Proceedings of the Twenty-Fifth AAAI Conference on Artificial Intelligence

Efficiency and Privacy Tradeoffs in Mechanism Design

Xin Sui
University of Toronto
Department of Computer Science
xsui@cs.toronto.edu

Craig Boutilier
University of Toronto
Department of Computer Science
cebly@cs.toronto.edu

Abstract

A key problem in mechanism design is the construction of protocols that reach socially efficient decisions with minimal information revelation. This can reduce agent communication, and further, potentially increase *privacy* in the sense that agents reveal no more private information than is needed to determine an optimal outcome. This is not always possible: previous work has explored the tradeoff between communication cost and efficiency, and more recently, communication and privacy. We explore a third dimension: the tradeoff between privacy and efficiency. By sacrificing efficiency, we can improve the privacy of a variety of existing mechanisms. We analyze these tradeoffs in both second-price auctions and facility location problems (introducing new incremental mechanisms for facility location along the way). Our results show that sacrifices in efficiency can provide gains in privacy (and communication), in both the average and worst case.

1 Introduction

Mechanism design deals with the creation of protocols that reach socially desirable (e.g., efficient) outcomes when self-interested agents have private information—typically *utilities over the outcome space*—relevant to the choice of an outcome (Mas-Colell et al. 1995). Much research assumes *direct revelation*, in which agents reveal their full utility functions to the mechanism. But direct mechanisms often elicit more information than required to make optimal choices, leading to communication and computational difficulties. Incremental mechanisms, commonly used in auctions, alleviate communication complexity in some cases (Zinkevich et al. 2003), though worst case results show that nearly complete information is needed in many settings (Nisan and Segal 2006). Alternatively, one can use *approximation*, sacrificing outcome efficiency to reduce communication complexity (Blumrosen and Nisan 2002; Hyafil and Boutilier 2007).

Direct revelation also requires a sacrifice of *privacy*: revealing its full utility function may be undesirable for an agent, especially when some of that information is *provably unnecessary* for computing the optimal outcome. Recent work—using techniques similar to those used in the analysis of communication complexity (Kushilevitz and Nisan 1997)—has analyzed privacy preservation of specific mechanisms in this sense; that is, where the degree to which an agent reveals more than is strictly needed to compute the outcome of the mechanism is the degree to which privacy has been lost.[1] For instance, Brandt and Sandholm(2008) showed that, for second-price auctions, the English auction preserves complete privacy—no agent reveals any more than is strictly necessary to determine the outcome—but that this comes at the cost of exponential communication complexity. More recently, Feigenbaum et al. (2010) proposed a general framework to analyze the *tradeoff between privacy and communication*, defining several forms of *privacy approximation*. They also showed how different mechanisms for second-price auctions (and several other problems) improve privacy at the expense of communication, and vice versa.

Previous work has addressed both the tradeoff between communication and efficiency, and the tradeoff between privacy and communication. In this work, we address a third tradeoff, that between efficiency and privacy, and provide a general framework for analyzing this tradeoff. Specifically, we consider *approximate mechanisms* that find ε-optimal solutions to choice problems, and show how agents' privacy improves as one increases the degree of approximation ε. Our contributions are as follows: In Sec. 3 we define a general framework for analyzing these tradeoffs, extending *privacy approximation ratios*, introduced by Feigenbaum et al.(2010), to the case of approximate mechanisms. In Sec. 4, we analyze the efficiency-privacy tradeoff in approximate versions of mechanisms for second-price auctions, including the English, sealed-bid, and bisection protocols in both the worst and average cases, and compare our ε-privacy approximation ratios with the exact ratios derived by Feigenbaum et al.(2010). We also generalize their analysis from 2-agent to n-agent auctions. In Sec. 5, we develop incremental protocols for *facility location problems* that imple-

[1] Note that the notion of *privacy* used here is quite different from *differential privacy*, which deals with the potential "leakage" of a user's private information associated with a particular set of queries to a database (Dwork 2006). Though some connections between differential privacy and mechanism design have been developed (McSherry and Talwar 2007), these have focused largely on how to exploit differential privacy to design approximately efficient and truthful mechanisms, and do not attempt to limit information revelation in the sense we pursue here.

ment the classic *median mechanism* (Schummer and Vohra 2007). We analyze the exact privacy approximation ratio for these new protocols, and again derive results demonstrating the efficiency-privacy tradeoffs induced by approximate versions of these protocols.

Approximate mechanisms will not just improve (increase) privacy, but also generally improve (reduce) communication complexity. While we have derived communication complexity results for our mechanisms, space precludes a full discussion. Furthermore, sacrificing efficiency usually breaks the incentive properties of standard mechanisms. We again defer a detailed analysis, but make brief remarks as appropriate. We simply note that all of our mechanisms are exactly truthful or ε-incentive compatible.[2]

2 Preliminaries

We assume an outcome must be chosen from a set X for a set of n agents. Each agent has a *type* or *valuation function* $t_i \in T_i$, with $t_i(x)$ denoting its utility for $x \in X$. Let $T = \times_i T_i$ be joint type space. We take each T_i to be finite and for convenience, let $T_i = \{1, \ldots, \nu_i\}$. The outcome is chosen to maximize some *social choice function (SCF)* $f : T \to X$. We focus on *social welfare* $sw(x,t) = \sum_i t_i(x)$, with the aim of choosing $x^* = \text{argmax}\, sw(t,x)$. Since t_i is *private information*, the goal in *mechanism design* is to construct protocols that incentivize each i to reveal enough about t_i to allow an optimal choice to be made (Mas-Colell et al. 1995). We will describe examples of such mechanisms for second-price auctions (SPAs) and facility location problems (FLPs) below.

Much of mechanism design deals with *direct revelation mechanisms*, in which each agent reveals its entire type to the mechanism. For simple outcome spaces (e.g., single-item auctions), the precision required by direct revelation is often unnecessary; in complex settings (e.g., combinatorial auctions, or CAs), X has exponential size, imposing significant burdens on communication. Incremental mechanisms have been proposed (e.g., ascending auctions (Parkes 1999) and adaptive elicitation (Zinkevich et al. 2003) for CAs) which, by eliciting only information that is "needed," can reduce this burden in practice, though not in general (Nisan and Segal 2006). In a different vein, one can use *informational approximation*, eliciting information about agent valuations that admits only an approximately optimal choice. For example, increasing the bid increment in an English auction reduces communication, but loses efficiency by inducing more "ties" due to additional imprecision introduced. This strategy has been examined for single-item allocation (Blumrosen and Nisan 2002), and for general quasi-linear, VCG-style mechanisms (Hyafil and Boutilier 2007).

The *communication complexity* model for multi-party computation (Kushilevitz and Nisan 1997) provides a useful framework for analyzing the costs of specific protocols, but can be adapted to quantify the degree of privacy revelation in mechanisms. One can think of an SCF f as a n-dimensional matrix (tensor) M^f whose entry at position $t = (t_1, \cdots, t_n)$ is $f(t) = f(t_1, \cdots, t_n)$.

Defn. 1. *Let f be an SCF. The* ideal monochromatic region *for $t \in T$ w.r.t. f is $R_f^I(t) = \{t' | f(t') = f(t)\}$. The* ideal monochromatic partition *of f is the set of (disjoint) ideal monochromatic regions w.r.t. T.*

Intuitively, $R_f^I(t)$ describes the set of type profiles t' that are indistinguishable from t relative to f: each such t' leads to the same choice $x = f(t)$. Thus the identity of the ideal monochromatic region is both necessary and sufficient to determine f's choice. A *(deterministic) communication protocol* p specifies the rules by which agents with private information share that information (with a third party or one another) to compute the outcome of a function (Kushilevitz and Nisan 1997). If the outcome $p(t)$ on inputs (types) t satisfies $p(t) = f(t)$, we say p *implements* SCF f. As such, a mechanism is simply a protocol. Define a *rectangle* of $M(f)$ to be a submatrix of $M(f)$.

Defn. 2. *Let p_f implement f. The p_f-*induced rectangle for *$t \in T$, $R_f^p(t)$, is the maximal submatrix S of M containing t s.t. the run of p_f is indistinguishable for any $t' \in S$.*[3]

The p-induced rectangles correspond to the information revealed by p. If a protocol p_f implements f, then we must have $R_f^p(t) \subseteq R_f^I(t)$. Feigenbaum et al.(2010) use the ratio of the sizes of the ideal (maximal) regions of f and the regions (rectangles) induced by p_f to characterize the degree to which p_f discloses *extraneous* private information.[4] We present the definitions using two agents with type vector $t = (t_1, t_2)$ (as in Feigenbaum et al.2010), though they extend to n agents in the obvious way (see below):

Defn. 3. *(Feigenbaum et al. 2010) The* worst case privacy approximation ratio (WPAR) *of protocol p_f for SCF f is:*

$$\mathit{wpar}(p_f) = \max_{(t_1,t_2)\in T} \frac{|R_f^I((t_1,t_2))|}{|R_f^p((t_1,t_2))|}.$$

Let D be a distribution over T. The average privacy approximation ratio (APAR) *of p_f is:*

$$\mathit{apar}(p_f) = E_D\left[\frac{|R_f^I((t_1,t_2))|}{|R_f^p((t_1,t_2))|}\right].$$

We can think of perfect privacy as revealing *only enough* about the type profile of the agents to compute f (i.e., reveal only the ideal region). These ratios (PARs) then measures how much *additional* information a protocol p_f reveals about the type vector (in the worst case, or on average given some distribution over types). A smaller PAR indicates that p offers a greater degree of privacy, with the smallest PAR value of 1 meaning that p offers *perfect privacy*. A PAR

[2] Indeed, when one factors in incentives, there is a more complex four-way tradeoff between efficiency, privacy, communication complexity and incentives. We discuss these issues and communication complexity results in a longer version this paper.

[3] The fact that indistinguishable regions of p_f must be rectangles is a consequence of the communication model (Kushilevitz and Nisan 1997) (e.g., see Fig. 1).

[4] These definitions are *objective privacy approximation ratios*; subjective variants can be defined (Feigenbaum et al. 2010), but we do not use these here.

value of $k > 1$ means that (either in the worst case or on average) the protocol learns that the joint type lies in a region that is k times smaller than required to compute f.

Sandholm and Brandt (2008) show that for SPAs, the English protocol is the only perfect privacy preserving protocol for two bidders, though it bears exponential communication cost; furthermore, perfect privacy is not possible for $n > 2$ bidders. Feigenbaum et al.(2010) demonstrate interesting *tradeoffs* between privacy and communication complexity in two-bidder SPAs by analyzing sealed-bid, bisection, and bounded bisection protocols. We discuss these results below when defining approximate versions of these protocols.

3 Efficiency-Privacy Tradeoffs

The work described above studies the tradeoff between privacy and communication. There has also been research analyzing the tradeoff between efficiency and communication. For example, *priority games* (Blumrosen and Nisan 2002) model single-item auctions in which agents express their valuations with limited precision, and provide allocations (and prices) that sacrifice efficiency (since true types are unknown) for communication savings; they are also strategyproof. *Partial revelation VCG mechanisms* (Hyafil and Boutilier 2007) apply in any setting (social welfare, quasilinear) where VCG can be used, but again limit revelation and sacrifice efficiency. Without efficient outcome selection, such mechanisms are not strategyproof; but with approximate variants of VCG pricing, ε-efficiency induces ε-incentive compatibility in dominant strategies.

Apart from those above, a third natural tradeoff suggests itself, namely, that between efficiency and privacy. We exploit the notion of approximate solution (Blumrosen and Nisan 2002; Hyafil and Boutilier 2007) and show how it can be used to improve the privacy approximation ratios of Feigenbaum et al.(2010): that is, how much additional privacy can be preserved if we allow an ε sacrifice in efficiency. We first define ε-approximation and ε-implementation:

Defn. 4. *We say an SCF $\tilde{f}$* ε-approximates *an SCF f if $|sw(f(t),t) - sw(\tilde{f}(t),t)| \leq \varepsilon$, $\forall t \in T$. If protocol $p_{\tilde{f}}$ implements such an $\tilde{f}$, we say $p_{\tilde{f}}$* ε-implements f.

In other words, $\tilde{f}$ (and any corresponding protocol $p_{\tilde{f}}$) approximates f if the difference in the social welfare between the two is no more than ε for any type profile.

We can now introduce privacy approximation ratios relative to approximate implementations of a SCF f.

Defn. 5. *Let $p_{\tilde{f}}$ be a protocol that ε-implements f with SCF $\tilde{f}$. The* ε-worst case privacy approximation ratio *of $p_{\tilde{f}}$ is:*

$$\varepsilon\text{-}\mathsf{wpar}(p_{\tilde{f}}) = \max_{t \in T} \frac{|R^I_f(t)|}{|R^p_{\tilde{f}}(t)|}.$$

Let D be a distribution over T. The ε-average case privacy approximation ratio *of $p_{\tilde{f}}$ is:*

$$\varepsilon\text{-}\mathsf{apar}(p_{\tilde{f}}) = E_D\left[\frac{|R^I_f(t)|}{|R^p_{\tilde{f}}(t)|}\right].$$

These definitions are similar to those in Defn. 3 except that we compare the ideal monochromatic regions of an SCF f to the regions (or rectangles) induced by a protocol for its *ε-approximation* $\tilde{f}$. Our definitions in fact reduce to Defn. 3 when $\varepsilon = 0$ (i.e., when $\tilde{f} = f$). As above, smaller values of ε-par indicate a greater degree of privacy preservation. Unlike exact par which has a minimum value of 1 (perfect privacy), ε-par *can be less than 1*, indicating that strictly less information than required for computing f is revealed. Indeed, this is only possible because of approximation. While both measures are interesting, we believe the average case measure ε-apar (using appropriate distributions in specific applications) may be more useful in practice.

These definitions can be recast to minimize ε-par over *all* ε-implementations of f, measuring the tradeoffs *inherent* in f; but we focus here on the analysis of specific families of protocols. Mechanisms for specific problems, e.g., SPAs, can be parameterized by the degree of approximation ε they offer, especially by limiting the precision with which agents reveal their valuations, hence improving ε-par by sacrificing efficiency. We now explore this tradeoff.

4 Tradeoffs in Second-Price Auctions

We illustrate the usefulness of our framework by analyzing the efficiency-privacy tradeoffs for approximate versions of three mechanisms used in *second-price auctions (SPAs)*, the English auction, the sealed-bid auction, and the bisection auction. Our contributions are two-fold: first, we generalize the two-agent analysis of Feigenbaum et al.(2010) by providing privacy approximation ratios (or bounds) for n-agent SPAs (whose analysis is somewhat more involved). Second, we demonstrate the additional privacy savings obtained by admitting approximate efficiency.[5]

Consider a setting with n agents, and each agent $i \leq n$ has a valuation v_i for some item. Let $v[h]$ be the h-th highest valuation in (multiset) $V = \{v_1, \cdots, v_n\}$ and $a[h]$ the agent with valuation $v[h]$ (ties broken lexicographically). The SPA allocates the item to $a[1]$ for price $v[2]$. The *sealed-bid mechanism* is a one-shot protocol for SPAs: each agent submits its valuation to the mechanism, which awards the item as required. The *English auction* is incremental: a (small) bid increment δ is chosen, and the price p is raised by δ at each round; i can drop out in any round (strategically, when $p > v_i$); when one agent remains, it is awarded the item at the current price. Ties (i.e., when more than one agent drops out at the last round) are broken lexicographically (at the prior price, which all final agents "accepted"). The *bisection auction* (Grigorieva et al. 2007) uses a binary search (asking each i whether v_i is above specific values) to determine a value b that lies between $v[1]$ and $v[2]$. Once $v_i < v[2]$ is proven, no further queries are asked of i. Once b is identified, binary search on the interval containing $v[2]$ is used to identify $v[2]$ to a desired precision σ. Following Feigenbaum et al.(2010), we treat the valuation space as discrete, representable with k bits, allowing $\nu = 2^k$ distinct valuations. We assume, w.l.o.g., that $v_i \in \mathbf{V}^k = \{0, \ldots, 2^k - 1\}$.

[5] We omit analysis of the bounded bisection protocol (Feigenbaum et al. 2010) for space reasons.

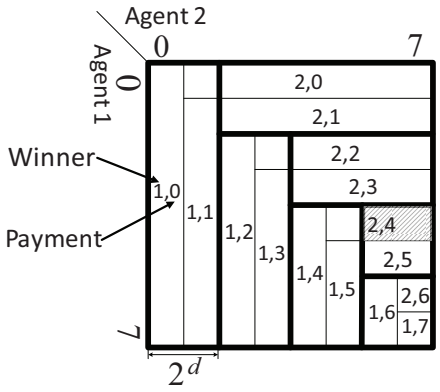

Fig. 1: Partitions induced by the English auction for 2-bidder SPAs when $\delta = 1$ ($\varepsilon = 0$, thin line) and $\delta = 2$ ($\varepsilon = 1$, thick line). When $\delta = 1$, this is also the ideal monochromatic partition. The shaded region indicates the inputs from which ε-wpar is derived. The numbers indicate the outcome for each ideal rectangle (e.g., in the leftmost rectangle, the item is allocated to agent 1 for a price of 0).

English Protocol The English protocol with an "exact" bid increment $\delta = 1$ has exponential communication complexity $O(2^k)$ (Brandt and Sandholm 2008): simply consider the case of $v[2] = 2^k - 1$. But this high cost allows for very strong privacy: for two agents, par (both worst and average case) is 1, i.e., it is perfectly privacy preserving (though for $n > 2$ agents, perfect privacy is not possible (Brandt and Sandholm 2008)). The thin line in Fig. 1 illustrates the ideal monochromatic partition for a two-agent SPA.

We can approximate the English auction by simply increasing the bid increment, setting $\delta = 1 + \varepsilon = 2^d$ for some precision $d > 0$.[6] Clearly this *ε-English protocol*, denoted p_E^ε, ε-approximates SPA, with suboptimal allocation happening only when multiple agents drop out at the last round; but all such agents have values within an interval of size $(1+\varepsilon)$, guaranteeing ε-efficiency. The price paid is also within ε of that dictated by the SPA, and p_E^ε is incentive compatible. The thick line in Fig. 1 illustrates the protocol-induced partition for the ε-English auction when $\varepsilon = 1$. Notice that for some type profiles, the outcome is different from that of the exact protocol (e.g., with profile $(2, 3)$, p_E^ε allocates the item to agent 1 for a price of 2, while the exact protocol allocates efficiently to agent 2 for a price of 2). It is easy to verify that, for any t, the protocol induced rectangle for p_E^ε is at least as large as that induced by the exact English protocol, indicating privacy savings. The shaded area denotes the profiles from which we derive ε-wpar: the ideal monochromatic region has size 3 while the protocol-induced rectangle has size 4. Note that ε-wpar $= \frac{3}{4} < 1$, indicating *better than perfect privacy*.

These intuitions can be generalized to n-agent SPAs:

Theorem 6. *For n-agent SPAs,*

$$\varepsilon\text{-wpar}_p(p_E^\varepsilon) = \frac{(2^k)^{n-1} - (2^k - 1)^{n-1}}{(1+\varepsilon)^n}$$

Worst-case privacy savings of p_E^ε relative to the exact protocol are $(1+\varepsilon)^n$, as one would expect ($1+\varepsilon$ per agent).

[6] We use powers of 2 for convenience only.

Suppose we have a uniform distribution D over type profiles (all average-case analysis in the sequel uses this D). We can bound ε-apar for the ε-English protocol:

Theorem 7. *For n-agent SPAs,*

$$(\lceil \tfrac{n}{2} \rceil)^2 \frac{(2^{k-1})^{n-3}}{\binom{n-3}{\lceil \frac{n}{2} \rceil - 2}(1+\varepsilon)^{n-2}} \leq \varepsilon\text{-apar}(p_E^\varepsilon) \leq 2\binom{n}{2}\frac{(2^k)^{n-2}}{(1+\varepsilon)^{n-1}}$$

In the ε-English protocol, the valuations of at least $n-1$ agents are identified with precision $1+\varepsilon$, so privacy savings are at least $(1+\varepsilon)^{n-1}$ relative to exact implementation. However, Thm. 7 bounds the savings by $(1+\varepsilon)^{n-1}$, so the average privacy savings of p_E^ε are exactly $(1+\varepsilon)^{n-1}$. We compare ε-apar of p_E^ε with that of other protocols below.

Bisection Protocol A natural way to approximate the bisection protocol is to use early termination, stopping when we identify $v[2]$ with some desired precision σ (i.e., when the bisection interval containing $v[2]$ is no larger than σ). We then allocate to $a[1]$ using the price at the low end of $v[2]$'s interval (ties broken lexicographically). To ensure ε-efficiency, thus defining the *ε-bisection protocol* p_B^ε, we must have $\sigma \leq 2^{\lfloor \log_2(\varepsilon+1) \rfloor}$. This mechanism is ε-incentive compatible (an agent's gain by misreporting is at most ε). We can derive bounds for wpar:

Theorem 8. *For n-agent SPAs,*

$$\frac{(n-1)(2^{k-1})^{n-1}}{(2^{\lfloor \log_2(\varepsilon+1) \rfloor})^n} \leq \varepsilon\text{-wpar}(p_B^\varepsilon) \leq \frac{(2^k)^n}{(2^{\lfloor \log_2(\varepsilon+1) \rfloor})^n}$$

This worst case occurs when all agents have values clustered in the interval containing $v[2]$: thus each reports its valuation with the maximum precision, so ε-wpar is exponential in both k and n. The privacy savings of p_B^ε relative to exact implementation is precisely $(2^{\lfloor \log_2(\varepsilon+1) \rfloor})^n \approx (\varepsilon+1)^n$.

For ε-apar, we have the following bounds:

Theorem 9. *For n-agent SPAs,*

$$\frac{nk}{(1+\varepsilon)^n} \leq \varepsilon\text{-apar}(p_B^\varepsilon) \leq (n+1)\binom{n}{\lfloor \frac{n}{2} \rfloor}\frac{(k+1)^n}{1+\varepsilon}$$

We see that apar for (exact and approximate) bisection is polynomial in k (and exponential in n), which compares favorably to the English protocol (which is exponential in both k and n). Depending on the number of agents whose values fall in the bisection interval containing $v[2]$ at termination, the privacy savings for p_B^ε range from $2^{\lfloor \log_2(\varepsilon+1) \rfloor}$ to $(2^{\lfloor \log_2(\varepsilon+1) \rfloor})^n$. We compare exact average case savings of p_B^ε with those of other protocols below.

Sealed-Bid Protocol The *ε-sealed-bid protocol* p_S^ε approximates the usual sealed-bid protocol by simply "coarsening" the valuation space, asking for reports v_i with limited precision σ. The bound $\sigma \leq 2^{\lfloor \log_2(\varepsilon+1) \rfloor}$ also holds for p_S^ε, requiring termination only when $v[2]$ is known to lie within an interval of length $2^{\lfloor \log_2(\varepsilon+1) \rfloor}$. ε-wpar for p_S^ε in n-agent SPAs is identical to that for ε-bisection, since it induces the same rectangles in the worst case.

ε	Second Price Auctions		
	$n=3$	$n=4$	$n=5$
$\varepsilon=0$	32 / 15 / 410	1225 / 72.4 / 11254	46563 / 350 / 333760
$\varepsilon=1$	8.1 / 5.11 / 51.25	156 / 20.7 / 703.4	2942 / 86 / 10430
$\varepsilon=3$	2.05 / 1.56 / 6.4	19 / 5 / 44	173 / 17.3 / 325.9
$\varepsilon=7$	0.48 / 0.4 / 0.8	1.95 / 0.96 / 2.75	8.0 / 2.6 / 10.2

Table 1: ε-apar for SPAs with different n and ε when $k=5$ bits. The three values in each cell indicate ε-apar for the ε-English, ε-bisection and ε-sealed-bid protocols, respectively.

Theorem 10. *For n-agent SPAs,*

$$\frac{(n-1)(2^{k-1})^{n-1}}{(2^{\lfloor \log_2(\varepsilon+1)\rfloor})^n} \leq \varepsilon\text{-wpar}(p_S^\varepsilon) \leq \frac{(2^k)^n}{(2^{\lfloor \log_2(\varepsilon+1)\rfloor})^n}$$

Despite the same worst case behavior, ε-sealed-bid is much worse on average than ε-bisection:

Theorem 11. *For n-agent SPAs,*

$$(\lceil \frac{n}{2} \rceil)^2 \frac{(2^{k-1})^{n-3}}{\binom{n-3}{\lceil \frac{n}{2} \rceil - 2}(1+\varepsilon)^{n-2}} \leq \varepsilon\text{-apar}(p_S^\varepsilon) \leq \frac{(2^k)^n}{(2^{\lfloor \log_2(\varepsilon+1)\rfloor})^n}$$

Our current lower bound for ε-apar(p_S^ε) is quite loose; but we can use ε-apar(p_E^ε) in its place: for each profile, p_E^ε induces a larger rectangle p_S^ε, so ε-apar$(p_S^\varepsilon) \geq \varepsilon$-apar$(p_E^\varepsilon)$. Hence, ε-apar(p_S^ε) is exponential in both k and n. In addition, since the size of all induced rectangles is $(2^{\lfloor \log_2(\varepsilon+1)\rfloor})^n$, in both the worst and average case, p_S^ε offers privacy savings of $(2^{\lfloor \log_2(\varepsilon+1)\rfloor})^n \approx (\varepsilon+1)^n$ over exact sealed-bid.

Summary The average case ε-privacy approximation ratios for SPAs of different sizes, computed numerically, are shown in Table 1. Recalling that smaller ε-apar indicates better privacy, we see that our ε-approximate protocols offer significant privacy savings relative to their exact counterparts. For instance, when $n=3$ and $\varepsilon=1$, the ε-English protocol reveals a fraction $\frac{8.1}{32} \approx \frac{1}{4}$ of the information revealed by the exact protocol, while $\varepsilon=3$ requires only $\frac{1}{16}$ of that information. We also see that ε-bisection has the smallest ε-apar, preserving much more privacy than either ε-English or ε-sealed-bid; e.g., when $\varepsilon=3$ and $n=4$, ε-bisection requires revelation of only $\frac{5}{19}$ and $\frac{5}{44}$ of the information required by ε-English and ε-sealed-bid, respectively. We also notice that ε-apar, and the privacy savings of the approximate protocols over their exact counterparts, grows exponentially with n. This is consistent with our theoretical results. Moreover, though our current proven lower bound for ε-bisection is linear in ε, these numerical results suggest that the actual savings are much greater. We conjecture that the true savings are $O((1+\varepsilon)^{\frac{n}{c}})$, for some constant $c>1$.

To summarize, we have derived privacy approximation ratios for the n-agent versions of three key protocols for SPAs. We have also shown that approximate variants of these protocols allow for savings in privacy over their exact counterparts that is exponential in the number of agents n and polynomial in the degree of approximation ε in almost all cases (both worst and average case).

5 Tradeoffs in Facility Location

We now consider another classic domain in mechanism design, *facility location problems (FLPs)* (Schummer and Vohra 2007). We must locate a facility (e.g., warehouse, or public park) to satisfy the needs of n agents. We have a finite set of locations on the real-line, which we take to be integers $\mathbf{L} = \{0, \ldots, 2^k - 1\}$. Agent i's type indicates its preferred location $q_i \in \mathbf{L}$, and its cost (negative utility) for any location y is $c_i(y) = c(y, q_i) = |y - q_i|$. The objective is to select an optimal location that maximizes social welfare by minimizing the SCF $f(q) = \sum_i c(y, q_i)$, i.e., the sum of distances faced by the agents. FLP has also been the subject of approximate implementation, e.g., to ensure strategyproofness given different social objectives (Procaccia and Tennenholtz 2009).

The *median mechanism* is a strategyproof mechanism that selects the optimal location y for FLPs (Schummer and Vohra 2007). We use FLP to refer to the problem of implementing the median mechanism in incremental and approximate fashion in the sequel. For ease of exposition, we assume an odd number of agents $n = 2m-1$; we also assume (w.l.o.g.) that agents are sorted by preferred location: $q_1 \leq \ldots \leq q_n$. The median mechanism asks each agent i to report q_i and locates the facility at the median q^M of the reported values. Generally FLPs are tackled using sealed-bid-like direct mechanisms; however, incremental elicitation of the q_i can be accomplished using mechanisms much like those for SPAs. We can define an *English protocol for FLPs*: beginning with a *current location* $p=0$, we increment p by $\delta = 1$, asking i if $q_i \geq p$, stopping when at least m agents have dropped out, thereby identifying the median. The *bisection protocol for FLPs* simply conducts a binary search to find the median q^M: at any stage, if we know either that $q_i \geq q^M$ or that $q_i \leq q^M$, agent i is asked no further queries. Notice that we need not know the exact value of q_i or q^M, merely that at least half of the agents have locations less than (respectively, greater than) q_i. Approximate versions of both protocols (as well as "sealed-bid") are defined analogously to the case of SPAs.

Before describing results regarding privacy approximation ratios, we first provide a general negative result:

Theorem 12. *There is no perfect privacy preserving protocol for the median mechanism for n-party FLPs, for any $n \geq 2$.*

Intuitively, this holds because any protocol requires the (indirect) revelation of the identity of an agent with the median value in at least *some* instances.

Claim 13. *Let q be a type profile with median q^M. Then:*

$$R_f^I(q) = \sum_{t=0}^{m-2} \binom{n}{t+1} [\sum_{s=0}^{t} \binom{n-1-t}{m-1-t+s} \cdot (q^M)^{m-1-t+s} \cdot (2^k - 1 - q^M)^{m-1-s}] + \sum_{t=0}^{m-1} \binom{n}{t} (2^k-1)^t$$

This result shows that the size of the ideal monochromatic region for FLPs is a function $Z(q^M)$ of q^M, not the entire profile q. Note $Z(r) = O(r^{m-1}(2^k - r)^{m-1})$, with

its largest value when $r = 2^{k-1}$. The first term reflects when fewer than m agents have location q^M, and the second when at least m agents have location q^M.

English Protocol We first analyze the exact English protocol for FLPs (i.e., where $\delta = 1$).

Theorem 14. *Let p_{EF} be the English protocol for n-agent FLPs. Then* $\mathsf{wpar}(p_{EF}) = Z(2^k - 2)$.

Worst case PAR is obtained when $m - 1$ agents prefer location $2^k - 1$, and m prefer $2^k - 2$: then p_{EF} induces a rectangle of size 1, while the ideal region has size $Z(2^k - 2)$.

The ε-English protocol p^{ε}_{EF} uses a bid increment $\delta > 1$, identifying the median with precision δ when the protocol stops, and randomly selecting a location within this δ-interval. To ensure ε-approximation, δ cannot be too large:

Lemma 15. *p^{ε}_{EF} ε-implements FLP only if the bid increment δ satisfies $\delta \leq 1 + \frac{\varepsilon}{n}$.*

The distinction with SPAs, which allow increments of $(1+\varepsilon)$, is due to the fact that an ε-misplacement of the facility can impact all n agents (not just the winner as in SPAs). The mechanism is $\frac{\varepsilon}{n}$-incentive compatible.

By Thm. 14 and Lem. 15, we have:

Corollary 16. *For n-agent FLPs,*

$$\lim_{n\to\infty} \varepsilon\text{-}\mathsf{wpar}(p^{\varepsilon}_{EF}) = \lim_{n\to\infty} \frac{Z(2^k - 2 - \frac{2\varepsilon}{n})}{(1 + \frac{\varepsilon}{n})^n} = \frac{Z(2^k - 2)}{e^{\varepsilon}}$$

For the ε-English protocol, each agent's location is identified with a precision of $1 + \frac{\varepsilon}{n}$ in the worst case, so the size of the protocol induced rectangle is $(1 + \frac{\varepsilon}{n})^n$ and converges to e^{ε} as $n \to \infty$. However, the $Z(q^M)$ term indicates that ε-wpar is still exponential in both k and n.

We begin our average case analysis with the exact protocol, providing upper and lower bounds:

Theorem 17. *For n-agent FLPs,*

$$m\binom{m-1}{\frac{m}{2}}(2^{k-1})^{m-2} \leq \mathsf{apar}(p_{EF}) \leq m\binom{n}{m-1}(2^k)^{m-1}$$

This result allows us to show that the average case privacy savings of p^{ε}_{EF} relative to exact p_{EF} are at most $(1 + \frac{\varepsilon}{n})^m$. However, in p^{ε}_{EF}, we "coarsen" the revealed locations of at least m and at most n agents, which means that the savings are exactly $(1 + \frac{\varepsilon}{n})^m$, and converges to $e^{\varepsilon/2}$ as $n \to \infty$. These exact savings can be multiplied by the terms in the bounds of Thm. 17 to derive bounds on ε-$\mathsf{apar}(p^{\varepsilon}_{EF})$.

Bisection Protocol We now consider the bisection protocol p_{BF} for FLPs and analyze its privacy approximation ratios before considering its ε-approximate implementation. We first consider wpar for p_{BF}:

Theorem 18. *The bisection protocol for n-agent FLPs has* $\mathsf{wpar}(p_{BF}) = Z(2^{k-1})$.

This worst case occurs with a type profile with $m - 1$ agents having ideal location $2^{k-1} + 1$ and m agents preferring location 2^{k-1}: the rectangle induced by p_{BF} has size 1 while the ideal monochromatic region has size $Z(2^{k-1})$. Hence, wpar is exponential in both k and n.

The approximate *ε-bisection protocol* p^{ε}_{BF} for FLPs identifies the median only to some desired precision, but uses a *dynamic precision* parameter to determine termination. Specifically, we terminate when the median is proven to lie in some interval $[q^M_-, q^M_+)$, and a random point in that interval is selected for the facility. The mechanism is $\frac{\varepsilon}{n}$-incentive compatible. To ensure ε-efficiency, we require:

Lemma 19. *Let l and r be the number of agents in $[q^M_-, q^M_+)$ whose desired location is left of (less than) and right of (greater than) of q^M, respectively. p^{ε}_{BF} ε-implements FLPs iff $(q^M_+ - q^M_- - 1)(2\max\{l, r\} + 1) \leq \varepsilon$.*

This means the "precision" of the final interval $[q^M_-, q^M_+)$ is determined by p^{ε}_{BF} dynamically: if, when the median value interval is identified, no other agents' locations lie within $[q^M_-, q^M_+)$, the protocol can stop when the interval is narrowed to $q^M_+ - q^M_- \leq 2^{\lfloor \log_2(\varepsilon+1) \rfloor} \approx 1 + \varepsilon$; but if $m - 1$ agents remain in the interval, and are left of q^M, then the protocol stops only when $q^M_+ - q^M_- \leq 2^{\lfloor \log_2(1+\frac{\varepsilon}{n}) \rfloor} \approx 1 + \frac{\varepsilon}{n}$. This mechanism is also $\frac{\varepsilon}{n}$-incentive compatible. By Thm. 18 and Lem. 19, we have the following corollary for the ε-bisection protocol:

Corollary 20. *For n-agent FLPs, ε-bisection satisfies:*

$$\lim_{n\to\infty} \varepsilon\text{-}\mathsf{wpar}(p^{\varepsilon}_{BF}) = \lim_{n\to\infty} \frac{Z(2^{k-1})}{(1 + \frac{\varepsilon}{n})^n} = \frac{Z(2^{k-1})}{e^{\varepsilon}}$$

For average case analysis, we again begin with exact bisection, providing upper and lower bounds:

Theorem 21. *For n-agent FLPs,*

$$\binom{n}{m}k^{m-1} \leq \mathsf{apar}(p_{BF}) \leq m\binom{n}{m}(k+1)^n$$

As with SPAs, apar for bisection in FLPs is polynomial in k, offering significant privacy savings relative to the English protocol. With ε-approximation, we can show that privacy savings range from $2^{\lfloor \log_2(1+\varepsilon) \rfloor}$ to $(2^{\lfloor \log_2(1+\frac{\varepsilon}{n}) \rfloor})^n$, depending on the number of agents whose locations fall into the bisection interval as q^M. We compare average case savings across these different protocols below.

Sealed-Bid Protocol The sealed-bid protocol p_{SF} for FLPs has each agent reveal her preferred location and returns the median.

Theorem 22. *For n-agent FLPs,* $\mathsf{wpar}(p_{SF}) = Z(2^{k-1})$.

The ε-sealed-bid protocol p^{ε}_{SF} asks for locations with limited precision σ. In the worst case, when all locations lie in the interval of q^M, Lem. 19 needs precision $\sigma \leq 2^{\lfloor \log_2(1+\frac{\varepsilon}{n}) \rfloor}$, and ε-$\mathsf{wpar}(p^{\varepsilon}_{SF})$ is identical to that for p^{ε}_{BF}.

Corollary 23.

$$\lim_{n\to\infty} \varepsilon\text{-}\mathsf{wpar}(p^{\varepsilon}_{SF}) = \lim_{n\to\infty} \frac{Z(2^{k-1})}{(1 + \frac{\varepsilon}{n})^n} = \frac{Z(2^{k-1})}{e^{\varepsilon}}$$

We have upper and lower bounds on apar for exact p_{SF}:

ε	Facility Location Problems	
	$n=3$	$n=5$
$\varepsilon=0$	96 / 42 / 1228	8776 / 1514 / 1.50E+06
$\varepsilon=10$	6.1 / 3.6 / 19.2	374 / 154 / 46766
$\varepsilon=15$	3.0 / 3.6 / 19.2	152 / 61 / 1461
$\varepsilon=22$	1.34 / 0.97 / 2.4	86 / 36 / 1461

Table 2: ε-apar for FLPs with different n and ε when $k = 5$ bits. The three values in each cell indicate ε-apar for the ε-English, ε-bisection and ε-sealed-bid protocols, respectively.

Theorem 24. *For n-agent FLPs,*

$$\binom{n}{m}(2^k-1)^{m-1} \leq \mathit{apar}(p_{SF}) \leq (2^k)^n$$

In p^{ε}_{SF}, each rectangle has size $2^{\lfloor \log_2(1+\frac{\varepsilon}{n}) \rfloor}$ (compared to size 1 for exact p_{SF}), so average privacy savings are $(2^{\lfloor \log_2(1+\frac{\varepsilon}{n}) \rfloor})^n \approx (1+\frac{\varepsilon}{n})^n$, converging to e^{ε} as $n \to \infty$. However, ε-apar(p^{ε}_{SF}) is still exponential in both k and n.

Summary As with SPAs above, Table 2 shows average case ε-privacy approximation ratios for FLPs of different sizes computed numerically. Results are similar to those for SPAs, so we omit a detailed discussion.

To summarize, we have proposed two incremental mechanisms for FLPs, the English and bisection protocols. Together with the sealed-bid protocol, we have provided upper and lower bounds on worst and average case par, showing, as with SPAs, that the bisection protocol offers relatively strong privacy guarantees compared to the other two protocols (polynomial in k and exponential in n). With ε-approximation, even stronger privacy savings are possible (exponential in ε as $n \to \infty$).

6 Concluding Remarks

We have presented a framework for analyzing the natural tradeoff between efficiency and privacy in mechanism design. Within this model, we have analyzed second-price auctions and facility location problems, and for each investigated the extent to which privacy is preserved for a variety of different protocols. We have shown that the bisection protocol offers significant privacy advantages over other protocols, and also demonstrated the degree to which additional privacy preservation can be gained through ε-approximation of these protocols over their exact implementations, using both worst and average case analyses.

Our framework can be generalized in several ways. While we have presented our work in the context of mechanism design, it can be applied to any form of distributed function computation. One might also consider other forms of approximate privacy that account for, say, different sensitivity to the reports of different agents, or from different regions of type space. Our analysis can also be extended in several ways, including deriving average case results for more realistic distributions of valuations; and broadening the class of mechanisms and social choice functions.

We view this work as simply a first step in the deeper exploration of a complicated four-way tradeoff between communication, efficiency, incentives and privacy in the design of mechanisms. Developing optimization models that explicitly trade off these criteria against one another will be important in the automated design of privacy-preserving mechanisms. Finally, a critical extension will be the analysis of multi-attribute and combinatorial domains, where sophisticated preference elicitation strategies are required (Hyafil and Boutilier 2007). Incremental mechanisms such as those discussed here should have even greater potential to offer practical—if not (worst-case) theoretical—privacy and communications savings. In such domains, there will also be a general need to take advantage of AI techniques for preference assessment.

Acknowledgements Thanks to the reviewers for their helpful comments. This work was supported by the Natural Sciences and Engineering Research Council (NSERC).

References

Blumrosen, L., and Nisan, N. 2002. Auctions with severely bounded communication. *43rd Annual Symposium on Foundations of Computer Science (FOCS-02)*, 406–416.

Brandt, F., and Sandholm, T. 2008. On the existence of unconditionally privacy-preserving auction protocols. *ACM Transactions on Information and System Security* 11(2). Article 6.

Dwork, C. 2006. Differential privacy. *33rd International Colloquium on Automata, Languages and Programming*, 1–12.

Feigenbaum, J., Jaggard, A. D. and Schapira, M. 2010. Approximate privacy: Foundations and quantification (extended abstract). *11th ACM Conference on Electronic Commerce (EC'10)*, 167–178.

Grigorieva, E., Herings, P. J.-J., Müller, R., and Vermeulen, D. 2007. The private value single item bisection auction. *Economic Theory* 30(1):107–118.

Hyafil, N., and Boutilier, C. 2007. Mechanism design with partial revelation. *20th International Joint Conference on Artificial Intelligence (IJCAI-07)*, 1333–1340.

Kushilevitz, E., and Nisan, N. 1997. *Communication Complexity*. Cambridge: Cambridge University Press.

Mas-Colell, A., Whinston, M. D., and Green, J. R. 1995. *Microeconomic Theory*. New York: Oxford University Press.

McSherry, F., and Talwar, K. 2007. Mechanism design via differential privacy. *48th Annual IEEE Symposium on Foundations of Computer Science (FOCS-07)*, 94–103.

Nisan, N., and Segal, I. 2006. The communication requirements of efficient allocations and supporting prices. *Journal of Economic Theory* 129(1):192–224.

Parkes, D. C. 1999. ibundle: An efficient ascending price bundle auction. *1st ACM Conference on Electronic Commerce (EC'99)*, 148–157.

Procaccia, A. D., and Tennenholtz, M. 2009. Approximate mechanism design without money. *10th ACM Conference on Electronic Commerce (EC'09)*, 177–186.

Schummer, J., and Vohra, R. V. 2007. Mechanism design without money. In Nisan, N.; Roughgarden, T.; Tardos, E.; and Vazirani, V. V., eds., *Algorithmic Game Theory*. Cambridge University Press. 209–241.

Zinkevich, M., Blum, A., and Sandholm, T. 2003. On polynomial-time preference elicitation with value queries. *4th ACM Conference on Electronic Commerce (EC-03)*, 176–185.