingly, in this paper, we explore the possibility of a wisdom of the crowds effect for another complex problem solving task, known as the Minimum Spanning Tree Problem (MSTP). First, as for TSPs, we develop an aggregation method that is based on easily observed features of individual solutions. Then, we apply the method to previously collected data for several MSTPs. We observe a strong wisdom of the crowds effect, in which the aggregate solution is closer to optimal than any individual solution. Finally, we examine how many individual solutions are needed for good aggregation, and discuss how our approach could be extended, modified, and applied to more general problems.

Minimum Spanning Tree Problems

In MSTPs, participants are required to find the shortest possible network that links together a set of nodes in some spatial configuration. An example stimulus and optimal solution for an MSTP is shown in Figure 1. In contrast to the TSP, there is no constraint on the paths that can be formed. Each node can be connected to multiple nodes. The optimal solution is an open, branching path system or tree, in which nodes can be linked to one or more other stimulus nodes.

Finding the optimal solution to MSTPs has an obvious real-world engineering application in regards to finding the minimal length network of cables or pipes needed to join discrete geographical locations (e.g., Borůvka, 1926). However, MSTPs are also of interest from a psychological perspective, providing insight into human decision-making, individual differences in cognitive abilities, and visuo-perceptual organization (e.g., Burns, Lee & Vickers, 2006; Vickers, Mayo, Heiman, Lee & Hughes, 2004). Specifically, the MSTP belongs to a class of difficult visual optimization problems such as the TSP and the Generalized Steiner Tree Problem (GSTP). Despite the apparent difficulty (and in some cases intractability) of these optimization problems, human observers are often able to find optimal or close-to-optimal solutions in a time frame that increases as a linear function of problem size (e.g., Dry, Lee, Vickers & Hughes, 2006; Graham, Joshi, & Pizlo, 2000).

An important finding from the literature on human solutions to MTSPs is that there are meaningful individual differences (e.g., Burns et al., 2006). As Surowiecki (2004) and others have emphasized, a precondition for the wisdom of the crowds effect is that there is variation between individuals. Intuitively, the hope is that some individuals complete some parts of an MSTP optimally or near-optimally, while other individuals complete different parts well. In this scenario, the aggregation of the individual solutions could potentially improve on both.

Dataset

The data were taken from Burns et al (2006). In that study, as part of a larger battery of optimization tasks and cognitive abilities tests, 101 participants completed 3 MSTPs, with 30, 60 and 90 nodes. The problems were comprised of black nodes on a uniform white background and were presented on color computer monitors.

The participants generated spanning trees by pointing and clicking with the mouse cursor, and were allowed to add or remove links as they saw fit. They were instructed to connect the nodes by making a system of links, using as many links as they felt necessary, under the condition that the resulting system had the minimum overall possible length. The participants worked without time limits and were asked to be as accurate as possible. The results of the empirical solutions are displayed in Figure 2, expressed as the percentage above optimal solution length (PAO = 100*[empirical length/optimal length - 1]). Participants provided solutions that were on average around 6% longer than the optimal solution. Importantly however, there were significant individual differences with some individuals providing solutions that were much closer to the optimal solution. Despite the large number of participant solutions available, there was no case in any problem where any participant’s solution exactly matched that of another participant.

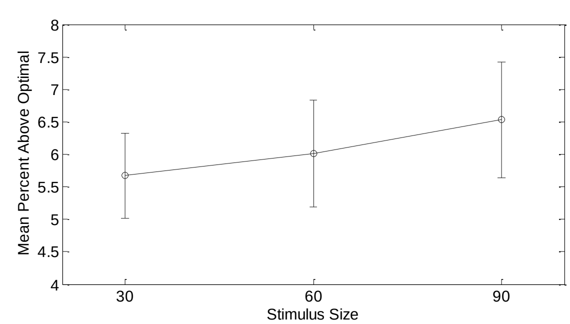

Figure 2. Mean empirical PAO for MSTP with 30, 60 and 90 nodes; error bars indicate standard deviation of individual performance.

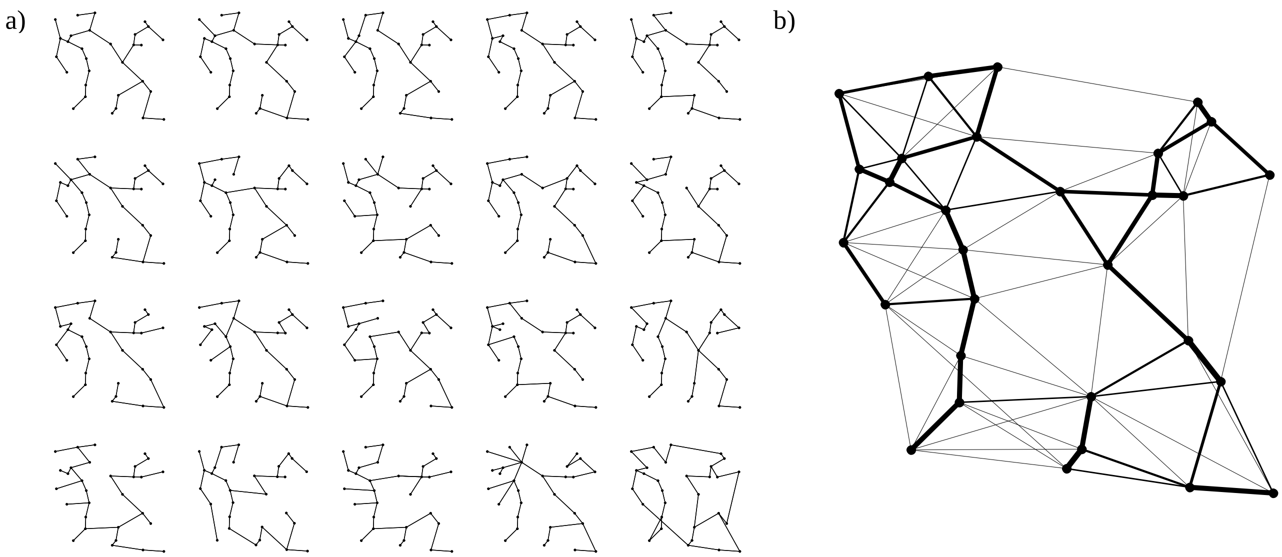

Figure 3. a) Representative subject solutions for the 30-node MSTP, the best subject solution in the upper left with decreasing performance across rows and the worst subject in the lower right. b) Visualization of agreement matrix on problem nodes. Vertices selected by at least one subject are drawn; thicker lines indicate higher agreement.

Aggregation Method

The data for the aggregation method were restricted to the information of which nodes each participant connected in their solutions. In particular, the method was not given any spatial information about the node locations, and so relied solely on the information contained in the participant solutions to create a proposed network. The aggregation method operates under the assumption that vertices between nodes that are better for inclusion in a MSTP solution tend to be selected by more participants. An aggregate solution that maximizes the degree of agreement with participant solutions can therefore be expected to have good performance.

In order to obtain an aggregate solution, we first arranged the solutions of all individuals in an $n \times n$ agreement matrix, where n is the number of nodes in the problem. Every cell a_{ij} in the matrix records the number of participants that connected nodes i and j in their solutions. A visualization of the agreement matrix is depicted in Figure 3b. We then derived a cost matrix of the same size with cell values $c_{ij} = k - a_{ij}$, where k was the total number of participants; connections with higher agreement would thus have lower costs. This cost matrix is then used as the input to a standard MSTP algorithm to obtain a proposal solution for the aggregate.

The MSTP can be solved optimally in polynomial time through the use of simple greedy algorithms such as Prim's algorithm (Jarník, 1930; Prim, 1957). In Prim's algorithm, a starting node is randomly selected from all nodes. At each step in the algorithm, the vertex with the smallest cost that connects an unconnected node to the already-connected nodes (or starting node, in the first step) is added to the network, until all nodes are connected. Despite the fact that the algorithm is greedy in nature, it is always guaranteed to output the minimum spanning tree depending on the cost metric being used. When the vertex costs are equal to the distances between nodes, Prim's algorithm is guaranteed to produce a spanning tree with the shortest total length. In our research, the vertex costs upon which Prim's algorithm is

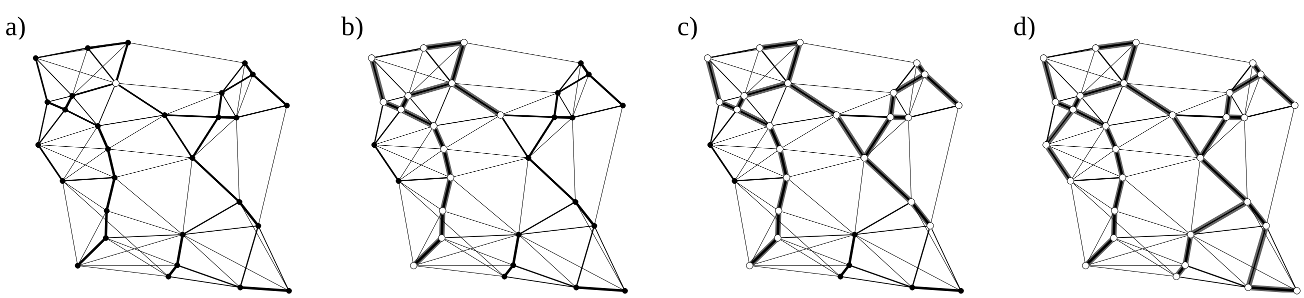

Figure 4. Example demonstration of Prim's algorithm on the 30-node MSTP. A random node is selected, shown in white (a.). At each step of the algorithm, vertices with the smallest cost (i.e., highest agreement) that connect an unconnected node (black) to those already connected (white) are added to the network until all nodes are connected (b-d.)

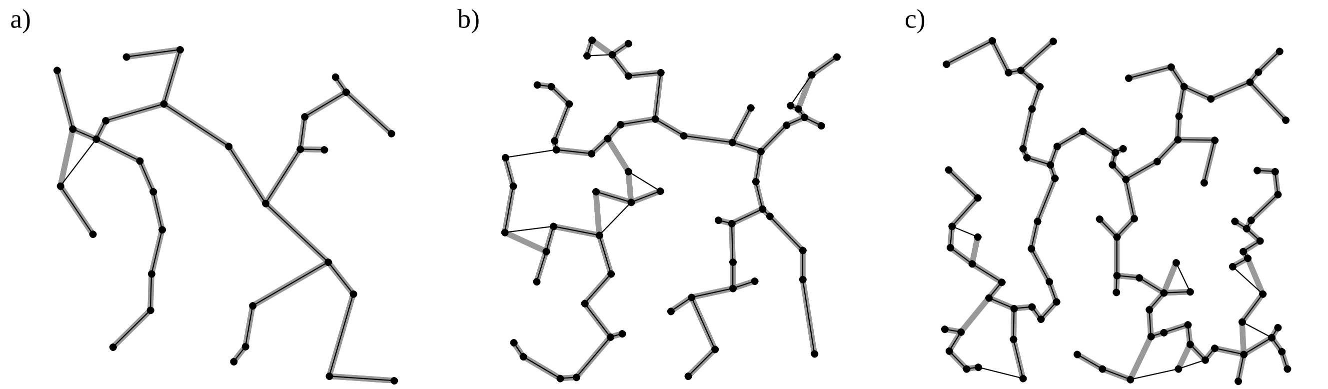

Figure 5. Solution paths for the aggregate method (thin black) and the optimal minimum spanning tree (thick gray) for the a) 30-node, b) 60-node, and c) 90-node MSTPs.

applied are set using the cost matrix based on subject agreement above. The algorithm will still produce a network with minimum total cost, but in this case, the network represents the spanning tree that has the highest agreement with the participant solutions. It is this solution that is selected by the aggregation method. A demonstration of the algorithm is shown in Figure 4.

Optimality of Prim's algorithm can be verified by considering the necessary conditions for a minimum spanning tree. For a solitary node, it is necessary for it to connect to its nearest neighbor using the vertex with the lowest cost. If a spanning tree is created without using such a vertex, and that node is connected to the others via some other vertex, it does not change the connectedness of the network by deleting that other vertex and instead connecting to the nearest neighbor, but it does reduce the total path length. This makes the first step of Prim's algorithm, connecting a random node to its nearest neighbor, a sensible action. The logic can be followed by induction to the sub-networks drawn by Prim's algorithm by treating each sub-network as if it were a single node, thus showing optimality. In cases where multiple potential vertices with the same cost may be selected for addition to the spanning tree, then any of the candidates may be chosen without affecting the solution's optimality.

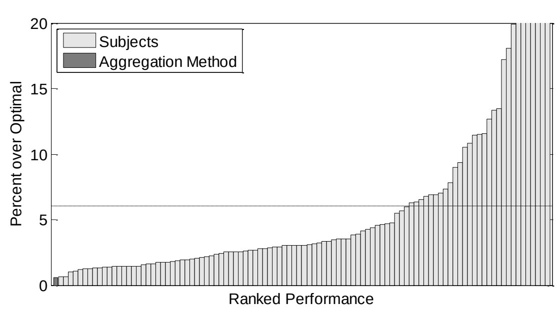

Figure 6: Ranked performance of subjects and the aggregation method averaged over all problems. Dashed horizontal line indicates mean subject performance.

Results

Figure 5 shows the optimal minimum spanning trees in thick gray lines and solutions selected by the aggregation method in thin black lines while participant and aggregate solution performance is provided in Table 1. Additional performance statistics are noted for the aggregate solutions: the amount of agreement the aggregate solutions had with subject solutions and a count of the number of participants whose performance is better than, worse than, or same as the aggregate. Subject agreement values were calculated as the proportion of subject vertices coinciding with vertices present in the aggregate solution; these can be obtained by noting the value of the aggregate path as measured on the agreement matrix, then dividing by $(n\text{-}1)k$, the number of vertices multiplied by the number of subjects. The aggregate

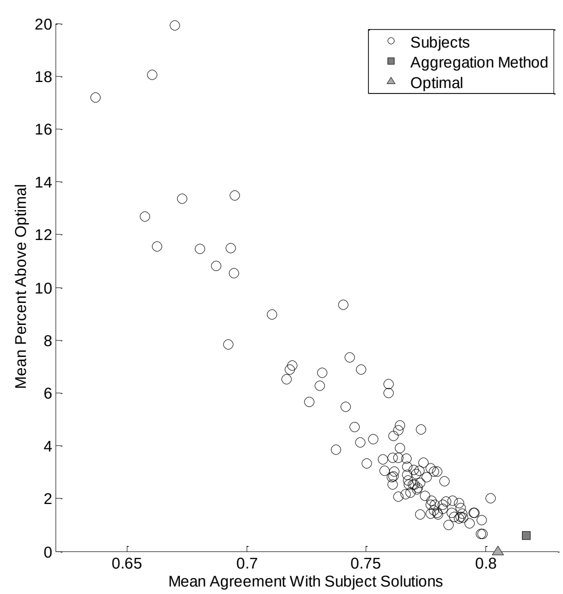

Figure 7: Performance averaged over all problems against mean solution agreement with subject solutions.

Table 1: Subject and Aggregate Method Performance on MSTPs (% network length over optimal)

Problem	subject performance		aggregate method performance				
	subj. best	subj. mean	path length	subj. agreement	# subj. better	# subj. same	# subj. worse
30 nodes	+0.000%	+5.672%	+0.059%	.7856	1	0	100
60 nodes	+0.037%	+6.010%	+1.410%	.8263	21	0	80
90 nodes	+0.235%	+6.533%	+0.310%	.8392	1	0	100
Overall	+0.644%	+6.072%	+0.593%	.8171	0	0	101

method solutions perform quite well, beating the average participant by a large margin. In the 30- and 90-node problems, the performance of the aggregate is bested only by a single participant out of the full set of 101. The aggregate performs relatively worse in the 60-node problem, but still better than most individuals. When performance is averaged over all problems, the aggregate performs better than any individual (Figure 6). Interestingly, the proportion of vertex agreement with participants increased with problem size, and solutions selected by the aggregate did not completely match any single individual on any problem. Figure 7 contains a plot of solution performance against the proportion of agreement with participant solutions averaged equally over all problems for all subjects, the optimal solution, and the aggregate solution. There is a clear correlation between individual performance and the amount of agreement their solutions had with other participants (r = -.9602). The optimal solution also has a high rate of coincidence with participant solutions, more than any individual.

Performance of the aggregation method under smaller sample sizes was also investigated. For each sample taken, subjects were selected randomly from the full dataset and aggregate solutions were created for all problems, their performances compared to the subjects in the sample that generated them. In cases where Prim's algorithm encountered a choice between vertices of the same cost, one was chosen at random to create the proposal solution. Solution performance for selected sample sizes is noted in Figure 8, averaged over 1000 random draws at each sample size. We find that for samples of as small as size 6, the aggregate is able to obtain performance that is, on average, significantly better than the mean subject and close to that of the best subject in the sample. Averaged over all problems, the aggregate was outperformed by about one participant at all sample sizes investigated. In certain cases for individual problems, the aggregate solution outperformed all participants in the sample; this was much more common for the 30-node and 90-node problems than the 60-node problem.

Conclusions

We have demonstrated a strong wisdom of the crowds effect for the MSTP using a simple aggregation method on participant solutions. The aggregation method was reliant only on the knowledge of which nodes were connected by each participant, requiring no information regarding the spatial characteristics of the problems themselves. In addition, the simple greedy algorithm used to generate solutions required no input parameters to run. The aggregation method solutions generally have performances ranking among the best participants on individual problems, and perform better than any individual when averaged over all problems. Even when the number of available participants was reduced down to as low as 6, the aggregation method was still able to extract enough information to propose solutions that produced performances significantly better than the mean subject and exceeding most or all participants in the sample.

While performance of the aggregation method is quite good, there are potential areas for expanding on the method. It was noted that there was a clear correlation between a

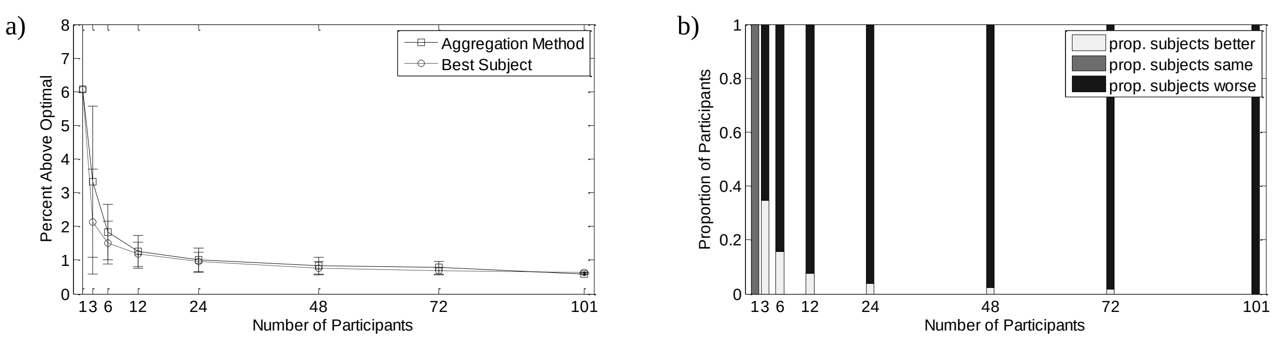

Figure 8. Performance of the aggregate method for selected sample sizes, taken across problems. a) Mean PAO for aggregate and best subject in each sample, error bars indicate standard deviation of individual samples. b) Proportion of subjects with better, same, or worse performance than the aggregate.

participant's performance and the amount of agreement they had with other participants. It may be useful if it were possible to identify 'experts' in the data and weight their responses over that of others. This approach of amplifying expertise may be most useful for when sample sizes are small. Due to the fact that there are so few participant solutions to draw from, there may be many networks that can potentially be chosen by the algorithm that share the same agreement with participant solutions, but carry very different performances in terms of actual distance. If participants can be weighted differently, then there will be less ambiguity. However, with the complexity of the problem structure, it is a difficult problem to create a formal system in which this can be done.

More generally, the results presented here, when coupled with those presented by Yi et al. (submitted) for the TSP, suggests that it may be possible to achieve wisdom of the crowds effects for complicated and only partly defined problems. While the MSTP does have a simple solution algorithm, and the TSP has good approximate solution algorithms for small numbers of nodes, our results show that near-optimal performance can be obtained from simple properties of the sub-optimal sets of solutions produced by a group of people.

In other words, our results show that there is an alternative route to solving these problems, not based on complicated algorithms, detailed stimulus information, and precise performance metrics. Instead, we have shown that the orders people produce can be combined to achieve near-optimality. Of course, for TSPs and MSTPs, there is not much reason to go to the effort of collecting human solutions when good algorithms are available. But our approach will continue to apply for different sorts of difficult problems where, for example, the stimuli or problem space is hard to represent in a formal way. This representational burden is borne by the individual providing solutions, and there is no need for any formal attempt to characterize the problem space. Even more intriguingly, our approach will apply in situations, such as some types of aesthetic judgment, where people agree on what constitutes a good answer once it is produced, but cannot define exactly what metric they are using. Since our aggregation approach just uses the patterns of relationships between individual judgments, and does not need a performance measure, it is equally applicable to these poorly defined problems.

We are currently investigating the use of the wisdom of the crowds approach described in this paper to the "wisdom of the crowds within", the idea that one can aggregate over multiple judgments from a single individual to obtain performance better than the individual judgments alone (Vul & Pashler, 2008). By applying transformations to MSTPs, we can easily test an individual on multiple repetitions of the same problem while minimizing bias from their responses on previous trials. We are also looking at applying the aggregation approach to a less-well defined aesthetic judgment task. Participants were asked in Dry, Navarro, Preiss, and Lee (2009) to connect point stimuli based off of constellations into perceived structures. It is possible that a structure created by aggregating over individuals is perceived as more aesthetically pleasing than individual patterns. The application of our approach to aggregation to these sorts of challenging problems seems a promising direction for further wisdom of the crowds research.

References

Applegate, D. L., Bixby, R. E., Chvátal, V., & Cook, W. J. (2006). *The Traveling salesman problem: A computational study*. Princeton, NJ: Princeton University Press.

Borůvka, O (1926). O jistém problému minimálním. *Práce Moravské Přírodovědecké Společnosti*, 3, 37–58.

Burns, N. R., Lee, M. D., & Vickers, D (2006). Are Individual Differences in Performance on Perceptual and Cognitive Optimization Problems Determined by General Intelligence? *Journal of Problem Solving*, *1*(1), 5-19.

Dry, M. J., Lee, M. D., Vickers, D., & Hughes, P. (2006). Human Performance on visually presented traveling salesperson problems with varying numbers of nodes. *Journal of Problem Solving*, 1, 20-32.

Dry, M.J., Navarro, D.J., Preiss, K., Lee, M.D. (2009) The Perceptual Organization of Point Constellations. In N. Taatgen, H. van Rijn, J. Nerbonne, & L. Shonmaker (Eds.), *Proceedings of the 31st Annual Conference of the Cognitive Science Society*, 1151-1156. Austin, TX: Cognitive Science Society.

Galton, F. (1907). Vox Populi. *Nature*, *75*, 450-451.

Goldstone, R. L., Gureckis, T. M. (2009) Collective Behavior. *Topics in Cognitive Science*, *1*, 412-438.

Graham, S. M., Joshi, A., & Pizlo, Z. (2000). The Traveling Salesman Problem: A hierarchical model. *Memory & Cognition*, *28*(7), 1191-1204.

Jarník,V (1930). O jistém problému minimálním. *Práce Moravské Přírodovědecké Společnosti*, 6, 57-63.

Prim, R. C. (1957). Shortest connection networks and some generalizations. *Bell System Technical Journal*, 36, 1389-1401.

Steyvers, M., Lee, M.D., Miller, B., & Hemmer, P. (2009). The Wisdom of Crowds in the Recollection of Order Information. In J. Lafferty, C. Williams (Eds.), *Advances in Neural Information Processing Systems*, *23*. MIT Press.

Surowiecki, J. (2004). *The Wisdom of Crowds*. New York, NY: W. W. Norton & Company, Inc.

Vickers, D., Mayo, T., Heitmann, M, Lee, M. D., & Hughes, P. (2004). Intelligence and individual differences in performance on three types of visually presented optimization problems. *Personality and Individual Differences*, *36*, 1059-1071.

Vul, E. & Pashler, H. (2008). Measuring the Crowd Within: Probabilistic Representations Within Individuals. *Psychological Science*, *19*(7), 645-647.

Yi, S. K. M., Steyvers, M., Lee, M. D., & Dry, M. J. (submitted). Wisdom of the Crowds in Traveling Salesman Problems.