Wisdom of the Crowds in Traveling Salesman Problems

Sheng Kung Michael Yi

Mark Steyvers

Michael D. Lee

Department of Cognitive Science

University of California, Irvine

Matthew J. Dry

Discipline of Pharmacology

University of Adelaide

Abstract

The phenomenon of the ‘wisdom of the crowds’ refers to the finding that the aggregate of a set of proposed solutions from a group of individuals performs better than the majority of individual solutions. We investigated this effect in the context of planar Euclidean traveling salesperson problem (TSP). The goal in TSPs is to estimate the shortest tour through a number of cities, represented as points in a two-dimensional display. We develop and apply an aggregation method that finds a single tour by combining the solutions from a group of individuals. Despite the fact that the aggregation method ignores spatial information, we demonstrate for most of the TSP problems that the aggregate solution tends to be closer to the optimal solution than the majority of individual solutions. Averaged across all of the TSP problems, we observe a strong wisdom of crowds effect where the averaged performance of the aggregation method outperforms even the best individual.

Wisdom of the Crowds in Traveling Salesman Problems

When judgments are made by a group of individuals, the judgment obtained by aggregating their judgments is often as good as, or might even be better than, the best individual in the group. This phenomenon, known as a wisdom of the crowds effect, is usually demonstrated for simple tasks such as estimating physical quantities or providing answers to multiple choice questions (see Surowiecki, 2004, for an overview). More recently, the effect has also been demonstrated within a single individual (Vul & Pashler, 2008). In this research, it is shown that averaging multiple estimates given by the same individual at different points in time can lead to a better answer than the individual estimates themselves. Once again, however, the focus of this work is on answering simple general knowledge questions, which have a single number as their correct answer.

An important challenge for wisdom of the crowds research involves its application to problems that require much more complicated, multi-dimensional answers. Recently, for example, Steyvers, Miller, Hemmer and Lee (2009) found the wisdom of the crowds effect for ordering problems, such as listing chronologically the US presidents, or ranking cities according to their number of inhabitants. For these sorts of problems, it is not usually possible to take a simple mean or mode of individual answers to obtain a group answer. Instead, some sort of account of how people solve the problem, and the nature of individual differences, are both needed to develop an aggregation method. In this way, to tackle combinatorially complicated problems, wisdom of crowds modeling needs input from the theories and methods of cognitive science.

In this paper, we investigate the wisdom of crowds for a classic complex problem solving task from computer science and operations research --- which has also been a recent topic of study in cognitive science --- known as the Traveling Salesperson Problem (TSP). We develop a method for combining individual human solutions to TSP problems, and then measure the performance of the aggregate solutions relative to the individual solutions.

In TSPs, a number of nodes, or ‘cities’, must be visited in a closed cycle that visits each node once, with the goal of minimizing the distance covered over the total path. The TSP serves as a classic example of an NP-complete problem, where computationally scalable solution methods for guaranteed optimal solutions are not known (Applegate, Bixby, Chvátal, & Cook, 2006). As the problem size grows, optimal solution methods quickly require infeasible computational resources. Instead, in order to get close to optimal performance, various approximation algorithms are employed (e.g., Helsgaun, 2000, 2009). Despite the computational complexity present in TSPs, the evidence from studying human performance is that people are able to create solutions quickly, while still maintaining good performance, for at least some versions of the problem. In particular, for planar Euclidean TSPS (i.e., those where the cities can be represented as points in a low-dimensional space), people seem able to complete TSPs in approximately linear time over problem sizes (Dry, Lee, Vickers & Hughes, 2006; Graham, Joshi & Pizlo, 2000). This comes in contrast to computational approaches, whose solution times tend to be at least on the order $O(n \ln n)$ with problem size.

The solutions generated by people seem to follow some basic principles consistently. They tend to connect cities along the convex hull and avoid making intersections in the path, heuristics that promote good performance (MacGregor & Omerod, 1996; van Rooij, Stege & Schactman, 2003; MacGregor, Chronicle & Omerod, 2004). There is also evidence that human solvers are sensitive to proximity between cities, generally connecting cities with their nearest neighbors over others, (Vickers, Mayo, Heitmann, Lee & Hughes, 2004). TSP solutions have even been linked to the automatic perception of minimal structures and aesthetics. When people are asked to evaluate solutions to TSPs in terms of aesthetics, the solutions that are evaluated higher tend to also be those that have shorter lengths (Vickers, Lee, Dry, Hughes & McMahon, 2006). Earlier research by Vickers, Butavicius, Lee, and Medvedev (2001) also found similarities between solution paths created by subjects whose given goals were to create aesthetically pleasing circuits and paths created by subjects who performed the standard TSP task.

Despite these general principles often being followed, however, there is also evidence for stable and significant individual differences in human TSP performance. While early results gave conflicting accounts of the level and nature of individual differences (e.g., MacGregor & Ormerod, 1996, Vickers et al. 2001), a recent reconciliation seems to have been reached, which argues for the presence of individual differences at least for sufficiently complicated problems (Chronicle, Macgregor, Lee, Ormerod & Hughes 2008). The prospect of individual differences in human TSP solutions makes it a potentially fruitful application for the wisdom of crowds idea. In particular, it raises the question of whether it is possible to combine individual solutions to find a group solution that is closer to optimal than all, or the majority, of the individual solutions.

In this paper, we use previously collected data to test the wisdom of crowds idea for TSPs. In these data, each individual independently generated a solution to a given

TSP. We develop and apply a method for aggregating these individual solutions in order to create a single aggregate tour that captures the commonalities of the individual tours. We propose an aggregation process that is restricted in two important ways. First, we assume that the cost function to evaluate the quality of a solution is not available until *after* the final aggregate solution is proposed. Therefore, in this situation, it is not possible to refine the solution iteratively during the aggregation process in order to optimize the tour distance[1]. This restriction is important because, otherwise, it would be possible to ignore the human solutions altogether and just directly optimize the tours using computational means. The goal here is to see what information is collectively contained in the human solution, and the absence of the cost function during aggregation ensures that the human solutions are the only available source of information. Second, we assume that the aggregator does not have access to any spatial information, such as the location of cities. The only information available is the order in which the cities are visited on the tours proposed by a group of individuals. This restriction allows us to propose relatively simple aggregation procedures that analyze which cities tend to be connected by individuals, regardless of their spatial layout. If we can demonstrate good performance in spite of this restriction, we can argue that important information is contained in the order the cities are visited. By focusing on order information, we can develop more domain general aggregation techniques that might generalize better to other complex optimization problems that do not rely on a spatial configuration of points.

Dataset

[1] Human solvers of TSPs typically do not have access to the cost function either during problem-solving. The quality of the solution becomes known only after the individual submits the final solution.

The data analyzed in this work was collected and reported by Vickers, Bovet, Lee, and Hughes (2003). A brief summary of the experiment follows, and more details can be found in the original work. A total of 83 students from the University of Adelaide community volunteered to serve as individuals in the experiment. Individuals completed 10 traveling salesman problems in each of 2 different conditions. In one condition, individuals had to complete closed tours, where a cycle visiting each city once had to be created, starting and ending at the same city. In the other condition, individuals had to complete open tours, where a path visiting each city once had to be created, starting and ending at different cities. Problems were comprised of thirty cities placed in a square array, with each coordinate location for each city independently drawn from a uniform distribution. Individuals completed problems using a computer interface that allowed them to connect cities in any order, offering great flexibility in the strategies they could use. The optimal path, and percentage length the individual's path exceeded that of the optimal, were displayed after each problem to try and maintain task motivation.

For the analysis presented here, we used the individuals' solutions in the closed tours condition only. Within the ten problems performed in the closed condition, three were excluded from the analysis due to a high proportion of individuals obtaining optimal solutions (0.157, 0.301, and 0.410). For these problems, the wisdom of crowds idea is less interesting, since it is easy to select the best path based on simple agreement measures, and there is no potential to perform better than the best individual.[2] An optimal solution was found by participants in three of the remaining seven problems, but at most two participants on each, and the modal response often contained only two or three

[2] The aggregation method described in this article has been verified to return the optimal solution for the three excluded problems.

participants, and never contained as many as ten participants. Thus, for the seven problems on which we focus, there is no obvious agreement among participants, and clear room for improving their solutions.

Aggregation Method

We develop an aggregation method in which we first analyze how individuals tend to connect cities locally on their tours. We expect that the good local connections between cities will tend to be selected by more individuals than those connections which are part of bad solutions. We then propose to find good aggregate solutions by maximizing the collective agreement across cities on a tour.

Specifically, we first collect all individuals' solutions into a $n \times n$ agreement matrix, where n is the number of cities in the problem. Each cell a_{ij} in the matrix records the proportion of individuals that connect cities i and j. Instead of operating directly on the agreement proportions, we first apply a nonlinear monotonic transformation function on the agreement matrix values to transform agreements into costs. We focus on the function $c_{ij} = 1 - I^{-1}_{a_{ij}}(b_1, b_2)$, where $I^{-1}_{a_{ij}}(b_1, b_2)$ is the inverse regularized beta function with parameters b_1 and b_2, each taking a value of at least 1. A plot of the cost function for selected parameter values is shown in Figure 1. Costs range from 0 to 1, with higher agreements leading to lower costs. When $b_1 = b_2 = 1$, there is a linear relationship between agreement and cost. As we increase the parameter values, the cost function becomes more nonlinear which allows us to threshold the agreement values. This allows us to map agreement values above some threshold to a relatively low cost and those below a threshold to relatively high costs. The ratio between b_1 and b_2 influences the relative utility of selecting high-agreement connections and avoiding low-agreement connections.

The goal of the aggregation method is to propose tours that minimize the sum of the costs across all the cities that are connected on the tour. This problem is itself a TSP but instead of minimizing a sum of Euclidean distances, the goal is to minimize a sum of costs. Because the costs can be asymmetric and do not obey the regularities of Euclidian distances, this version of the TSP cannot be solved using many traditional TSP solvers. Instead, we solve for the lowest-cost paths using the LKH program, which solves TSPs using the Lin-Kernighan heuristic (Helsgaun, 2000, 2009)[3]. While the heuristic is not guaranteed to produce the optimal solution for extremely large problems, for small problems such as those being observed in this article, the heuristic implementation is able to consistently produce the optimal solution.

Results

Examples of paths chosen by the aggregation method for particular TSPs can be seen in Figure 2b. Black lines show the routes selected by individuals, with thicker lines corresponding to connections chosen by more individuals. Shaded lines show the tour selected by the aggregation method, using the optimized parameter values of $b_1 = b_2 = 3$ (see later discussion on the effect of parameter values). Figure 3 shows an alternate representation of a path selected by the aggregation method on the raw agreement matrix. Cells indicate the number of individuals electing to connect the row- and column-cities,

[3] The Lin-Kernighan heuristic operates as a generalization of a 2-opt algorithm. In a 2-opt algorithm, pairs of links in a problem tour are inspected. If there exists a pair of connections between the cities connected by those links that has a shorter length while maintaining the full tour, then the tour is changed to include those connections. The algorithm continues until no further improvement can be found. The Lin-Kernighan heuristic improves on the basic algorithm by looking at the general λ-opt algorithm, selecting an appropriate λ on each step.

with shaded cells corresponding to those selected in the aggregation method path. The upper-right subplot is taken from the raw counts according to the city indices as generated by the problem and the lower-left subplot sorts the cities in order of the aggregation method path. The aggregation method tends to include most of the paths with the highest agreement, though there are some paths with moderate amounts of agreement that are not selected if there are alternate paths that would have a smaller total cost.

We evaluated the solutions generated by the aggregation method in two ways. First, we evaluated it as compared to the optimal solution, measuring how much longer the proposed path is than the optimal path. Secondly, we compared this result to the path lengths of solutions provided by subjects, measuring how many subjects perform better, same, or worse than the aggregation method solution.

Figure 4 shows the performance of the aggregation method over a range of parameter values, together with the performance of individuals. On the vertical axis, performance is given as a percentage indicating how much longer the average path given by an individual or aggregate solution was than the optimal for the seven problems. On the horizontal axis, the proportion of connections in solution paths matching with other individuals across problems is given. Each circle indicates an individual, squares indicate aggregation method performances, and the cross on the horizontal axis shows the performance of the optimal algorithm. It is clear that there was a strong correlation between individual performance and path agreement between subjects, $r = -.897$, suggesting that the aggregation method should select solution paths that will perform relatively well.

The performance of the aggregation method over parameter values holds up to these expectations. Table 1 shows the performance of the aggregation method as compared to that of individuals. For parameter values $b_1 = b_2 = 1$, the aggregation method comes up with good solutions for four of the seven problems, but for three, the method selects a connection that no individual selects, resulting in paths that have lengths that perform worse than the median subject. As we increase the parameter values, the method is deterred from selecting connections not selected by any individual and performance improves. Meanwhile, the selected paths do not lose much in terms of agreement with individuals in the raw calculation. Figure 5 shows the performance of the aggregation method over a range of parameter values. The shading of each marker corresponds to performances plotted in Figure 4; darker markers correspond to better performance. Performance is maximized when the total parameter value is high and b_1 is at least as large as b_2. With these parameter values, the aggregation method favors a strategy of avoiding low-agreement connections in order to create good proposal paths. While the best solution paths for the aggregation method never correspond to the optimal path, they tend to rank among the best individuals. Most compelling, when considered over all of the available TSP problems, the aggregation method performs better than any individual. This is a demonstration of a strong wisdom of the crowds effect, with the aggregate solutions outperforming all of the individual solutions from which they are generated.

An interesting follow-up result considers how many individual solutions are needed to achieve good aggregate performance. If we limit the number of individuals available to the aggregation method, we can still obtain solution paths with good performance. Figure 6 shows estimates of the mean performance of the aggregation method for parameter values $b_1 = b_2 = 3$ for selected individual sample sizes. For sample sizes as small as 12 individuals, performance of the aggregation method can be expected to exceed that of the best individual in both the sample taken as well as the full dataset. For smaller sample sizes, the aggregation method solution’s performance is not expected to exceed that of the best individual in the sample taken, but the solution path can still extract enough information from the sampled paths to improve upon the performance of the average subject in the full dataset.

Discussion

Our results show that the aggregation method we have developed and applied for TSP problems is able to demonstrate wisdom of the crowds effects. Solution paths proposed by the method are created solely based on the combined city connections selected by individuals and are independent of spatial information regarding city locations. Despite the limited information available, paths selected by the aggregation method perform at a level that is among the best individuals on individual problems, and exceeds performance of the best individual when averaged over all problems.

One possible direction for future research, for those situations where people solve multiple TSP problems, involves identifying the better performed individuals, and increasing their contribution to the aggregate solutions. This identification of ‘experts’ can continue to be done without explicit feedback from a cost function, because, as we have observed, the better individuals tend to have higher agreements with the solutions of others. The challenge is to infer and share this information about expertise across all the problems, in some sort of hierarchical model.

A second potential line for future research is to consider TSP problems in the context of within-individual wisdom of the crowds research, sometimes known as “the crowd within” (Vul & Pashler, 2008). The basic idea would be to consider multiple solutions from the same person on the same TSP problem, and test whether the aggregation of these repeated solutions leads to better performance. One nice methodological feature of this problem is that, unlike general knowledge questions, it is relatively easy to test a person on multiple versions of the same TSP problem, by applying distance-preserving transformations to the visual problem representation.

More generally, given our demonstration of a wisdom of the crowds effect in human performance on the Traveling Salesman Problem, it seems possible or likely that there may be other complex problems in which an aggregation approach may be viable to find good solutions[4]. In particular, the niche for application is to problems that are difficult to solve by computational means but nonetheless can be solved reasonably well, with some inherent variability in performance, by people. Such problem might arise through combinatorial intractability, or through inherent challenges in formalizing aspects of a problem sufficiently to make them suitable for algorithms. Problems with these challenges may be able to be solved using the wisdom of the crowds idea, which cuts through the inherent computational complexity, the uncertainties in the specifics of a problem, or individual discrepancies in utility function for problem outcomes.

[4] One problem that may be of interest is in the perception of point constellations studied by Dry, Navarro, and Lee (2009). Participants in this task were asked to connect sets of points together according to the structure each individual perceived. These point sets corresponded to Ptolemaic constellations, as found in star atlases. The interest here is to see if an aggregation method can take in the individual solutions and retrieve structures that are close to those identified by Ptolemy in a wisdom of the crowds fashion. In this situation, the goals are not well defined and a best solution is subjective. Over individuals, however, we may be able to use aggregation to select a structure that has reasonable agreement with the group as well as the ‘accepted’ structure devised by Ptolemy.

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.

Chronicle, E.P., MacGregor, J.N., Lee, M.D., Ormerod, T.C., & Hughes, P. (2008). Individual differences in optimization problem solving: Reconciling conflicting results. *Journal of Problem Solving*, *2*(1), 41-49.

Dry, M., 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*(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.

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

Haxhimusa, Y., Kropatsch, W. G., Pizlo, Z., Ion, A. (2009). Approximative graph pyramid solution of the E-TSP. *Image and Vision Computing*, *27*(7), 887-896.

Helsgaun, K. (2000). An effective implementation of the Lin-Kernighan traveling salesman heuristic. *European Journal of Operational Research*, *126*, 103-130.

Helsgaun, K. (2009). General *k*-opt submoves for the Lin-Kernighan TSP heuristic. *Mathematical Programming Computation*, *1*, 119-163.

MacGregor, J.N., Chronicle, E.P., Ormerod, T.C. (2004). Convex hull or crossing avoidance? Solution heuristics in the traveling salesman problem. *Memory & Cognition*, *32*, 260-270.

MacGregor, J.N., Ormerod, T. (1996) Human performance on the traveling salesman problem. *Perception & Psychophysics*, *58*(4), 527-539.

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.

van Rooij, I., Stege, U., Schactman, A. (2003). Convex hull and tour crossings in the Euclidean traveling salesman problem: Implications for human performance studies. *Memory & Cognition*, *31*(2), 215-220.

Vickers, D., Bovet, P., Lee, M.D., Hughes, P. (2003). The perception of minimal structures: Performance on open and closed versions of visually presented Euclidean traveling salesman problems. *Perception*, *32*(7), 871-886.

Vickers, D., Butavicius, M., Lee, M., Medvedev, A. (2001). Human performance on visually presented Traveling Salesman problems. *Psychological Research*, *65*, 34-45.

Vickers, D., Lee, M.D., Dry, M., Hughes, P. (2003). The roles of the convex hull and the number of potential intersections in performance on visually presented traveling salesman problems. *Memory & Cognition*, *31*(7), 1094-1104.

Vickers, D., Lee, M.D., Dry, M., Hughes, P., McMahon, J.A. (2006). The aesthetic appeal of minimal structures: Judging the attractiveness of solutions to traveling salesperson problems. *Perception & Psychophysics*, *68*(1), 32-42.

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.

Table 1

Individual and Aggregate Performance on TSPs.

Problem	Subject performance			Aggregation performance, $b_1 = b_2 = 1$				Best aggregation performance, $b_1 = b_2 = 3$			
	# subj. optimal	subj. min	subj. mean	path length	# subj. better	# subj. same	# subj. worse	path length	# subj. better	# subj. same	# subj. worse
A30.14	2	+0.000%	+3.246%	+0.491%	11	8	64	+0.491%	11	8	64
A30.24	2	+0.000%	+4.791%	+1.434%	16	2	65	+1.434%	16	2	65
A30.48	0	+0.078%	+5.936%	+6.953%	55	0	28	+0.159%	1	1	81
B30.04	0	+0.121%	+5.502%	+5.494%	52	0	31	+0.193%	2	2	79
B30.11	1	+0.000%	+4.992%	+0.162%	1	5	77	+0.162%	1	5	77
B30.21	0	+0.043%	+5.325%	+0.043%	0	3	80	+0.043%	0	3	80
B30.27	0	+1.229%	+5.497%	+4.965%	46	0	37	+3.331%	22	2	59
All	0	+1.718%	+5.041%	+2.792%	14	0	69	+0.830%	0	0	83

Figures

Figure 1. Transformation functions from agreement matrix to cost matrix from the family $c_{ij} = 1 - I^{-1}_{a_{ij}}(b_1, b_2)$, where $I^{-1}_{a_{ij}}(b_1, b_2)$ is the inverse regularized beta function with parameters b_1 and b_2, for sample values of b_1 and b_2.

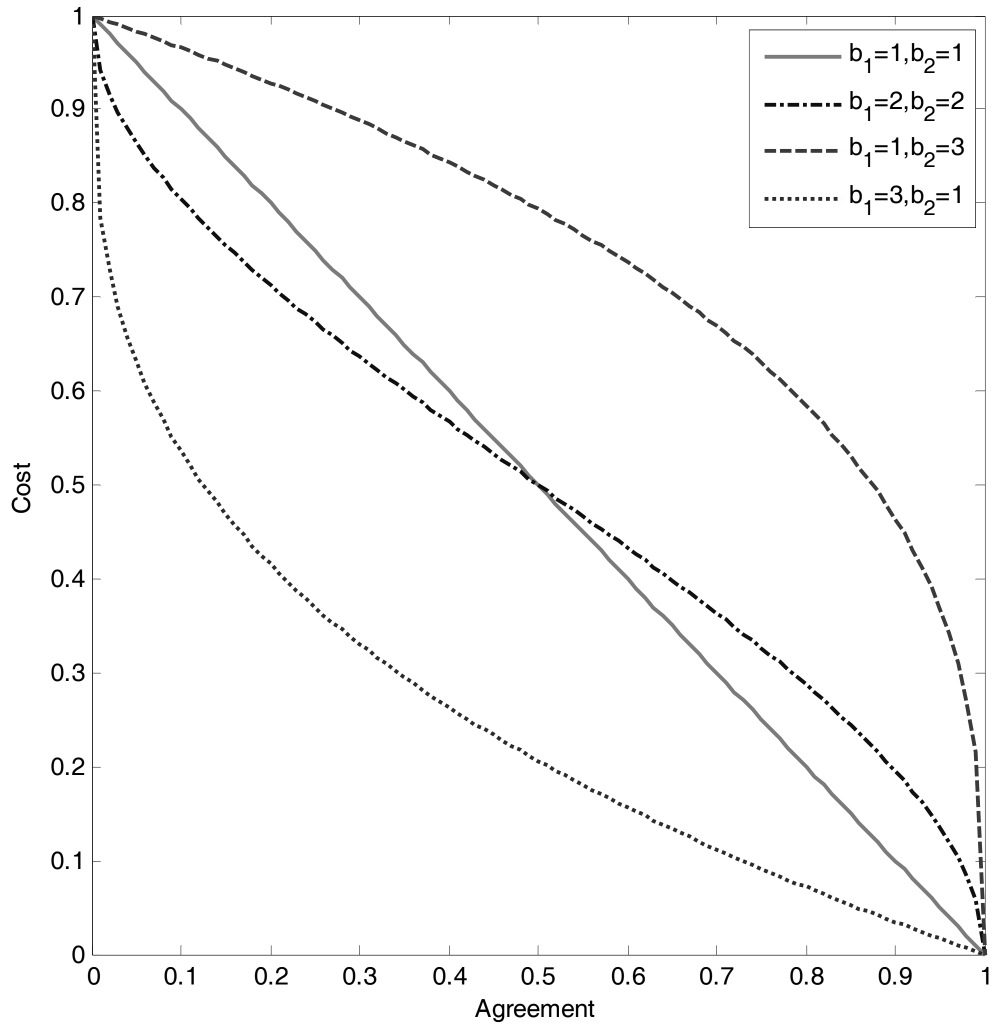
Figure 2. Sample Traveling Salesperson Problem, showing (a) the optimal solution, and (b) the aggregated subject paths in black, where thicker lines indicate more agreement, with aggregation method-selected path in gray.

Figure 3. Agreement matrix samples for problem B30.27. Cell values indicate the number of individuals who chose each path. The upper-right panel (a) shows the unsorted matrix and the lower-left panel (b) shows the sorted matrix according to the path selected by the best-performing aggregation method parameters. Shading is provided in each part of the figure to show selected connections.

Figure 4. Task performance vs. path agreement with other subjects over all problems. Circles show individual subject performances, squares show the aggregation method performance over parameter values, and the ‘x’ mark indicates how the optimal paths compare to subjects.

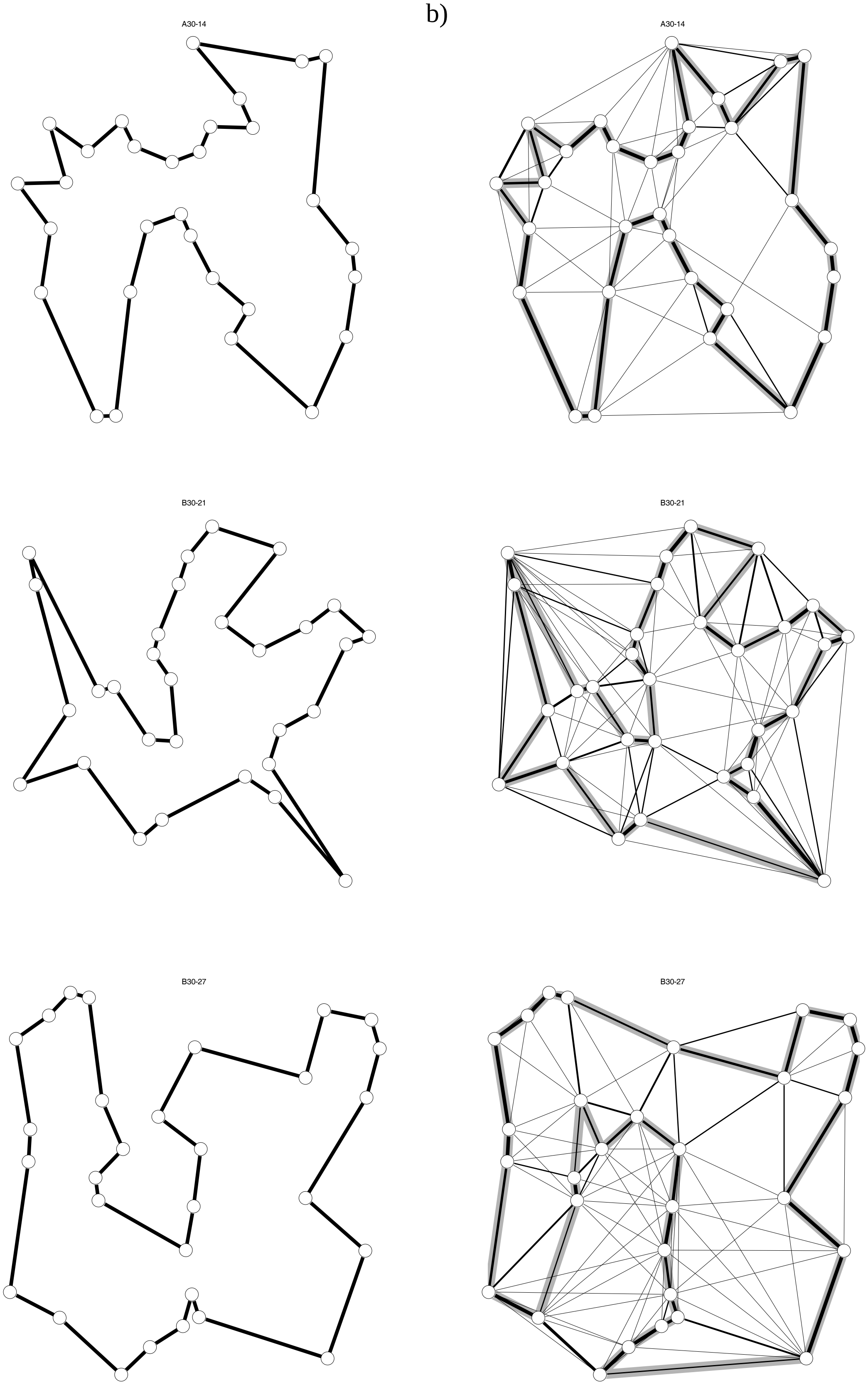
Figure 5. Aggregation method performance over parameter values. Darker and larger squares indicate better performance, corresponding to the shades in Figure 4.

Figure 6. Mean task performance vs. individual sample size for the aggregation method with parameters $b_1 = b_2 = 3$. Error bars extend one standard deviation in each direction of the mean for each sample size. Dashed line shows the expected performance of the best subject taken from a sample.

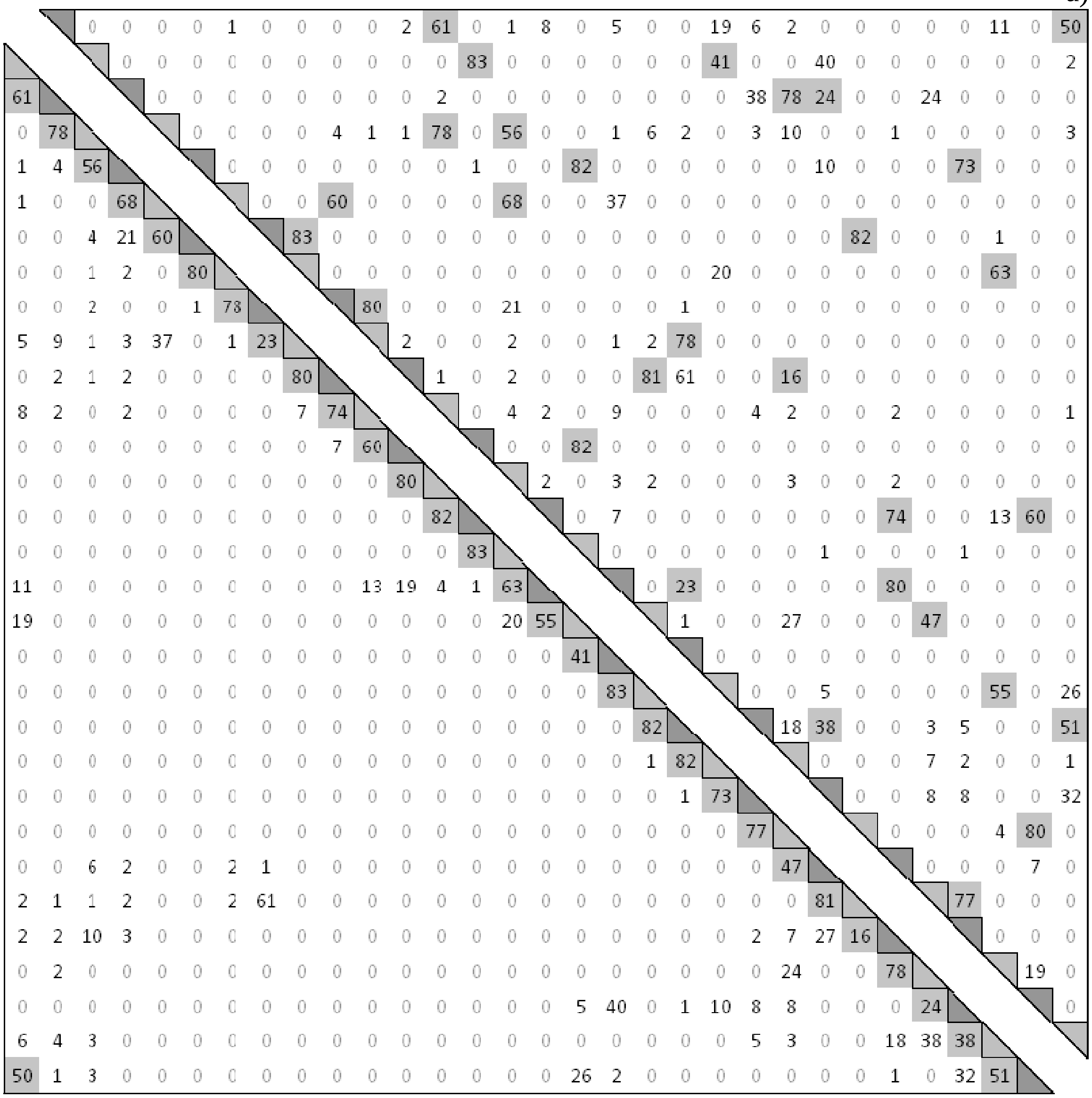


b1=1,b2=1
b1=2,b2=2
b1=1,b2=3
b1=3,b2=1
Cost
Agreement
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

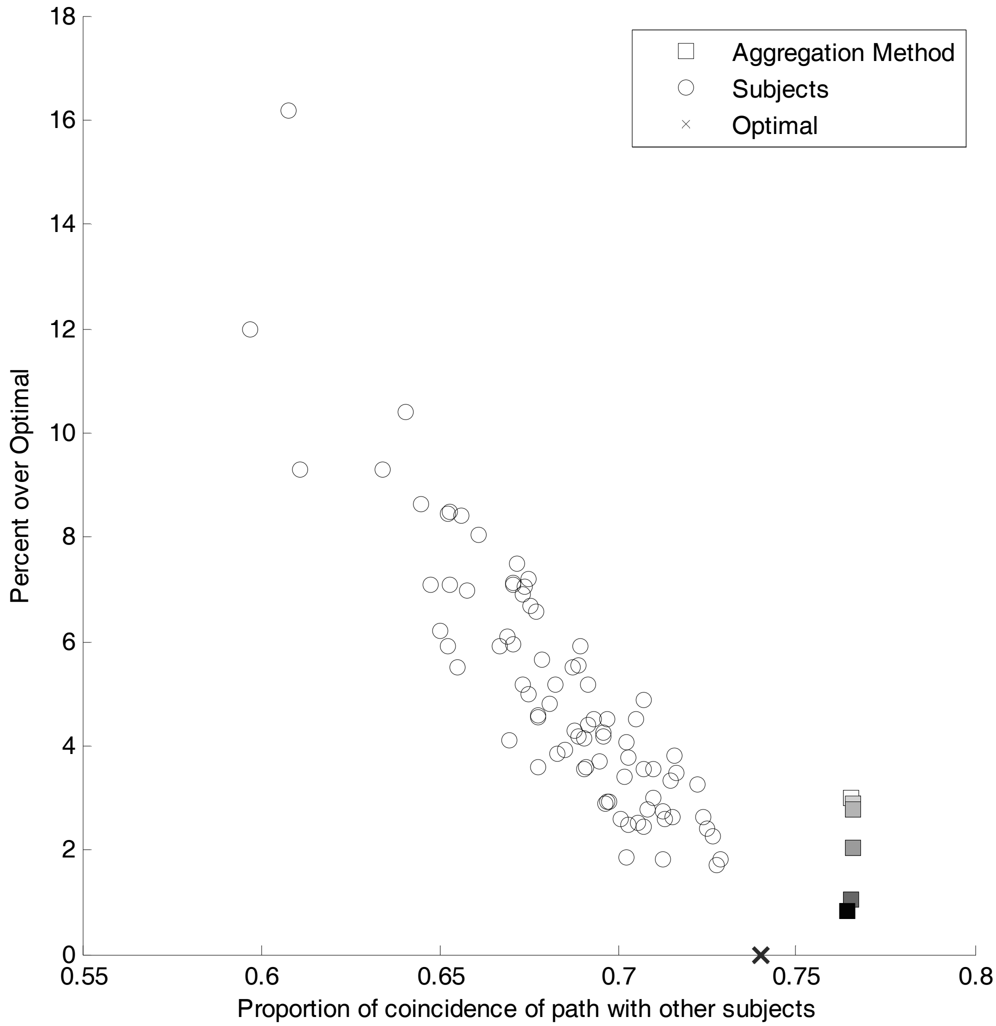
a) b)



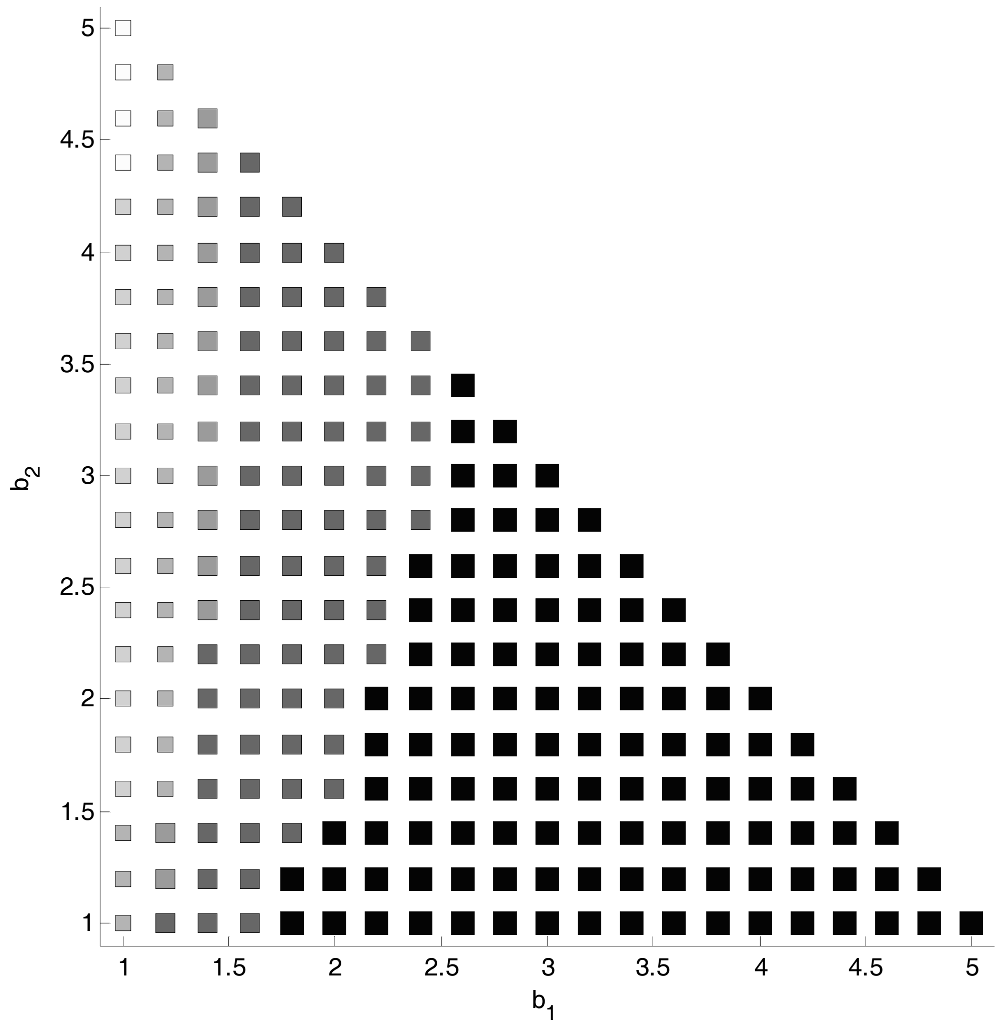
a)

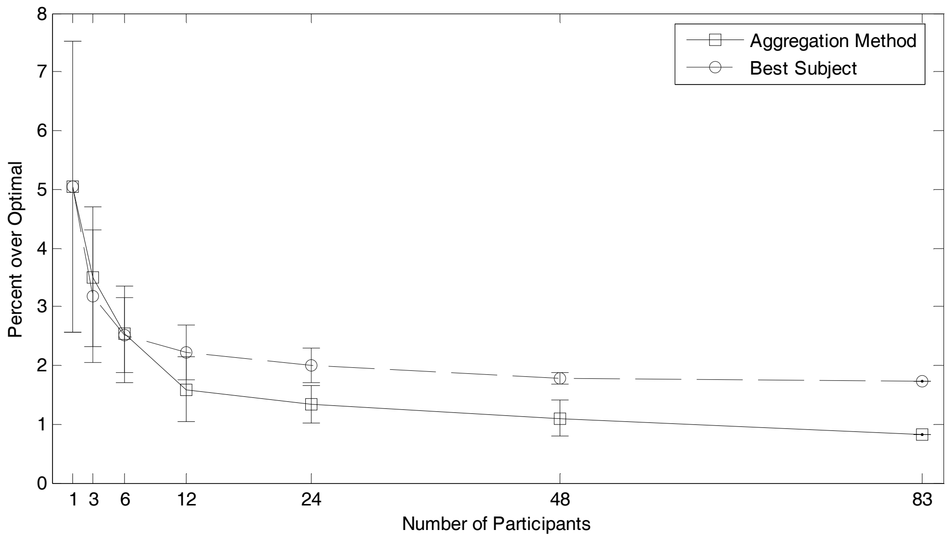


b)



18
16
14
12
10
8
6
4
2
0
Percent over Optimal
0.55
0.6
0.65
0.7
0.75
0.8
Proportion of coincidence of path with other subjects
Aggregation Method
Subjects
Optimal


5
4.5
4
3.5
3
2.5
2
1.5
1
b_2
1
1.5
2
2.5
3
3.5
4
4.5
5
b_1


Aggregation Method
Best Subject
Percent over Optimal
Number of Participants
0
1
2
3
4
5
6
7
8
1
3
6
12
24
48
83