Bounded Manipulation of Median Mechanisms in d Dimensions

Ankit Patel and Julie Farago

May 17, 2002

1 Abstract

In this paper we introduce the idea of *bounded strategy proofness.* We than apply this idea to the problem of allocating goods to agents with linear single peaked preferences in two dimensional space. In our paper we find that we can bound the amount of agent manipulation possible while still arriving at an efficient outcome. We introduce two mechanisms to determine a median, which will become the allocation, in two dimensional space. We find that the first and more simple algorithm introduced results in close approximation to the more complex mechanism. Additionally, we prove bounds of expected manipulation within the simple mechanism to the $O(\frac{1}{n})$. Finally, we introduce many possible path for future work with regards to the ideas laid out in the paper.

2 Introduction

We introduce the concept of bounded manipulation in which an agent's ability to manipulate the outcome of a mechanism, given that all other agents are truthful, is bounded (with high probability). This concept deviates from the usual strategy proof mechanisms in which the agent has no incentive whatsoever to lie. In a mechanism that has this bounded manipulation property, there is some incentive to lie, but this incentive is bounded since the amount of benefit the agent can obtain is bounded. In particular, we find efficient, bounded manipulation mechanisms for certain voting problems in $\mathbb{R}^2$ and we conjecture that these also extend to problems in $\mathbb{R}^d$.

3 Related Work

When determining a fair allocation for linear single peaked preferences, we know that an efficient and strategy-proof solution was to take the median of the peaks [Moulin]. We intend to extend this finding to n dimensional space. We know however that it is not possible to create a strategy-proof, efficient, non-dictatorial, social choice function in higher dimensions [Barbera and Jackson], however we would like to explore a class of social choice functions that lead to bounded manipulation. In order to achieve this goal, we prove bounds on agent manipulation and then test them via simulation.

4 Motivation

The application for a two dimensional social choice function is especially obvious in geographical elections, where the space of possibilities is 2-dimensional. More generally, it is applicable any space where there is a spectrum (e.g. the political spectrum from liberal to conservative). If we can find a way to bound potential agent manipulation, we can have a reasonable method of allocation in such spaces.

5 Bounded Manipulation

Definition. (Bounded Manipulation) A mechanism is ϵ-bounded in agent manipulation if, ex-post, an agent can only change the outcome of the mechanism by at most ϵ given that all other agents are truthful.

6 Median Mechanisms

Before we describe the different mechanism that we use, we will lay down the terminology needed to understand them. A projection of a set S onto a line l is $proj_l S$. Let S be the set of sample points.

6.1 A Simple Median Mechanism

Our initial approach to finding a median in two dimensional space is centered around a fairly simple mechanism. The algorithm for determining the median among a set of discrete points is:

Step 1. Project S onto the x-axis and find the median. Call this $Median_x$.

Step 2. Project S onto the y-axis and find the median. Call this $Median_y$.

Step 3. The *Ideal Median* is the point $(Median_x, Median_y)$

Step 4. Find the point in S closest to the *Ideal Median* and call this point the *Median.*

6.1.1 Bounds for Uniform Distribution

Theorem(Bounded Manipulation). Given that agents values are uniformly distributed in the unit square, an agent can manipulate the mechanism's outcome by at most $O(\frac{1}{n})$ with high probability, given that all other agents are truthful.

Proof. Consider an agent i, after the mechanism has chosen an outcome. Does agent i have any incentive to have reported something other than his true value v_i? Yes. WLOG, suppose agent i's true value v_i is in the first quadrant, where the origin is defined to be the ideal median m_0 of S. Also, suppose that the outcome of the mechanism is denoted by m. By a simple geometric argument, one can see that the very best that i can do is to "sneak" right inside the circle centered at m_0 that contains m on its circumference. If i were to report this point right inside the above circle as its value, then the new outcome m' will be as close to v_i as is possible. Hence, the amount of manipulation that i can do is equal to the distance from m_0 to the point in S closest to m_0. But this distance, denoted by R_{min}, is a random variable, since the values of the agents is a probability distribution. We want to show that the expectation of R_{min} is $O(\frac{1}{n})$ where n is the number of agents.

Define X_i, Y_i as random variables uniformly distributed on [-1,1]. Then the distance from the median [(0,0) in the case of a uniform distribution centered at (0,0)] is $R_i = \sqrt{X_i^2 + Y_i^2} \leq |X_i| + |Y_i|$ by the triangle inequality. Hence, we can get an upper bound by considering $R_i = (X_i + Y_i)$ instead, where X_i, Y_i are uniformly distributed on [0,1]. So, by convolving the uniform density function $f(z) = 1, 0 \leq z \leq 1$ with itself, we get that $f_{R_i}(r) = \int f(z)f(r-z)dz = 1 - |r-1|, 0 \leq r \leq 2$. Now, define $R_{min} = min\{R_1, ..., R_M\}$ and so the cumulative density of R_{min} is $F_{R_{min}}(r) = Pr[R_{min} \leq r] = Pr[\forall i\, R_i \geq r] = \Pi_i Pr[R_i \geq r] = \Pi_i(1 - F_{R_i}(r))$. After some integration and simplification, one gets:

$$Pr[R_{min} \geq r] = 1 - F_{R_{min}}(r) = \begin{cases} (1 - \frac{r^2}{2})^n & 0 \leq r \leq 1 \\ (2 - 2r + \frac{r^2}{2})^n & 1 \leq r \leq 2 \end{cases}$$

Of course, $E[R_{min}] = \int rPr[R_{min} \geq r]dr$ however, this cannot be evaluated easily because of the exponent n. Nonetheless, via Maple we discovered that the expectation is indeed $O(\frac{1}{n})$, specifically it is bounded above by $\frac{\sqrt{2}}{n}$. Is this intuitive? Well, yes it is since we know that in one dimension the expectation of the minimum of n uniformly random samples on [0,1] is $\frac{1}{n+1} = O(\frac{1}{n})$ and by the Pythagorean Theorem it should be $O(\frac{\sqrt{2}}{n})$ in two dimensions. It is by this logic and some more Maple tests that we conclude that in general, in d dimensions, this expected minimum is $O(\frac{\sqrt{d}}{n})$, though a rigorous proof seems daunting. Note that from the density function $f_{R_{min}}(r) = \frac{\partial}{\partial r}F_{R_{min}}(r)$ we can calculate confidence intervals, i.e. values $r_{p,n}$ s.t. $Pr[R_{min} \leq r_{p,n}] \geq p$, where n is the number of agents participating in the mechanism. ■

Do we know whether this mechanism yields an efficient outcome? The answer turns out to be yes, in the case of linear single-peaked preferences, and we can prove it in a straightforward manner. The more general case of single-peaked preferences is also conjectured to be true but we only sketch a proof of this proposition here.

Theorem(Efficiency). The above mechanism is efficient and maximizes the total utility of n agents with single-peaked linear preferences in the plane.

Proof. Let $\ell_i(m_1, m_2) = |(x_i, y_i) - (m_1, m_2)| = |x_i - m_1| + |y_i - m_2|$ be the loss function of agent i, where $m = (m_1, m_2)$ is the outcome of the mechanism, and $I = (I_1, I_2)$ is the ideal median of S. Then the total loss is:

$$L(S, m) = \Sigma_i \ell_i(m) = \Sigma_i |x_i - m_1| + |y_i - m_2|$$

By differentiating with respect to m_i, we find that $\frac{\partial}{\partial m_i}L = \#\{i : z_i > m_i\} - \#\{i : z_i < m_i\} = 0$, where $i = 1, 2$ and $z_1 = x, z_2 = y$. Thus, the minimizer of total loss is the outcome $(m_1, m_2) = (Median_x, Median_y) = I$, the ideal median. However, since the ideal median may not be in the data set, we must instead choose some other point. Which one? Well, by Triangle Inequality:

$$\begin{aligned} L(S,m) &= \Sigma_i \ell_i(m) = \Sigma_i |x_i - m_1| + |y_i - m_2| \\ &= \Sigma_i |x_i - I_1 + I_1 - m_1| + |y_i - I_2 + I_2 - m_2| \\ &\leq \Sigma_i(|x_i - I_1| + |y_i - I_2|) + \Sigma_i(|I_1 - m_1| + |I_2 - m_2|) \\ &= L(S,I) + n(|I_1 - m_1| + |I_2 - m_2|) \end{aligned}$$

Now, since $L(S,I)$ is already minimized, the best we can do is minimize the second term, which is just the "Manhattan distance" between the ideal median I and m. Thus minimization is equivalent to choosing the point m in S s.t. the Manhattan distance $(|I_1 - m_1| + |I_2 - m_2|)$ is minimum. But this is exactly what our median mechanism does! Note that this proof easily generalizes to d dimensions. ■

Proof Sketch. (General single-peaked preferences) A utility function $u_i(m_1, m_2)$ that is single-peaked must also be single-peaked in each variable. Since the choice of median is efficient in one dimension [Moulin], we can conclude that the choice of median in each dimension is also efficient since we can treat the x_i independently of the y_i. Thus, the choice of the ideal median I is efficient. The problem, however, is that I may not be in S. From here, we must prove that the choice of the point in S that is "closest" to I is the best we can do. Though we do not provide this proof here, we believe it results from a straightforward generalization of Moulin's argument for the one dimensional case. [Moulin] ■

6.2 A Generalized Mechanism

We extend our initial approach by allowing medians to be taken on M non-parallel lines in the plane. The median will be defined as the point in S that is the median on the most number of non-parallel lines. The new algorithm for determining the generalized median of a set of discrete points is as follows:

Step 1. Choose M non-parallel lines, $L_1, ..., L_M$, s.t. L_i has slope $-\frac{\pi}{2} + \frac{i\pi}{M}$.

Step 2. $\forall i$ Project S onto L_i, and compute the median. Call this median m_i.

Step 3. Find the mode of $m_1, m_2, ..., m_M$. Call this point the *Median.*

This mechanism is a generalization of the previous, simpler mechanism.

Proving Bounded Manipulation and Efficiency is much more difficult and we do not yet know hot to fully do it. We can, however, give intuition as to why we believe these conjectures are true. When an agent looks back on his decision, he will not be able to easily compute an alternative solution with more benefit to him. The computation seems very daunting and complex, since there are M lines, where M is very large. An agent would have to "move" to a point where he would simultaneously become the median for many lines that he was not a median for before, enough lines s.t. he would become the new mode or that soem other point closer to him becomes the new mode. Finding such a point seems difficult and the set of all such points seems very small (we conjecture exponentially small in M). Thus, this computation seems very hard indeed and therefore an agent has even more incentive to be truthful.

As for efficiency, since we have only a 2-dimensional space, the medians of the M lines in the plane are "dependent" upon the medians on the x- and y-axes. Since the latter case is efficient by our earlier Thm in a previous section (proven for linear single-peaked utilities, sketched for the general case), we conjecture that it is also true for the generalized mechanism. Another piece of intuition that points toward efficiency is that as $M \rightarrow \infty$, the mode of all medians should be the most "central" point in S since it is defined as the point that is the "center" along almost all lines in the plane. This notion of centrality brings this generalized median closer to is ideal median, which we know is efficient from earlier.

Furthermore, we found, through simulation (code attached) that the generalized mechanism for a large M seems to largely coincide with the simple mechanism. The resulting median from the generalized mechanism (approximately 80% of the time) exactly matched the median from the simple mechanism (which again, is efficient), and the in the rest of the trials the different median results were very close to each other. This suggests that the simple mechanism is a reasonable approximation to the generalized mechanism.

7 Conclusions

We succeeded in defining *Bounded Manipulation* and applied it to the problem of finding a median in two dimensional space. Our simple mechanism was proved to have an expected bounded strategy proofness of $\frac{1}{n}$. Additionally, we proved that the allocation found by this method is efficient. Though

Barbera and Jackson prove that we cannot find a non-dictatorial, efficient and strategy-proof social choice function, we did show a heuristic for approximating strategy-proofness and efficiency. We believe that our simple mechanism is a good way to tackle an otherwise difficult problem.

The difficulty comes from complex mechanisms like the generalized mechanism. We determined that the heuristic from the simple mechanism approximates this more general one. But, this approximation in location of the point chosen does not give us any information about the bounds of strategy-proofness or possible agent manipulation. Though they may choose the same answer, one may be more prone to manipulation than the other.

The most important results from this paper include the newly introduced idea of a problem being strategy proof within certain bounds, and our simple mechanism design for allocating to agents with single peaked preferences in two dimensional space. Additionally, we believe that these ideas can easily be extended to fit other situations. Bounded strategy proofness seems to be a very widely applicable idea, and our 2 dimensional mechanisms can easily be extended to d dimensions.

8 Future Work

The potential for future work on this topic is very large. To begin, we will explore the possibilities within the median problem.

There is a lot of room for exploring how non-uniform distributions would affect the outcome and the bounds of strategy-proofness. For example, if all agents were distributed around the circumference of a circle, then the bounds of possible manipulation would be the diameter of that circle. How do other distributions effect our bounds? Additionally, we did not include a rigorous proof of bounded strategy proofness in d dimensions of the simple mechanism and would like to see the intuition provided above proven. Finally, we showed that our simple allocation was efficient in linear, single peaked preferences, but we believe this can easily be expanded to general single peaked preferences, via Moulin's argument. As for the generalized mechanism, we have neither proven bounded manipulation nor efficiency. We conjecture that these propositions are true and also give intuition for why.

Beyond the constraints of this particular problem there is a lot of possibility for applying the ideas above to other problems. What types of problems are suited well for bounded strategy proofness? How can one coerce a non-strategy-proof problem without bounds into a bounded one? Can we apply this idea to any current mechanisms?