ε-Privacy:
Data Publishing against Realistic Adversaries

Speaker:

Michaela Götz

Joint work with:

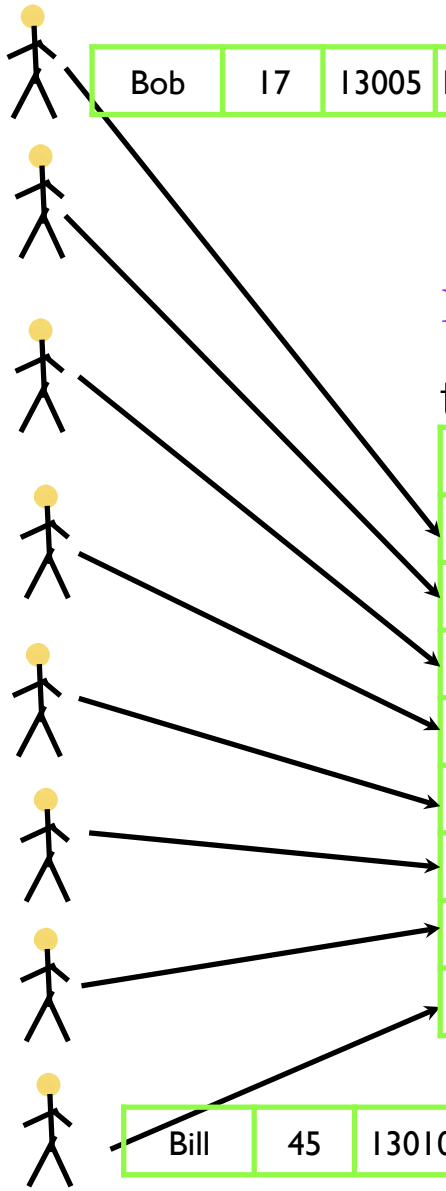
Ashwin Machanavajjhala and Johannes Gehrke

YAHOO! RESEARCH

Cornell University

Setting

Individuals



Bob	17	13005	Heart Disease

Data Curator

table T

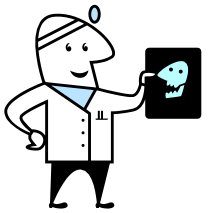
Name	Age	Zip	Disease
Bob	17	13005	Heart Disease
Jim	19	13000	Viral Infection
Cathy	20	14850	Cancer
Anne	24	14850	Heart Disease
Joe	29	14850	Viral Infection
Marie	34	13005	Cancer
Dana	39	13005	Cancer
Bill	45	13010	Cancer

Bill	45	13010	Cancer

Published table T'

Age	Zip	Disease
< 20	1300*	Heart Disease
< 20	1300*	Viral Infection
2*	14850	Cancer
24	14850	Heart Disease
29	14850	Viral Infection
34	130**	Cancer
39	130**	Cancer
45	130**	Cancer

Users



Privacy - Overview

- What is sensitive information?
  - "Bob has ulcer"
  - "Bob has some stomach disease"
- What is privacy?
  - Adversary *does not learn much about* Bob's sensitive information. [perfect privacy, t-closeness, alpha-beta privacy, …]
  - Adversary *learns the same* about Bob whether or not that Bob's information is part of the release. [differential privacy]
- What does the adversary know about T ?

- *Goal:* Data Publishing Mechanism

Adversarial knowledge

Pr[Bob has Cancer] = 1/3
Pr[Bob has Heart Disease] = 1/3
Pr[Bob has a Viral Infection] = 1/3

Cathy	20	14850	Cancer
Anne	24	14850	Heart Disease

anti-corruption privacy

weak adversaries

Adversary's strength

extremely strong adversaries

Adversarial knowledge

Pr[Bob has Cancer] = 1/3
Pr[Bob has Heart Disease] = 1/3
Pr[Bob has a Viral Infection] = 1/3

Bob	17	13005	[redacted]
Jim	19	13000	Viral Infection
Cathy	20	14850	Cancer
Anne	24	14850	Heart Disease
Joe	29	14850	Viral Infection
Marie	34	13005	Cancer
Dana	39	13005	Cancer
Bill	45	13010	Cancer

Fixed distribution over sensitive values

as in T

uniform

t-closeness
l-diversity
proximity privacy
anti-corruption privacy

differential privacy

weak adversaries

Adversary's strength

extremely strong adversaries

Adversarial knowledge

Pr[Bob has Cancer] = 0
Pr[Bob has Heart Disease] = .7
Pr[Bob has a Viral Infection] = .3

Bob	17	13005	[redacted]
Jim	19	13000	Viral Infection
Cathy	20	14850	Cancer
Anne	24	14850	Heart Disease
Joe	29	14850	Viral Infection
Marie	34	13005	Cancer
Dana	39	13005	Cancer
Bill	45	13010	Cancer

t-closeness
l-diversity
proximity privacy
anti-corruption privacy

differential privacy

weak adversaries

Adversary's strength

extremely strong adversaries

Adversarial knowledge

t-closeness
l-diversity
proximity privacy
anti-corruption privacy

?

differential privacy

weak adversaries

Adversary's strength

extremely strong adversaries

Outline

- ε-Privacy – definition
  - Realistic adversaries
  - Privacy guarantee
- A privacy-preserving mechanism
  - Generalization algorithm
  - Utility experiments
- Instantiation of other privacy guarantees

ε-Privacy: Adversaries

- Knowledge about the individuals in T
  - Complete information about a few individuals in T.
- Knowledge about the Population:

Where does the prior belief come from?
External data.

  - Adversary is forming her prior based on external data.
  - Given the published table T' she updates her belief
  - How much her belief changes depends on her "stubbornness"

Adversary's statistical knowledge

- Some probability distribution p over sensitive values generates the sensitive values for the population.
  - Example: $p = (.2, .5, .3)$, but maybe $p = (.2, .45, .35)$
- Uncertainty about p depends on size of external data
  - Example: pretty sure $p = (.2, .5, .3)$
- 2 step process:
  1. choose distribution p over sensitive values
  2. for each individual choose sensitive value i w.p. p_i

- Natural choice for categorical attributes: Dirichlet Distribution $D(\sigma_1, \ldots, \sigma_s)$
  - shape $\sigma_1, \ldots, \sigma_s$, stubbornness $\sigma = \Sigma\, \sigma_i$

Knowledge about Population

Disease	Count
Cancer	2 M
Viral Infection	5 M
Heart Disease	3 M

Adversary's statistical knowledge

- Dirichlet Distribution $D(\sigma_1, \ldots, \sigma_s)$
  - shape $\sigma_1, \ldots, \sigma_s$, stubbornness $\sigma = \Sigma\, \sigma_i$
- Adversary is forming her prior based on external data.
  - Table T -> $D(\sigma_1, \ldots, \sigma_s)$, e.g. D(1000, 3000, 500)
- Given the published table T' she updates her belief
  - Conditioning, e.g. Pr[Bob has Cancer | T', D(1000, 3000, 500)]
- How much her belief changes depends on her "stubbornness"
  - Parameter σ in Dirichlet

Knowledge about Population

Disease	Count
Cancer	1000
Viral Infection	3000
Heart Disease	500

Privacy definition

- Differential privacy for restricted adversaries:

An adversary in class A learns roughly the same about an individual no matter whether or not that individual's data is contained in the release.

Privacy definition

table T

Name	Age	Zip	Disease
Bob	17	13005	Heart Disease
Jim	19	13000	Viral Infection
Cathy	20	14850	Cancer
Anne	24	14850	Heart Disease
Joe	29	14853	Viral Infection
Marie	34	13005	Cancer
Dana	39	13005	Cancer
Bill	45	13010	Cancer

table T without Bob

Name	Age	Zip	Disease
Jim	19	13000	Viral Infection
Cathy	20	14850	Cancer
Anne	24	14850	Heart Disease
Joe	29	14853	Viral Infection
Marie	34	13005	Cancer
Dana	39	13005	Cancer
Bill	45	13010	Cancer

table T'

Age	Zip	Disease
< 20	1300*	Heart Disease Viral Infection
2*	14850	Viral Infection Cancer Heart Disease
>20	130**	Cancer Cancer Cancer

Adversary's posterior belief that Bob has Cancer is roughly the same in both cases.

Age	Zip	Disease
< 20	1300*	Viral Infection
2*	14850	Viral Infection Cancer Heart Disease
>20	130**	Cancer Cancer Cancer

Adversarial reasoning - Example

- Prior: D(1000, 3000, 500)
- Posterior belief about:

Bob	17	13005

Knowledge about Population

Disease	Count
Cancer	1000
Viral Infection	3000
Heart Disease	500

table T' (with Bob)

Age	Zip	Disease	Count
< 20	1300*	Heart Disease Viral Infection	2001 3000
2*	14850	Viral Infection Breast Cancer	7000 1000
>30	130**	Viral Infection Breast Cancer Heart Disease	500 2000 700

Pr[Bob has a Heart Disease | T', D]
= 2001/5001
Pr[Bob has a Viral Infection | T', D]
= 3000/5001
Pr[Bob has Cancer | T', D]
= 0

Adversarial reasoning - Example

- Prior: D(1000, 3000, 500)
- Posterior belief about:

Bob	17	13005

Knowledge about Population

Disease	Count
Cancer	1000
Viral Infection	3000
Heart Disease	500

table T' (without Bob)

Age	Zip	Disease	Count
< 20	1300*	Heart Disease Viral Infection	2000 3000
2*	14850	Viral Infection Breast Cancer	7000 1000
>30	130**	Viral Infection Breast Cancer Heart Disease	500 2000 700

Pr[Bob has a Heart Disease | T', D]
= (2000+500)/(5000+4500)
Pr[Bob has a Viral Infection | T', D]
= (3000+3000)/(5000+4500)
Pr[Bob has Cancer | T', D]
= (0+1000)/(5000+4500)

Adversarial reasoning - Example

- Prior: D(1000, 3000, 500)
- Posterior belief about:

Bob	17	13005

	table T'	T'- Bob
Pr[Bob has a Heart Disease]	0.40	0.26
Pr[Bob has a Viral Infection]	0.60	0.63
Pr[Bob has a Cancer]	0	0.11

Adversarial reasoning - Example

- Prior: D(1000, 3000, 500) D(500, 1500, 250)

- Posterior belief about:

Bob	17	13005

	table T'	T'- Bob	
Pr[Bob has a Heart Disease]	0.40	0.26	0.31
Pr[Bob has a Viral Infection]	0.60	0.63	0.62
Pr[Bob has a Cancer]	0	0.11	0.07

Knowledge about Population

Disease	Count	
Cancer	1000	500
Viral Infection	3000	1500
Heart Disease	500	250

Adversarial reasoning

Observation:

If generalization T' preserves ε-privacy against adversary $D(\sigma_1, \ldots, \sigma_s)$

then it also preserves ε-privacy against adversary $D(r^*\sigma_1, \ldots, r^*\sigma_s')$ for $r^* < 1$.

Smaller Stubbornness -> easier to achieve ε-privacy.

Adversarial reasoning - Example

- Prior: D(r*1000, r*3000, r*500) take r -> ∞
- Posterior belief about:

Bob	17	13005

	Posterior Belief table T'	T'- Bob
Pr[Bob has a Heart Disease]	0.40	0.22
Pr[Bob has a Viral Infection]	0.60	0.67
Pr[Bob has a Cancer]	0	0.11

Adversarial reasoning

Observation:

Infinitely stubborn adversaries belief that Pr[Bob has Disease i] = σ_i/σ

Infinitely stubborn adversaries do not update their belief about the population given T'.

Higher Stubbornness -> less the adversary learns from T' about population.

Adversarial classes

	Stubbornness	Shape
Class I:	σ	σ(Heart), σ(Virus), σ(Cancer)
Class II:	σ	arbitrary
Class III:	∞	σ(Heart), σ(Virus), σ(Cancer)
Class IV:	∞	arbitrary

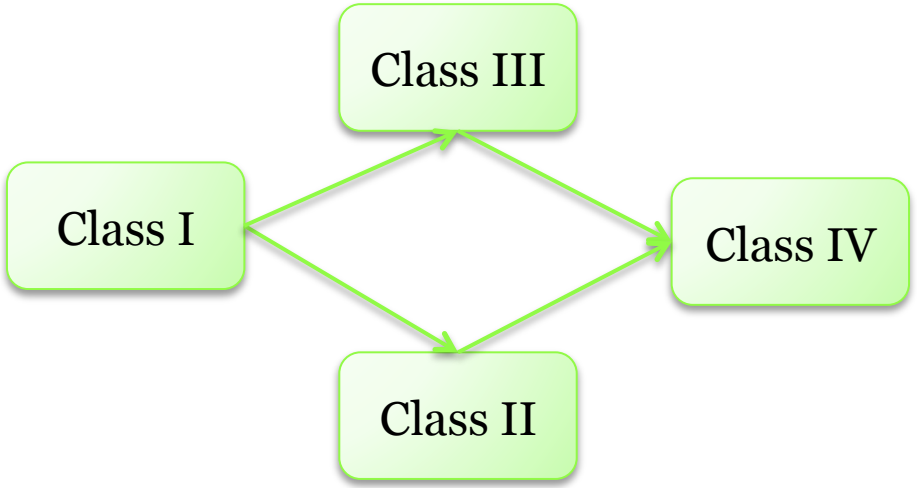
} *realistic*

{ t-closeness
l-diversity

differential privacy

Adversarial classes

	Stubbornness	Shape
Class I:	$\leq \sigma$	σ(Heart), σ(Virus), σ(Cancer)
Class II:	$\leq \sigma$	arbitrary
Class III:	$\leq \infty$	σ(Heart), σ(Virus), σ(Cancer)
Class IV:	$\leq \infty$	arbitrary



Outline

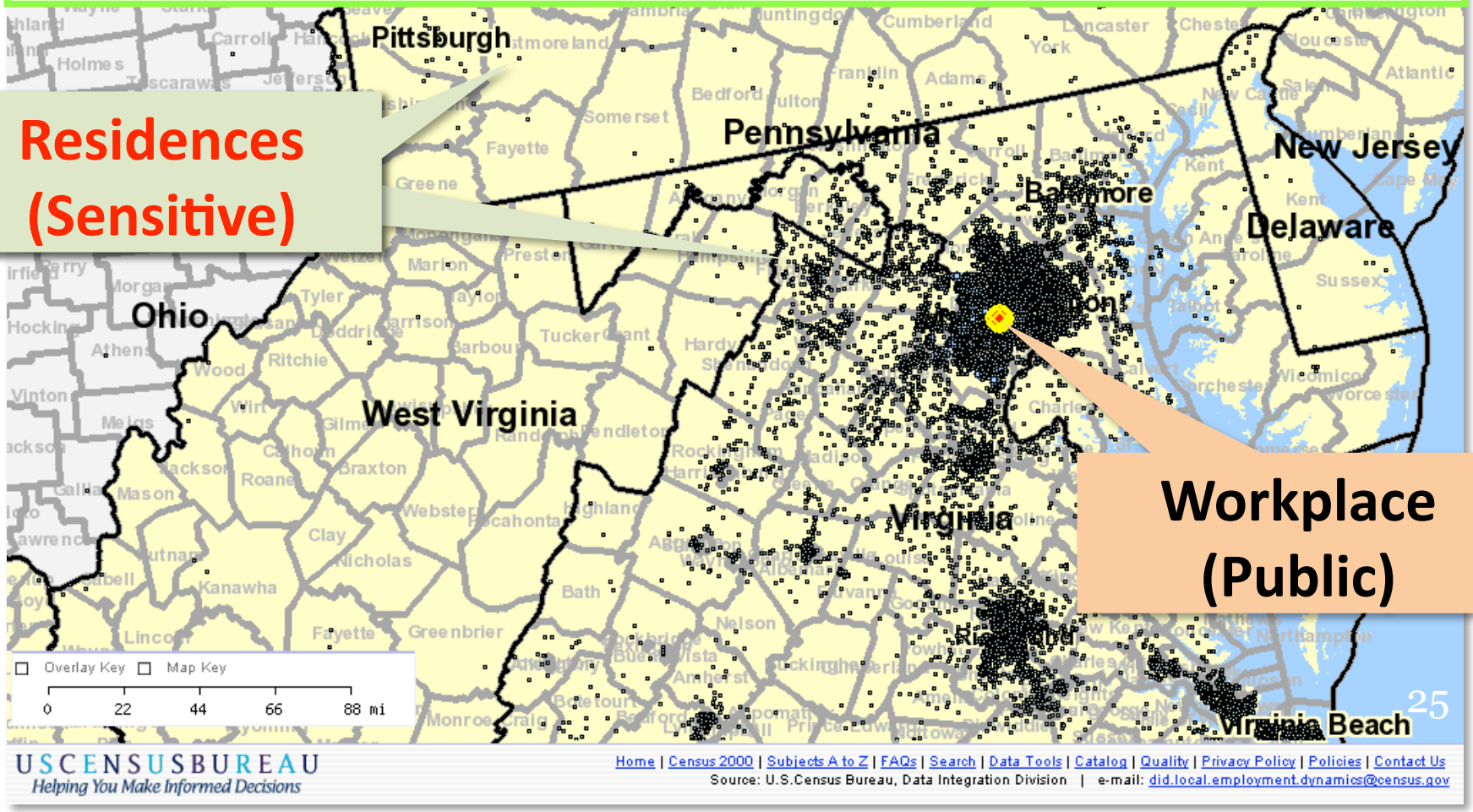
- ε-Privacy - definition
  - Adversaries with statistical knowledge
  - Privacy guarantee
- An ε-private mechanism
  - Generalization algorithm
  - Utility experiments
- Instantiation of other privacy guarantees

An ε-private generalization algorithm

- Input:
  - Table T
  - Specification of sensitive information!
  - Choice of adversaries!
    - $D(\sigma_1, \ldots, \sigma_s)$: shape $\sigma_1, \ldots, \sigma_s$, stubbornness σ
    - Complete Knowledge about a few individuals in T
  - Choice of privacy parameter ε!
- Output:
  - Generalization T'
    - ε-private
    - useful

Choosing the adversarial class

http://lehdmap3.dsd.census.gov/



Choosing the adversarial class

- Example: U.S. Census wants to publish ε-private commute patterns.
- 1a) Based on previous releases set upper bound on stubbornness.
  - Example: Set stubbornness = number of individuals in previous versions of commute patterns.
- 1b) Fix shape if possible.
  - Example: Either set shape = distribution in previous releases or do not make assumptions about the shape.
- 2 Upper bound number of individuals the adversary has complete knowledge about.

Create a generalized table T'

- a) Check T' preserves ε-privacy against an adversary with belief $D(\sigma_1, \ldots, \sigma_s)$:
  - All non-sensitive groups with n tuples out of which n(s) have sensitive value s:

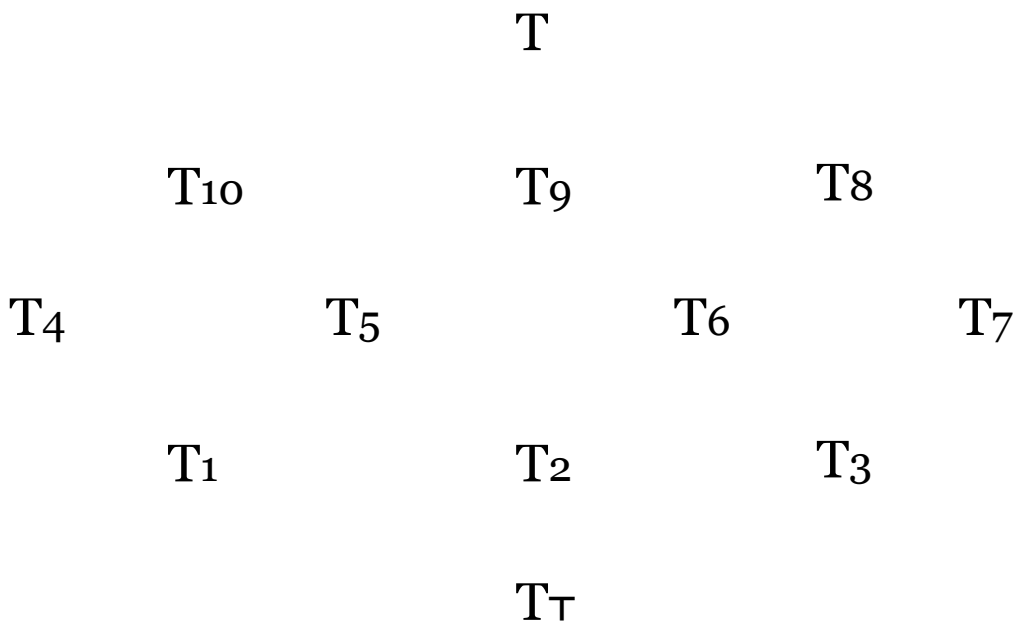
    $n \geq \Phi(\sigma, D, \varepsilon)$

    $n(s)/n \leq \Phi'(\sigma, D, \varepsilon, n)$

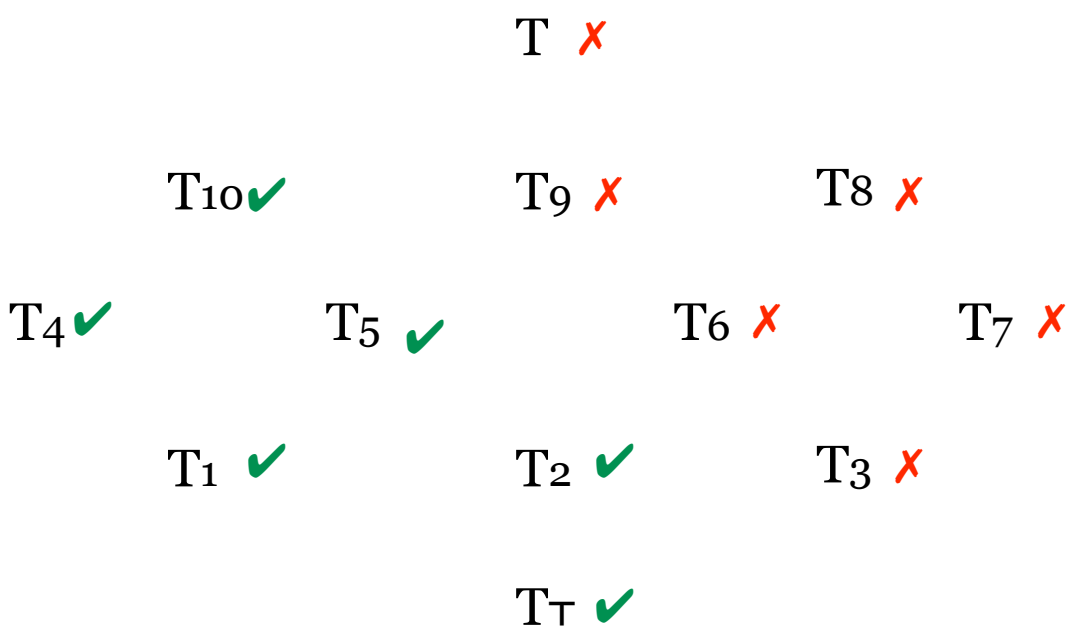
    - Easy to check.
    - Can derive condition for the other classes.

- b) Pick the one that maximizes utility.

a) Check privacy of ALL generalized tables

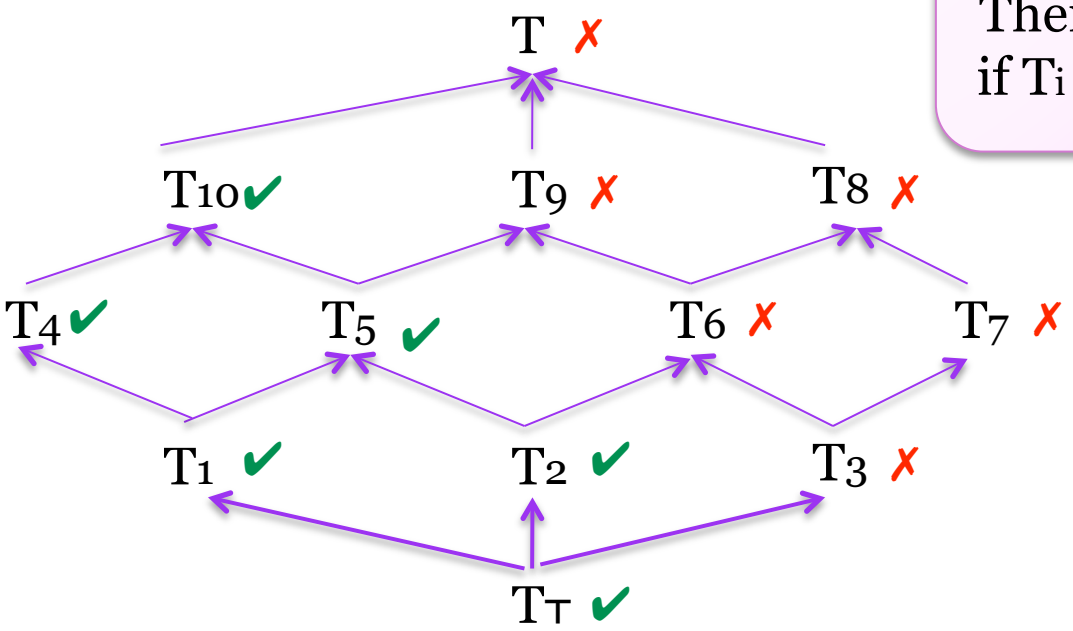


a) Check privacy of ALL generalized tables



a) Check privacy

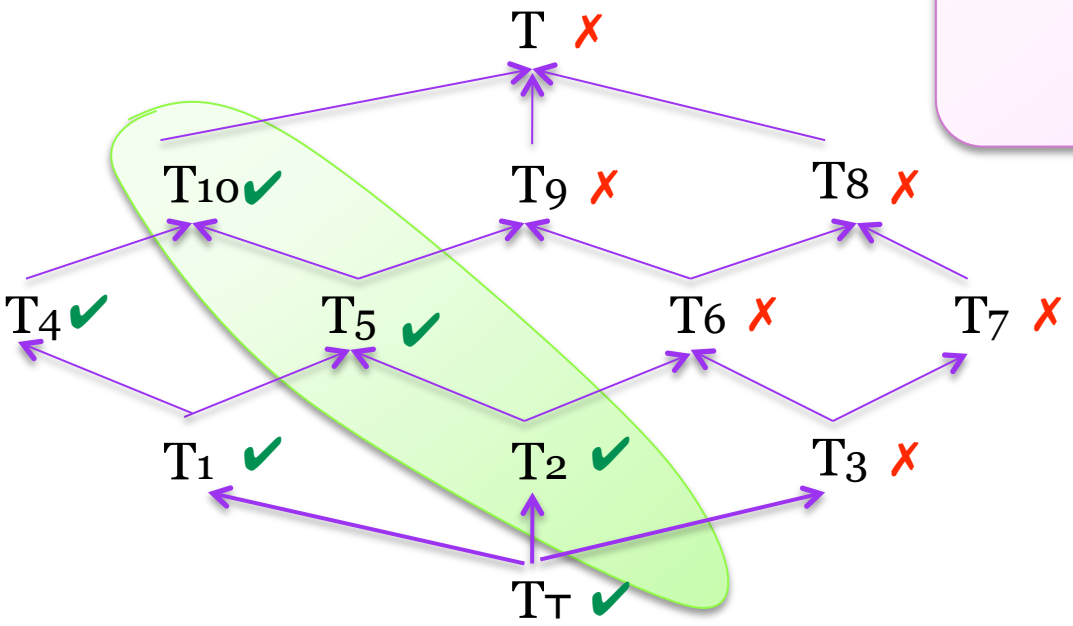
Generalization Lattice:
There is a path from T_i to T_j if T_i is a generalization of T_j.



- Observation: Privacy is monotonic.
- Assumption: Utility function is monotonic.

b) Maximize utility

Use Incognito or Mondrian to find a privacy preserving generalization with maximum utility.



- Observation: Privacy is monotonic.
- Assumption: Utility function is monotonic.

Experiments

- Compare privacy-utility tradeoff
  - Across classes of adversaries
  - Across privacy definitions (l-diversity, t-closeness)
- Utility
  - Metric: discernibility, Avg. group size

Experiments

- Compare privacy-utility tradeoff
  - Across classes of adversaries
  - Across privacy definitions (l-diversity, t-closeness)
- Utility
  - Metric: discernibility, Avg. group size
- Data: American Community Survey ~ 3 million tuples

Attribute	|Domain|	Generalization	Height
Age	73	Ranges – 5, 10, 20, 40, *	6
Marital St.	6	Taxonomy	3
Race	9	*	2
Gender	2	*	2
Salary class	2	*Sensitive Attr.*	-

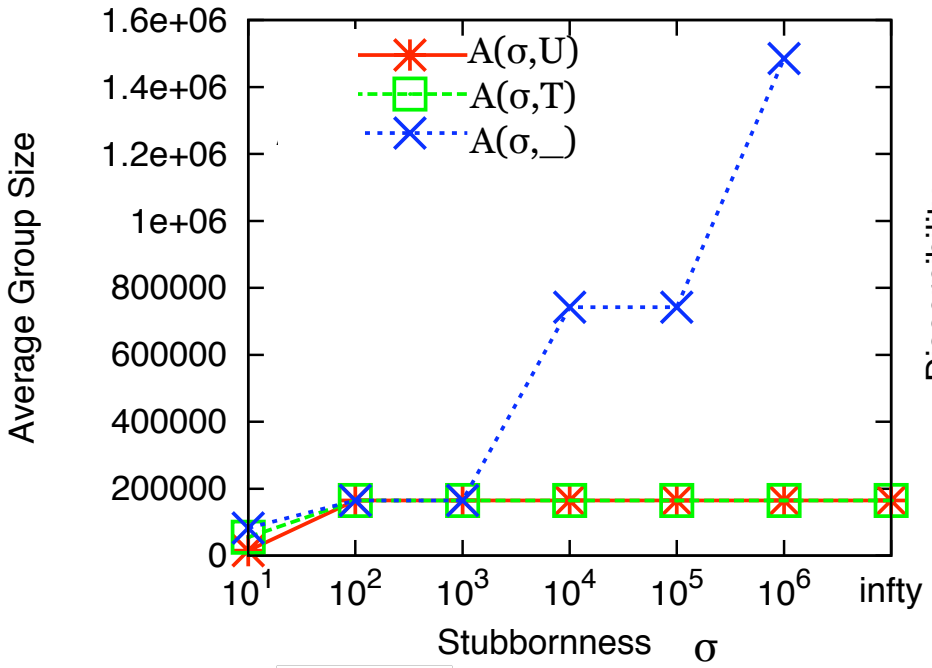
Realistic vs. Unrealistic Adversaries

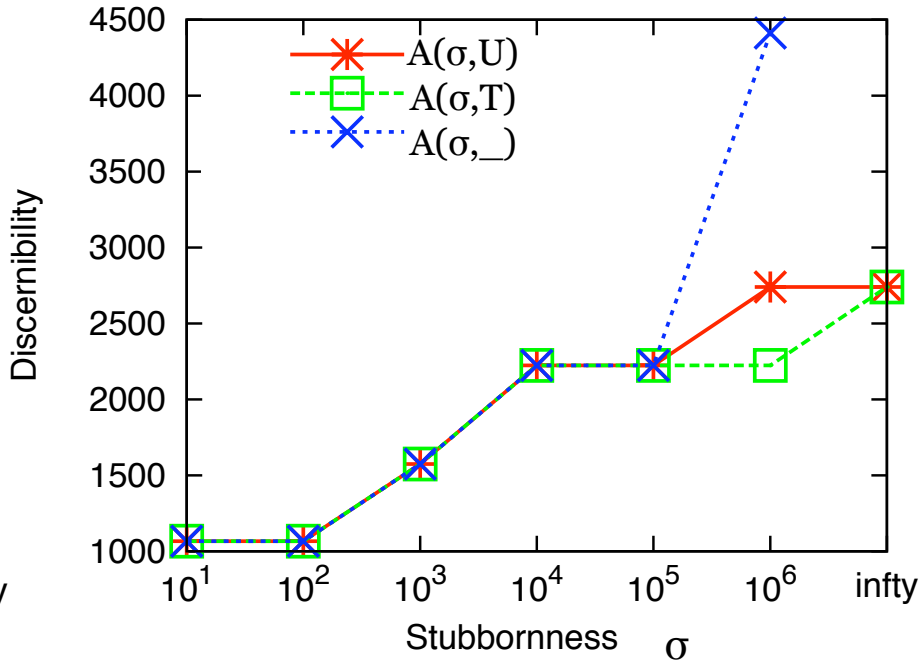
- Classes
  - Prior: Uniform, as in T, arbitrary _
  - Stubbornness: σ ≤ {10, 10^2, …, 10^6, ∞}

  - Class I: A(U, σ), A(T, σ), for σ ≤ {10, 10^2, …, 10^6}
  - Class II: A(_, σ), for σ ≤ {10, 10^2, …, 10^6}
  - Class III: A(U, ∞), A(T, ∞)
  - Class IV: A(_, ∞)

Realistic vs. Unrealistic Adversaries

- The effect of the stubbornness on utility





ε = 2.5

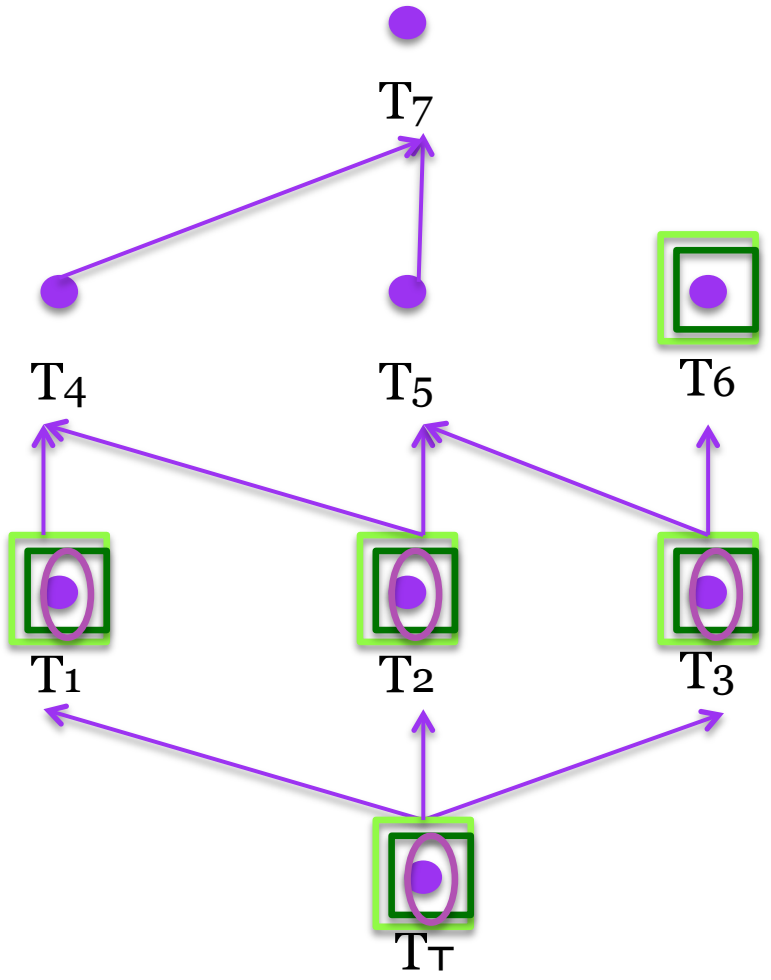
Comparison to other Privacy Guarantees

(4,2)-diversity

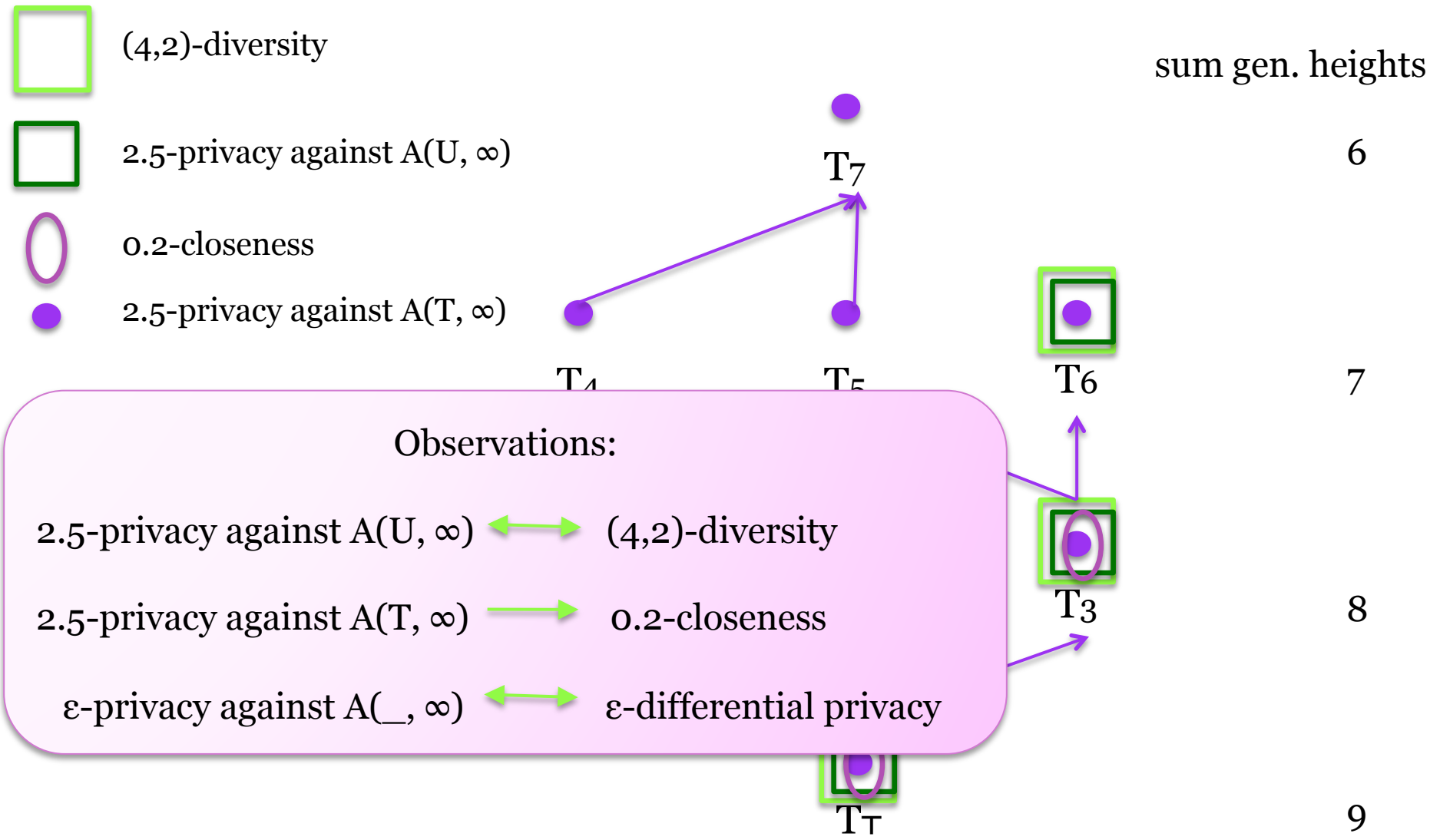
2.5-privacy against A(U, ∞)

0.2-closeness

2.5-privacy against A(T, ∞)



Comparison to other Privacy Guarantees



Summary

- Realistic Adversaries
  - Have statistical knowledge about the population
  - Form prior based on external data
  - Update their belief
- Publishing Generalizations:
  - Practical Trade-offs between Privacy and Utility
  - Instantiate other guarantees ($\sigma \rightarrow \infty$)

Future Work

- Extend Background Knowledge:
  - Prior over non-sensitive attributes
  - Negation statements
- Study other Sanitization Algorithms:
  - Synthetic data
  - Interactive queries

Questions?

B. *de Finetti*. "Funzione caratteristica di un fenomeno aleatorio." *Mathematice e Naturale, 1931*.

F. *Bacchus et al.*, "From statistics to beliefs." *AAAI 1992*

L. *Sweeney*, "k-Anonymity: A Model for Protecting Privacy", *IJUFKS, 2002*

A. *Evfimievski et al.*, "Limiting Privacy Breaches in Privacy Preserving Data Mining", *PODS 2003*

G. *Miklau et al.*, "A Formal Analysis of Information Disclosure in Data Exchange", *SIGMOD 2004*

K. *LeFevre et al.*, "Incognito: Efficient Full Domain k-Anonymity", *SIGMOD 2005*

A. *Machanavajjhala et al.*, "L-Diversity: Privacy beyond K-Anonymity", *ICDE 2006*

C. *Dwork*. "Differential privacy", *ICALP 2006*.

N. *Li et al.*, "t-Closeness: Privacy beyond K-Anonymity and L-Diversity", *ICDE 2007*

Y. *Tao et al.*, "On anti-corruption privacy preserving publication.", *ICDE 2008*