The Hardness of Being Private

Anil Ada, Arkadev Chattopadhyay, Stephen Cook,
Lila Fontes, Michal Koucký, Toniann Pitassi

IEEE Conference on Computational Complexity 2012
Porto, Portugal

Communication complexity

Two-player model

- each player has a private input (Alice has $x \in X$, Bob has $y \in Y$)
- players communicate over a channel
- players follow a protocol to compute $f : X \times Y \to Z$
- the last message sent is the value $f(x, y) = z$

Communication complexity

Two-player model

- each player has a private input (Alice has $x \in X$, Bob has $y \in Y$)
- players communicate over a channel
- players follow a protocol to compute $f : X \times Y \to Z$
- the last message sent is the value $f(x, y) = z$

The **communication cost** of a protocol is the worst-case length of the full transcript.

Matrix M_f has entries $M_f[x, y] = f(x, y)$.

Matrix M_f has entries $M_f[x, y] = f(x, y)$.
A submatrix is **monochromatic** if f is constant on inputs in the submatrix.

Matrix M_f has entries $M_f[x, y] = f(x, y)$.
A submatrix is **monochromatic** if f is constant on inputs in the submatrix.
A deterministic protocol computing f repeatedly partitions M_f into **rectangles** (submatrices) until every rectangle is monochromatic.

Matrix M_f has entries $M_f[x, y] = f(x, y)$.
A submatrix is **monochromatic** if f is constant on inputs in the submatrix.
A deterministic protocol computing f repeatedly partitions M_f into **rectangles** (submatrices) until every rectangle is monochromatic.

Vickrey auction

The 2-player Vickrey auction is defined as $f : X \times Y \to Z$ where $X = Y = [2^n]$, $Z = [2^{n+1}]$ and $f(x, y) = \begin{cases} (x, B), & \text{if } x \leq y \\ (y, A) & \text{if } y < x \end{cases}$

Matrix M_f has entries $M_f[x, y] = f(x, y)$.
A submatrix is **monochromatic** if f is constant on inputs in the submatrix.
A deterministic protocol computing f repeatedly partitions M_f into **rectangles** (submatrices) until every rectangle is monochromatic.

Vickrey auction

The 2-player Vickrey auction is defined as $f : X \times Y \to Z$ where $X = Y = [2^n]$, $Z = [2^{n+1}]$ and $f(x, y) = \begin{cases} (x, B), & \text{if } x \leq y \\ (y, A) & \text{if } y < x \end{cases}$

	1	2	3	4	…	$2^n - 1$	2^n
1	$(1, B)$	$(1, B)$	$(1, B)$	$(1, B)$	…	$(1, B)$	$(1, B)$
2	$(1, A)$	$(2, B)$	$(2, B)$	$(2, B)$	…	$(2, B)$	$(2, B)$
3	$(1, A)$	$(2, A)$	$(3, B)$	$(3, B)$	…	$(3, B)$	$(3, B)$
4	$(1, A)$	$(2, A)$	$(3, A)$	$(4, B)$	…	$(4, B)$	$(4, B)$
⋮	⋮	⋮	⋮	⋮	⋱	⋮	⋮
$2^n - 1$	$(1, A)$	$(2, A)$	$(3, A)$	$(4, A)$	…	$(2^n - 1, B)$	$(2^n - 1, B)$
2^n	$(1, A)$	$(2, A)$	$(3, A)$	$(4, A)$	…	$(2^n - 1, A)$	$(2^n, B)$

Matrix M_f has entries $M_f[x, y] = f(x, y)$.
A submatrix is **monochromatic** if f is constant on inputs in the submatrix.
A deterministic protocol computing f repeatedly partitions M_f into **rectangles** (submatrices) until every rectangle is monochromatic.

Vickrey auction

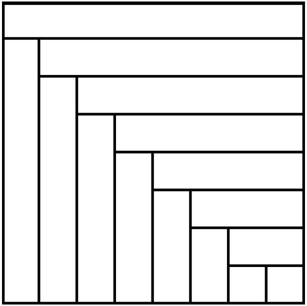
The 2-player Vickrey auction is defined as $f : X \times Y \to Z$ where
$X = Y = [2^n]$, $Z = [2^{n+1}]$ and $f(x, y) = \begin{cases} (x, B), & \text{if } x \leq y \\ (y, A) & \text{if } y < x \end{cases}$

	1	2	3	4	...	$2^n - 1$	2^n
1	$(1, B)$	$(1, B)$	$(1, B)$	$(1, B)$	...	$(1, B)$	$(1, B)$
2	$(1, A)$	$(2, B)$	$(2, B)$	$(2, B)$	...	$(2, B)$	$(2, B)$
3	$(1, A)$	$(2, A)$	$(3, B)$	$(3, B)$	...	$(3, B)$	$(3, B)$
4	$(1, A)$	$(2, A)$	$(3, A)$	$(4, B)$	...	$(4, B)$	$(4, B)$
⋮	⋮	⋮	⋮	⋮	...	⋮	⋮
$2^n - 1$	$(1, A)$	$(2, A)$	$(3, A)$	$(4, A)$	...	$(2^n - 1, B)$	$(2^n - 1, B)$
2^n	$(1, A)$	$(2, A)$	$(3, A)$	$(4, A)$	...	$(2^n - 1, A)$	$(2^n, B)$

Matrix M_f has entries $M_f[x, y] = f(x, y)$.
A submatrix is **monochromatic** if f is constant on inputs in the submatrix.
A deterministic protocol computing f repeatedly partitions M_f into **rectangles** (submatrices) until every rectangle is monochromatic.

Vickrey auction

The 2-player Vickrey auction is defined as $f : X \times Y \to Z$ where
$X = Y = [2^n]$, $Z = [2^{n+1}]$ and $f(x, y) = \begin{cases} (x, B), & \text{if } x \leq y \\ (y, A) & \text{if } y < x \end{cases}$



Matrix M_f has entries $M_f[x, y] = f(x, y)$.
A submatrix is **monochromatic** if f is constant on inputs in the submatrix.
A deterministic protocol computing f repeatedly partitions M_f into **rectangles** (submatrices) until every rectangle is monochromatic.

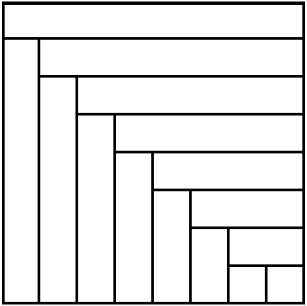
Vickrey auction

The 2-player Vickrey auction is defined as $f : X \times Y \to Z$ where $X = Y = [2^n]$, $Z = [2^{n+1}]$ and $f(x, y) = \begin{cases} (x, B), & \text{if } x \le y \\ (y, A) & \text{if } y < x \end{cases}$

Regions (preimages)

region $R_{x,y} = \{(x', y') \in X \times Y \mid f(x, y) = f(x', y')\}$

defined by **function** $\longrightarrow$



Matrix M_f has entries $M_f[x, y] = f(x, y)$.
A submatrix is **monochromatic** if f is constant on inputs in the submatrix.
A deterministic protocol computing f repeatedly partitions M_f into **rectangles** (submatrices) until every rectangle is monochromatic.

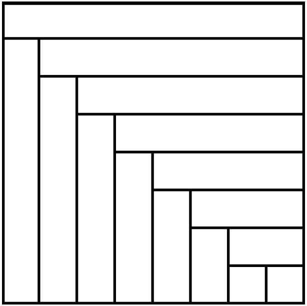
Vickrey auction

The 2-player Vickrey auction is defined as $f : X \times Y \to Z$ where $X = Y = [2^n]$, $Z = [2^{n+1}]$ and $f(x, y) = \begin{cases} (x, B), & \text{if } x \leq y \\ (y, A) & \text{if } y < x \end{cases}$

Regions (preimages)

region $R_{x,y} = \{(x', y') \in X \times Y \mid f(x, y) = f(x', y')\}$

defined by **function** $\longrightarrow$



Rectangles

rectangle $P_{x,y} = \{(x', y') \in X \times Y \mid f(x, y) = f(x', y')$ and $\pi(x, y) = \pi(x', y')\}$

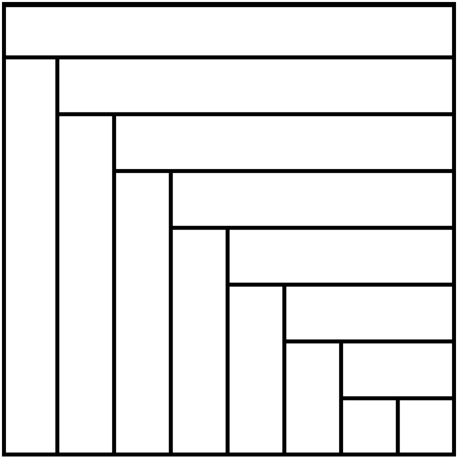
defined by **protocol**

Privacy against eavesdroppers

Can an eavesdropper learn about x and y, aside from $z = f(x, y)$?

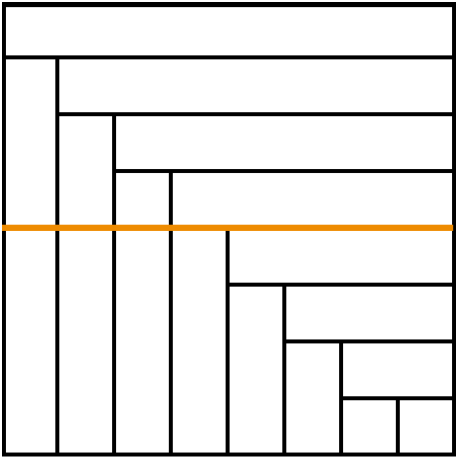
Privacy against eavesdroppers

Can an eavesdropper learn about x and y, aside from $z = f(x, y)$?



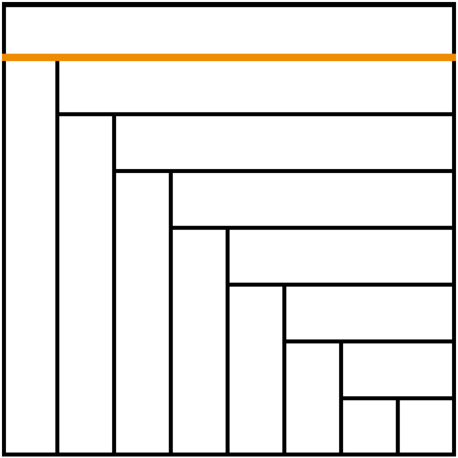
Privacy against eavesdroppers

Can an eavesdropper learn about x and y, aside from $z = f(x, y)$?



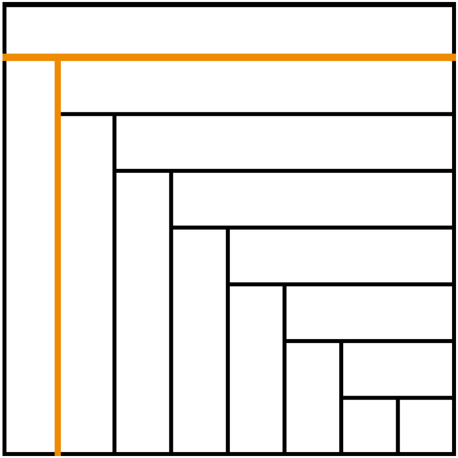
Privacy against eavesdroppers

Can an eavesdropper learn about x and y, aside from $z = f(x, y)$?



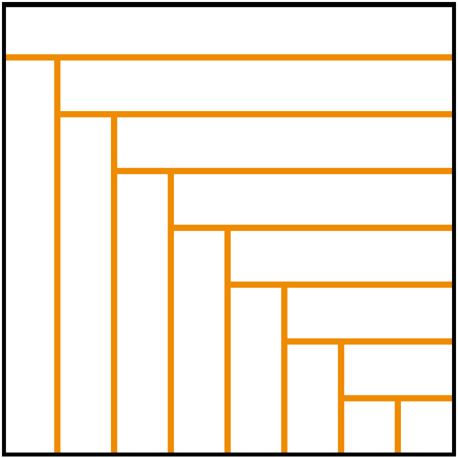
Privacy against eavesdroppers

Can an eavesdropper learn about x and y, aside from $z = f(x, y)$?



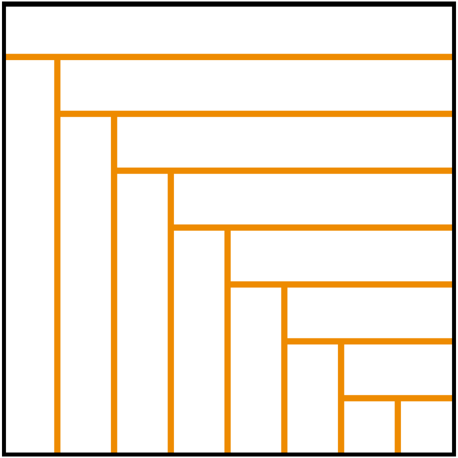
Privacy against eavesdroppers

Can an eavesdropper learn about x and y, aside from $z = f(x, y)$?



Privacy against eavesdroppers

Can an eavesdropper learn about x and y, aside from $z = f(x, y)$?



Ascending English bidding.

Perfect privacy

A protocol for 2-player function $f : X \times Y \rightarrow Z$ is **perfectly private** if every two inputs in the same **region** are partitioned into the same **rectangle**.

Perfect privacy

A protocol for 2-player function $f : X \times Y \to Z$ is **perfectly private** if every two inputs in the same **region** are partitioned into the same **rectangle**.

Characterizing perfect privacy (Kushilevitz '89)

The perfectly private functions of 2 inputs are fully characterized combinatorially. A private deterministic protocol for such functions is given by "decomposing" M_f.

Approximate privacy

Approximate privacy

Privacy approximation ratio (Feigenbaum Jaggard Schapira '10)

A protocol for f has **worst-case privacy approximation ratio**:

$$\text{worst-case PAR} = \max_{(x,y)} \frac{|R_{x,y}|}{|P_{x,y}|}$$

$$\text{average-case PAR} = \mathbb{E}_{(x,y)} \frac{|R_{x,y}|}{|P_{x,y}|} \text{ over distribution } \mathcal{U}$$

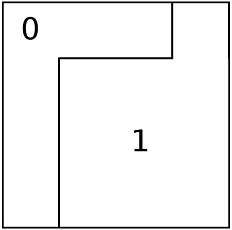
Approximate privacy

Privacy approximation ratio (Feigenbaum Jaggard Schapira '10)

A protocol for f has **worst-case privacy approximation ratio**:

$$\text{worst-case PAR} = \max_{(x,y)} \frac{|R_{x,y}|}{|P_{x,y}|}$$

$$\text{average-case PAR} = \mathbb{E}_{(x,y)} \frac{|R_{x,y}|}{|P_{x,y}|} \text{ over distribution } \mathcal{U}$$



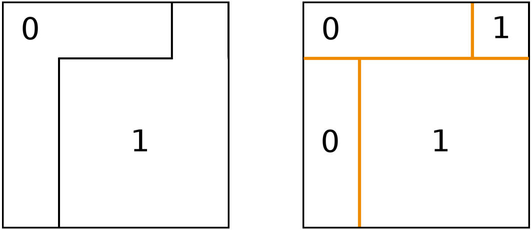
Approximate privacy

Privacy approximation ratio (Feigenbaum Jaggard Schapira '10)

A protocol for f has **worst-case privacy approximation ratio**:

$$\text{worst-case PAR} = \max_{(x,y)} \frac{|R_{x,y}|}{|P_{x,y}|}$$

$$\text{average-case PAR} = \mathbb{E}_{(x,y)} \frac{|R_{x,y}|}{|P_{x,y}|} \text{ over distribution } \mathcal{U}$$



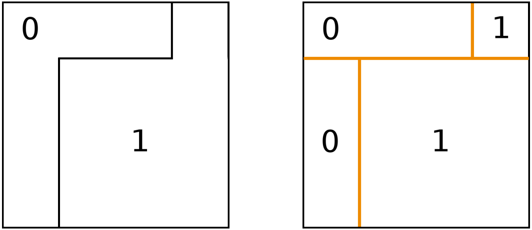
Approximate privacy

Privacy approximation ratio (Feigenbaum Jaggard Schapira '10)

A protocol for f has **worst-case privacy approximation ratio**:

$$\text{worst-case PAR} = \max_{(x,y)} \frac{|R_{x,y}|}{|P_{x,y}|}$$

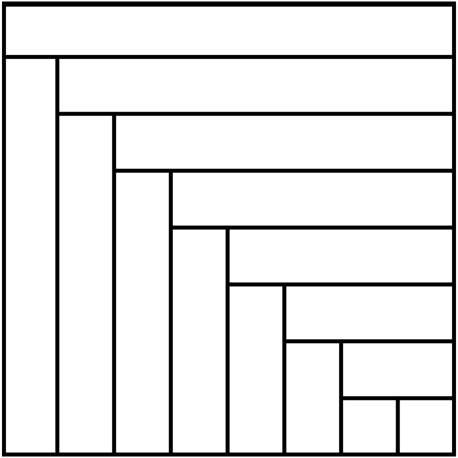
$$\text{average-case PAR} = \mathbb{E}_{(x,y)} \frac{|R_{x,y}|}{|P_{x,y}|} \text{ over distribution } \mathcal{U}$$



worst-case $\mathrm{PAR} = 10$
average-case $\mathrm{PAR} = 2$

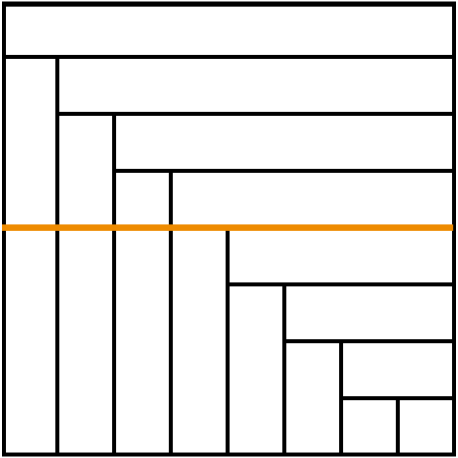
Two-player Vickrey auction

Bisection protocol



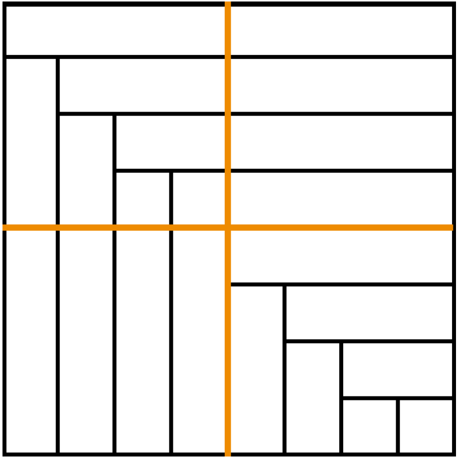
Two-player Vickrey auction

Bisection protocol



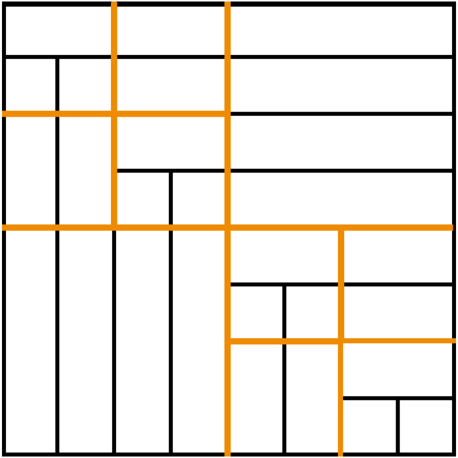
Two-player Vickrey auction

Bisection protocol



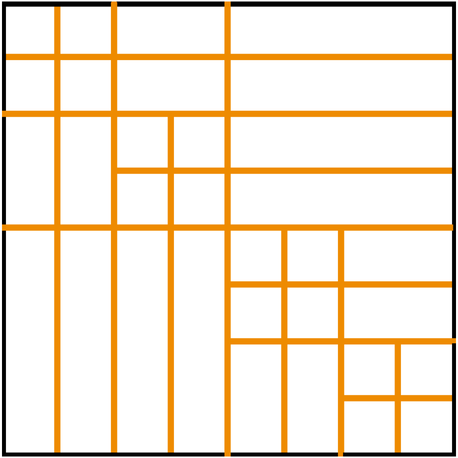
Two-player Vickrey auction

Bisection protocol



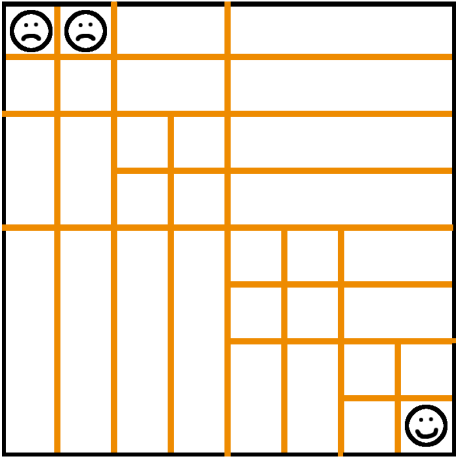
Two-player Vickrey auction

Bisection protocol



Two-player Vickrey auction

Bisection protocol



Upper bounds (Feigenbaum Jaggard Schapira '10)

	English bidding	bisection protocol
communication cost	2^n	$O(n)$
worst-case PAR	1	2^n
average-case PAR	1	$O(1)$

Our contributions

Theorem 1: worst-case lower bound

For all n, for all p, $2 \le p \le n/4$, any deterministic protocol for the n-bit two-player Vickrey auction obtaining PAR less than 2^{p-2} has length at least $2^{n/4p}$.

Our contributions

Theorem 1: worst-case lower bound

For all n, for all p, $2 \leq p \leq n/4$, any deterministic protocol for the n-bit two-player Vickrey auction obtaining PAR less than 2^{p-2} has length at least $2^{n/4p}$.

Theorem 2: average-case lower bound

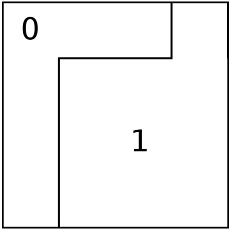
For all $n, r \geq 1$, any deterministic protocol of length at most r for the n-bit two-player Vickrey auction has average-case PAR greater than $\Omega(\frac{n}{\log(r/n)})$.

Privacy against players

Can Bob learn anything about Alice's private input x, beyond the fact that $z = f(x, y)$? Can Alice learn anything about Bob's private input y?

Privacy against players

Can Bob learn anything about Alice's private input x, beyond the fact that $z = f(x, y)$? Can Alice learn anything about Bob's private input y?



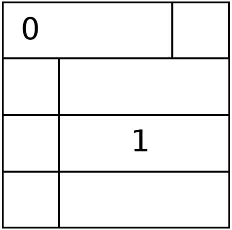
Privacy against players

Can Bob learn anything about Alice's private input x, beyond the fact that $z = f(x, y)$? Can Alice learn anything about Bob's private input y?

Subjective regions

region $R^A_{x,y} = \{(x, y') \in X \times Y \mid f(x, y) = f(x, y')\}$

defined by **function** Alice sees



Privacy against players

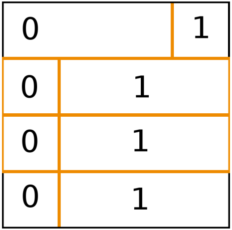
Can Bob learn anything about Alice's private input x, beyond the fact that $z = f(x, y)$? Can Alice learn anything about Bob's private input y?

Subjective regions

region $R^A_{x,y} = \{(x, y') \in X \times Y \mid f(x, y) = f(x, y')\}$

defined by **function**

Alice sees



Subjective rectangles

rectangle $P^B_{x,y} = \{(x, y') \in X \times Y \mid f(x, y) = f(x, y'), \pi(x, y) = \pi(x, y')\}$

defined by **protocol**

Alice sees

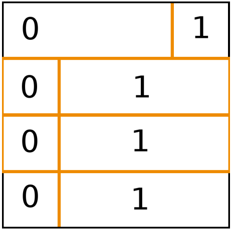
Privacy against players

Can Bob learn anything about Alice's private input x, beyond the fact that $z = f(x,y)$? Can Alice learn anything about Bob's private input y?

Subjective regions

region $R^A_{x,y} = \{(x,y') \in X \times Y \mid f(x,y) = f(x,y')\}$

defined by **function** Alice sees



Subjective rectangles

rectangle $P^B_{x,y} = \{(x,y') \in X \times Y \mid f(x,y) = f(x,y'), \pi(x,y) = \pi(x,y')\}$

defined by **protocol** Alice sees

Subjective privacy approximation ratio (Feigenbaum Jaggard Schapira '10)

$$\text{average-case } \mathrm{PAR}^{\text{sub}} = \max_{v=A,B} \mathbb{E}_{(x,y)} \frac{|R^v_{x,y}|}{|P^v_{x,y}|}$$

Theorem (Braverman '11): IC(DISJ)$= \Omega(n)$.

information cost IC

$$IC = I(\mathbf{X} : \Pi_P(\mathbf{X}, \mathbf{Y})|\mathbf{Y}) + I(\mathbf{Y} : \Pi_P(\mathbf{X}, \mathbf{Y})|\mathbf{X})\}$$

informational privacy PRIV_D (Klauck '02)

$$\mathrm{PRIV}_D(P) = \max\{I(\mathbf{X} : \Pi_P(\mathbf{X}, \mathbf{Y})|\mathbf{Y}, f(\mathbf{X}, \mathbf{Y})), I(\mathbf{Y} : \Pi_P(\mathbf{X}, \mathbf{Y})|\mathbf{X}, f(\mathbf{X}, \mathbf{Y}))\}$$

Theorem: $\mathrm{PRIV}_D - \log|Z| \leq IC \leq 2(\mathrm{PRIV}_D + \log|Z|)$
Theorem: $\mathrm{PRIV}_D(P) \leq \log(\mathrm{avg}_D \, \mathrm{PAR}^{sub}(P))$

Theorem 3

Any protocol P computing the n-bit Set Intersection $\mathrm{INTERSEC}_n$ has exponential average-case subjective PAR:

$$\mathrm{avg}_{\mathcal{U}} \, \mathrm{PAR}^{sub}(P) = 2^{\Omega(n)}$$

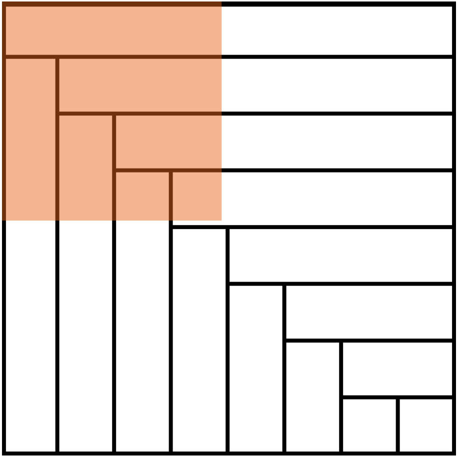
Theorem 1: worst-case lower bound

For all n, for all p, $2 \leq p \leq n/4$, any deterministic protocol for the n-bit two-player Vickrey auction problem obtaining PAR less than 2^{p-2} has length at least $2^{n/4p}$.

Theorem 1: worst-case lower bound

For all n, for all p, $2 \leq p \leq n/4$, any deterministic protocol for the n-bit two-player Vickrey auction problem obtaining PAR less than 2^{p-2} has length at least $2^{n/4p}$.

progress: steps that look like bisection.



useless: steps that look like English bidding.

Future directions

- a "good" unified definition of privacy
- length-privacy tradeoffs for other functions
- general results for length-privacy tradeoffs
- randomized protocols
- protocols with error
- approximate privacy hierarchy?
- more than 2 players
- privacy against coalitions?

Ball Partition Problem

For integers N and $r \geq 1$, there are N balls and r rounds. All of the balls begin in one big set. In each round, the balls in each current set are partitioned into (at most) two new sets. The cost of partitioning the balls in any set S into sets S_1 and S_2 is $\min(|S_1|, |S_2|)$. After r rounds, each of the N balls shall be in a singleton set. The total cost of the game is the sum of the cost, over all r rounds, of every partition made during each round. We denote the minimal possible cost by $B(N, r)$.

Theorem 17

For the Ball Partition Problem, $B(N, r) \geq \frac{N \log N}{4 \log(\frac{4r}{\log N})}$.

Average-case PAR

We define it slightly differently.

Theorem 2: average-case lower bound

For all $n, r \geq 1$, any deterministic protocol of length at most r for the n-bit two-player 2^n-Vickrey auction problem has average-case PAR greater than $\Omega(\frac{n}{\log(r/n)})$ (over the uniform distribution of inputs).

Proof: The Ball Partition problem simplifies the analysis of arbitrary protocols to an analysis of protocol trees and probability.

Answered Feigenbaum conjecture about set intersection:

Theorem 3

For all $n \geq 1$, and any protocol P computing the Set Intersection INTERSEC_n the average-case subjective PAR is exponential in n under the uniform distribution:

$$\mathrm{avg}_{\mathcal{U}} \, \mathrm{PAR}^{sub}(P) = 2^{\Omega(n)}$$

Relating PAR to info measures. Definitions of mutual information measures of privacy (nice because info=0 corresponds to perfect privacy) [Kla02, Bra11]. *Theorem 21* Info theoretic privacy $\leq$ log of average PAR.

$$\mathrm{PRIV}_D(P) \leq \log(\mathrm{avg}_D \, \mathrm{PAR}^{sub}(P))$$

Mark Braverman.
Interactive information complexity.
Electronic Colloquium on Computational Complexity, (123), 2011.

Felix Brandt and Tuomas Sandholm.
On the Existence of Unconditionally Privacy-Preserving Auction Protocols.
ACM Transactions on Information and System Security, 11(2):1–21, May 2008.

Joan Feigenbaum, Aaron D Jaggard, and Michael Schapira.
Approximate Privacy: Foundations and Quantification.
Proceedings of the 11th Conference on Electronic Commerce, pages 167–178, 2010.

Joan Feigenbaum, Aaron D Jaggard, and Michael Schapira.
Approximate Privacy: PARs for Set Problems.
DIMACS Technical Report 2010-01, pages 1–34, 2010.

Hartmut Klauck.
On quantum and approximate privacy.
Proc. STACS, 2002.

Eyal Kushilevitz.
Privacy and communication complexity.
30th Annual Symposium on Foundations of Computer Science, pages 416–421, 1989.

Andrew Chi-chih Yao.
Some Complexity Questions Related to Distributive Computing.
Proceedings of the 11th Annual ACM Symposium on Theory of Computing (STOC'79), pages 209–213, 1979.