How Not to Preserve Privacy

k-Anonymity: a Model For Protecting Privacy

Latanya Sweeny

Releasing a Database

Name	ID Number	Gender	Birth Date	Zip code	Medical Condition
Amy Colon	523950649	Male	11/02/1956	98097	Uncommon Cold
Inga Hull	991039441	Male	14/09/1967	48254	Heaped Piles
Jessica Walls	746510555	Male	02/04/1954	66950	Bloaty Head
Cameran Prince	272922661	Female	19/10/1953	89395	Uncommon Cold
Beatrice Oliver	367636643	Male	20/02/1950	58484	Slack Tongue
Stewart Schroeder	573424830	Female	12/08/1969	78345	Heaped Piles
Guy Cleveland	426525813	Female	05/01/1970	69107	Slack Tongue
Yoshi Sweet	744617659	Female	29/07/1960	66015	Uncommon Cold
Herman Wilkerson	355495414	Male	29/11/1970	12794	Uncommon Cold
Lara Shaffer	930512852	Female	18/06/1961	76031	Bloaty Head
Wynter Bryan	385448496	Female	09/02/1971	68597	Slack Tongue
Adria Mcbride	337515106	Female	15/11/1968	18392	Bloaty Head
Eugenia Key	322441746	Female	24/03/1967	46997	Uncommon Cold
Rowan Barrera	383749474	Male	31/05/1952	63570	Heaped Piles
Urielle Riley	795856737	Female	01/08/1985	08603	Uncommon Cold
Caesar Lancaster	995946734	Male	01/01/1986	93861	Uncommon Cold
Irene Curry	046498803	Male	09/04/1978	87454	Slack Tongue
Aline Hess	865009451	Female	05/06/1966	78956	Bloaty Head
Peter Calderon	336136140	Female	17/04/1987	60254	Bloaty Head
Hu Parrish	693587559	Male	03/02/1984	51213	Bloaty Head
Valentine Haynes	048717454	Female	10/04/1965	86362	Uncommon Cold
Amos Edwards	025759543	Male	13/07/1954	13197	The Squits

Tuple

Attribute

Releasing a Database

Name	ID Number	Gender	Birth Date	Zip code	Medical Condition
Amy Colon	523950649	Male	11/02/1956	98097	Uncommon Cold
Inga Hull	991039441	Male	14/09/1967	48254	Heaped Piles
Jessica Walls	746510555	Male	02/04/1954	66950	Bloaty Head
Cameran Prince	272922661	Female	19/10/1953	89395	Uncommon Cold
Beatrice Oliver	367636643	Male	20/02/1950	58484	Slack Tongue
Stewart Schroeder	573424830	Female	12/08/1969	78345	Heaped Piles
Guy Cleveland	426525813	Female	05/01/1970	69107	Slack Tongue
Yoshi Sweet	744617659	Female	29/07/1960	66015	Uncommon Cold
Herman Wilkerson	355495414	Male	29/11/1970	12794	Uncommon Cold
Lara Shaffer	930512852	Female	18/06/1961	76031	Bloaty Head
Wynter Bryan	385448496	Female	09/02/1971	68597	Slack Tongue
Adria Mcbride	337515106	Female	15/11/1968	18392	Bloaty Head
Eugenia Key	322441746	Female	24/03/1967	46997	Uncommon Cold
Rowan Barrera	383749474	Male	31/05/1952	63570	Heaped Piles
Urielle Riley	795856737	Female	01/08/1985	08603	Uncommon Cold
Caesar Lancaster	995946734	Male	01/01/1986	93861	Uncommon Cold
Irene Curry	046498803	Male	09/04/1978	87454	Slack Tongue
Aline Hess	865009451	Female	05/06/1966	78956	Bloaty Head
Peter Calderon	336136140	Female	17/04/1987	60254	Bloaty Head
Hu Parrish	693587559	Male	03/02/1984	51213	Bloaty Head
Valentine Haynes	048717454	Female	10/04/1965	86362	Uncommon Cold
Amos Edwards	025759543	Male	13/07/1954	13197	The Squits

Identifying

Sensitive

Quasi Identifiers

- The database cannot be after removing the identifying attributes
- Latanya Sweeny was able to find the medical records of the governor of Massachusetts from a database that was released for research purposes and the voter list of Cambridge Massachusetts using his zip code, gender and birth date
- Philippe Golle showed in a research that over 60% of the population in the US are uniquely identifiable from these attributes

Quasi Identifiers

- Quasi identifiers are the set of attributes that are unique for a specific tuple and enable identification of the object a tuple corresponds to
- This definition assumes knowledge of the type of data the attacker will use (attributes of the database the attacker has)

k-Anonymity

- Each tuple is indistinguishable from at least *k*-1 other tuples with respect to the quasi identifiers
- Guarantees: quasi identifiers cannot be used to link data to less than *k* tuples

k-Anonymity - Example

Race	Birth Date	Gender	ZIP Code	Problem
Black	1965	Male	0214*	short breath
Black	1965	Male	0214*	chest pain
Black	1965	Female	0213*	hypertension
Black	1965	Female	0213*	hypertension
Black	1964	Female	0213*	obesity
Black	1964	Female	0213*	chest pain
White	1964	Male	0213*	chest pain
White	1964	Male	0213*	obesity
White	1964	Male	0213*	short breath
White	1967	Male	0213*	chest pain
White	1967	Male	0213*	chest pain

Attacks on *k*-Anonymity

Original (Private) Table

Race	Birth Date	Gender	ZIP Code	Problem
black	09/20/1965	male	02141	short breath
black	02/14/1965	male	02141	chest pain
black	10/23/1965	female	02138	painful eye
black	08/24/1965	female	02138	wheezing
black	11/07/1964	female	02138	obesity
black	12/01/1964	female	02138	chest pain
white	10/23/1964	male	02138	short breath
white	03/15/1965	female	02139	hypertension
white	08/13/1964	male	02139	obesity
white	05/05/1964	male	02139	fever
white	02/13/1967	male	02138	vomiting
white	03/21/1967	male	02138	back pain

Linked Table

Race	Birth Date	Gender	ZIP Code	Problem
black	1965	male	02141	short breath
black	1965	male	02141	chest pain
black	1965	female	02138	painful eye
black	1965	female	02138	wheezing
black	1964	female	02138	obesity
black	1964	female	02138	chest pain
white	1964	male	02138	short breath
white	1965	female	02139	hypertension
white	1964	male	02139	obesity
white	1964	male	02139	fever
white	1967	male	02138	vomiting
white	1967	male	02138	back pain

Released Table #1

Race	Birth Date	Gender	ZIP Code	Problem
black	1965	male	02141	short breath
black	1965	male	02141	chest pain
person	1965	female	0213*	painful eye
person	1965	female	0213*	wheezing
black	1964	female	02138	obesity
black	1964	female	02138	chest pain
white	1964	male	0213*	short breath
person	1965	female	0213*	hypertension
white	1964	male	0213*	obesity
white	1964	male	0213*	fever
white	1967	male	02138	vomiting
white	1967	male	02138	back pain

Released Table #2

Race	Birth Date	Gender	ZIP Code	Problem
black	1965	male	02141	short breath
black	1965	male	02141	chest pain
black	1965	female	02138	painful eye
black	1965	female	02138	wheezing
black	1964	female	02138	obesity
black	1964	female	02138	chest pain
white	196*	male	02138	short breath
white	196*	human	02139	hypertension
white	196*	human	02139	obesity
white	196*	human	02139	fever
white	196*	male	02138	vomiting
white	196*	male	02138	back pain

How to Break Anonymity of the Netflix Prize Dataset

Arvind Narayanan
Vitaly Shmatikov

Netflix Prize

- Netflix, the largest online movie rental service, announced $1,000,000 prize for improving their movie recommendation service
- A database, consisting of 100,480,507 movie ratings (on a 1 to 5 scale, with dates dates the ratings were entered) created by 480,189 subsribers

Does Privacy of Ratings Matter?

- Movie rating can leak sensitive information
- Future privacy

Definitions

V – set of all movies
C – set of all subscribers
$\forall\, m \in N, c \in C$
$r_c(m)$ – c's rating of movie m
$d_c(m)$ – date rating entered
For a fixed c and some $M \subseteq V, \epsilon_d, \epsilon_r, \delta_d, \delta_r$ the attacker knows $\hat{r}_c(m), \hat{d}_c(m)\ \forall\, m \in M$ such that
$\mathbf{P}_{m \in M}(|r_c(m) - \hat{r}_c(m)| \leqslant \epsilon_r) \geqslant 1 - \delta_r$
$\mathbf{P}_{m \in M}(|d_c(m) - \hat{d}_c(m)| \leqslant \epsilon_d) \geqslant 1 - \delta_d$

Definitions

Neighborhood of c (with respect to M)

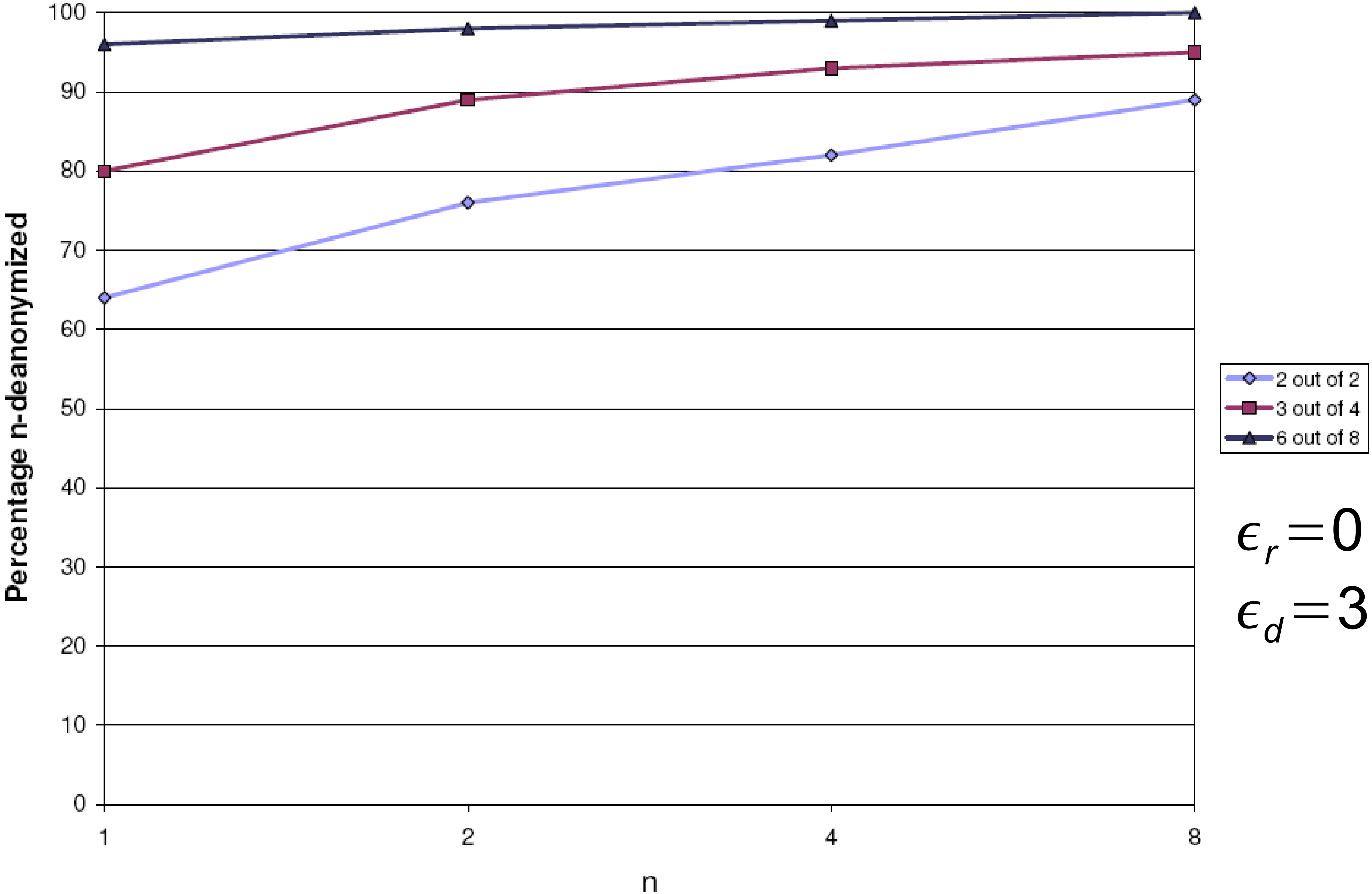
$$N_M(c) := \{c' \in C : \mathbf{P}_{m\in M}(|r_c(m) - \hat{r}_c(m)| \leqslant \epsilon_r) \geqslant 1 - \delta_r \wedge \\ \mathbf{P}_{m\in M}(|d_c(m) - \hat{d}_c(m)| \leqslant \epsilon_d) \geqslant 1 - \delta_d\}$$

$$n_M(c) := |N_M(c)|$$

M is uniformly chosen from c's rated movies and $|M| = k$, possibly with some restriction (not in the top 100 or 500 most rated movies)

$$\mu(n, k) = \mathbf{P}_{c\in C}(n_M(c) \leqslant n)$$

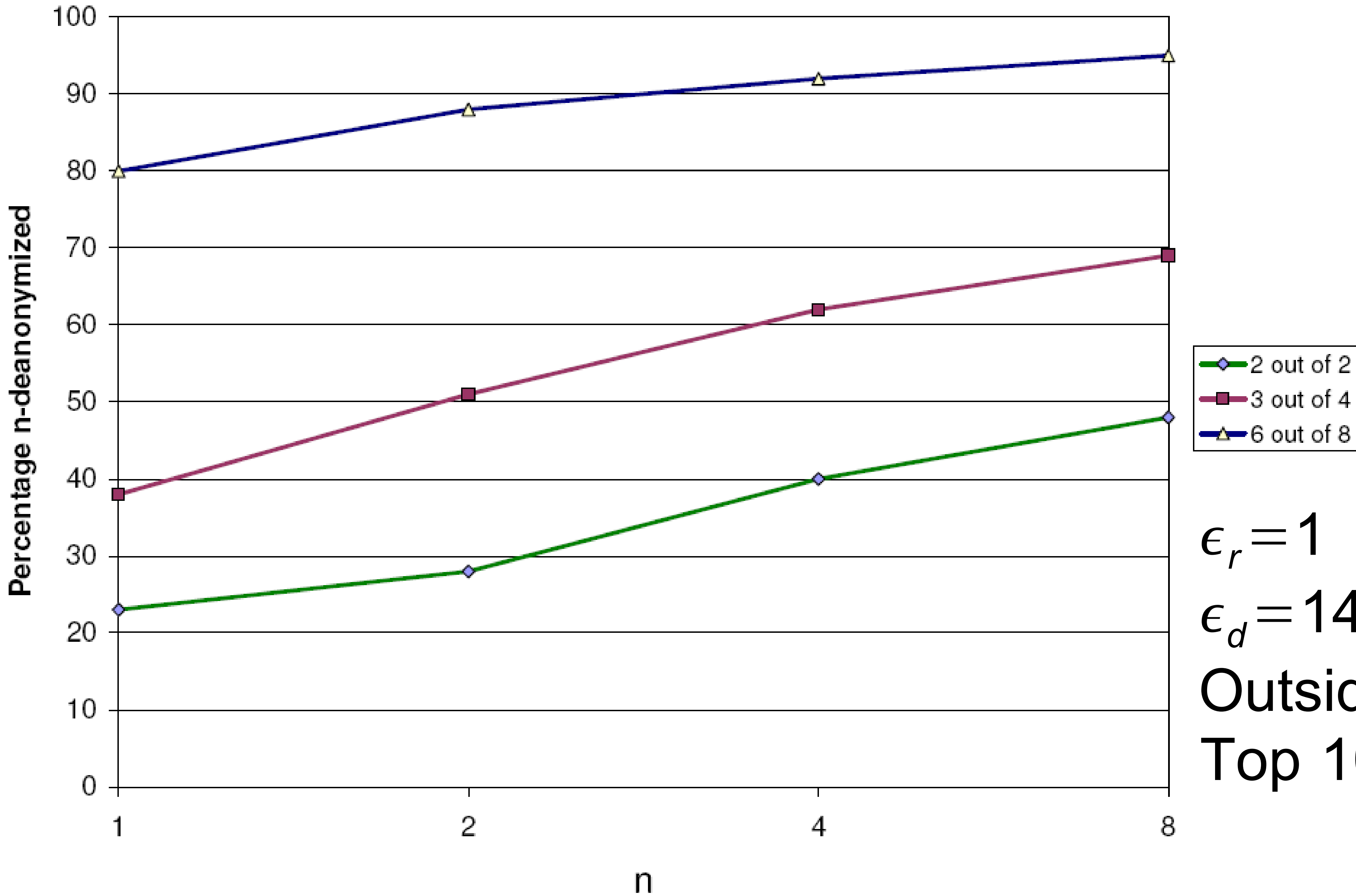
Results



$\epsilon_r = 0$

$\epsilon_d = 3$

Results

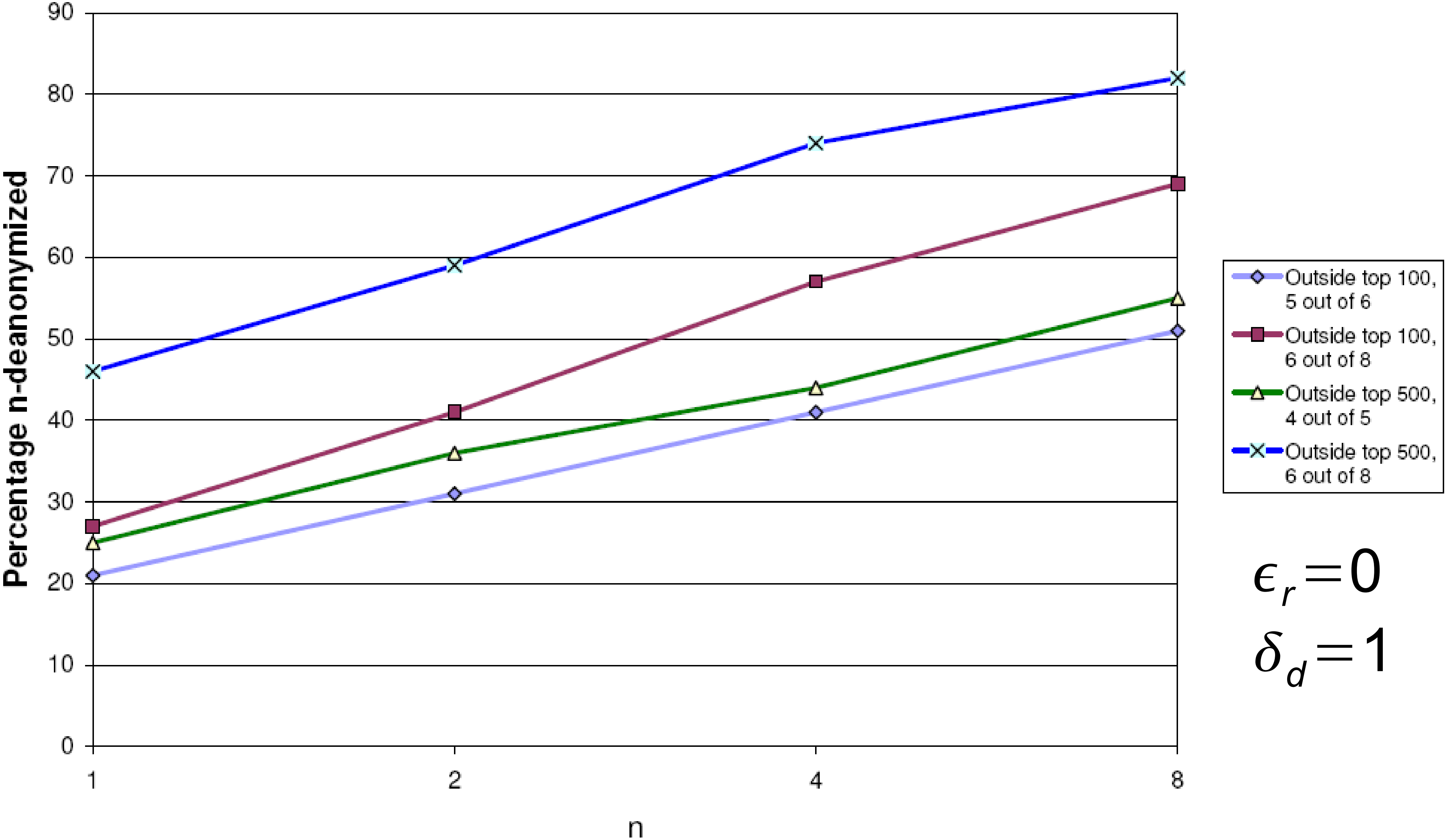


$\epsilon_r = 1$

$\epsilon_d = 14$

Outside Top 100

Results



$\epsilon_r = 0$

$\delta_d = 1$

De-anonymization using IMDb

- Movies that do not appear in both databases
- Users may enter only a comment, so rating might be missing
- Dates might not be correlated
- However, entire user's data in IMDb is available

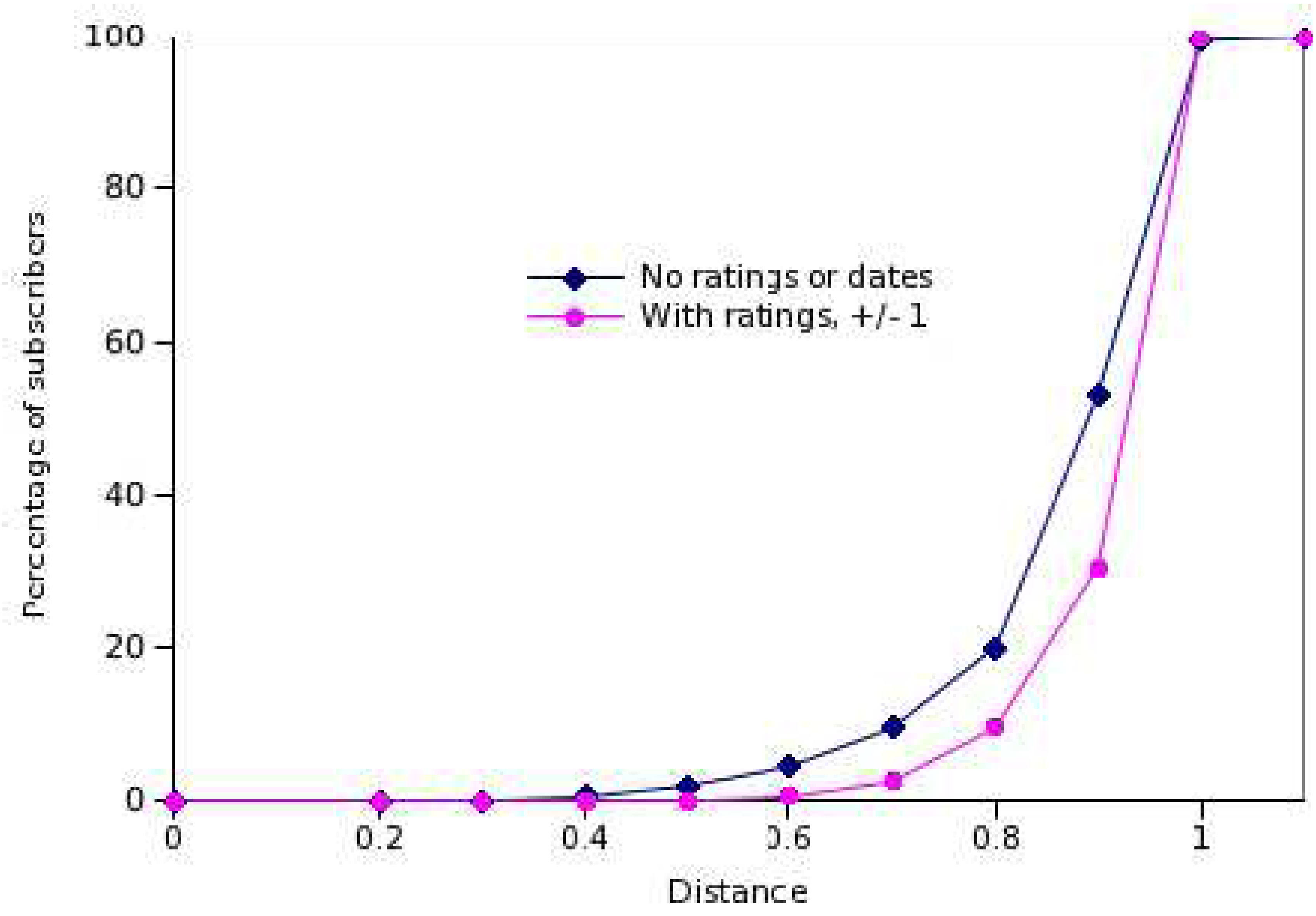
De-anonymization using IMDb

- The researchers manually extracted a few dozen IMDb users' records, defined a distance function and tried to match Netflix records
- With high confidence, two records were found to belong to Netflix subscribers and non public (and possibly sensitive) data was found from one of them

De-anonymization using IMDb

- Political views from ratings of “Power and Terror: Noam Chomsky in Our Times” and “Fahrenheit 9/11”
- Religious views from ratings of “Jesus of Nazareth” and “the Gospel of John”
- Sexual preferences from ratings of “Bent” and “Queer as Folk”

Is *k*-Anonymity Possible?



l-Diversity: Privacy Beyond *k*-Anonymity

Ashwin Machanavajjhala
Johannes Gehrke
Daniel Kifer
Muthuramakrishnan
Venkitasubramaniam

Is *k*-Anonymity Enough?

Zip Code	Age	Nationality	Condition
130**	<30	$*$	Heart Disease
130**	<30	$*$	Heart Disease
130**	<30	$*$	Viral Infection
130**	<30	$*$	Viral Infection
1485*	≥40	$*$	Cancer
1485*	≥40	$*$	Heart Disease
1485*	≥40	$*$	Viral Infection
1485*	≥40	$*$	Viral Infection
130**	3$*$	$*$	Cancer
130**	3$*$	$*$	Cancer
130**	3$*$	$*$	Cancer
130**	3$*$	$*$	Cancer

Is *k*-Anonymity Enough?

- *k*-anonymity does not protect linking sensitive attributes to a tuple
- *k*-anonymity does not provide any protection against background knowledge
- Simply outputting every record *k* times will satisfy *k*-anonymity

Bayes-Optimal Privacy

Q - single quasi-identifier
S - single sensitive attribute
T - private table
T^* - released table
$t[S]$, $t[Q]$ – victim's sensitive attribute and quasi-identifier in the private table
Attacker knows q, the victim's quasi-identifier, q^*, the generalized value of q and f, distribution of sensitive values according to quasi-identifiers

Bayes-Optimal Privacy

Attacker's prior belief that victim's sensitive attribute is s

$\alpha_{q,s} = P_f(t[S] = s \mid t[Q] = q)$

Attacker's posterior belief that victim's sensitive attribute is s

$\beta_{q,s,\ T^*} = P_f(t[S] = s \mid t[Q] = q,\ T^*)$

$$|\beta_{q,s,\ T^*} - \alpha_{q,s}| < \varepsilon$$

Bayes-Optimal Privacy

- We don't know what *f* the attacker knows
- We might not even know *f*
- We don't know attacker's data not modeled in *q*
- Still, it is possible to limit the attacker's belief that a tuple is associated with a certain sensitive attribute

Towards *l*-Diversity

$n_{(q^*,s)}$ – number of tuples in T^* with quasi–identifier q^* and sensitive attribute s

$$\beta_{q,s,T^*} = \frac{n_{(q^*,s)} \frac{f(s|q)}{f(s|q^*)}}{\sum_{s' \in S} n_{(q^*,s')} \frac{f(s'|q)}{f(s'|q^*)}}$$

$$\beta_{q,s,T^*} \approx 1 \quad \text{iff} \quad \forall s' \neq s \quad n_{(q^*,s)} \frac{f(s|q)}{f(s|q^*)} \gg n_{(q^*,s')} \frac{f(s'|q)}{f(s'|q^*)}$$

Towards *l*-Diversity

- It is only possible to change the number of tuples associated with a certain sensitive attribute
- Assumption: the attacker have less than *l*-1 “pieces of information”
- “A piece of information” - victim *X* does not have sensitive value *s*

Towards *l*-Diversity

- No single value can appear too frequently
- Attacker should not be able to dismiss sensitive values such that a single value will appear too frequently
- A table is *l*-diverse if every *q**-block contains at least *l* “well represented” sensitive values

Entropy l-Diversity

A table is entropy l-diverse if for every q^* block

$$-\sum_{s \in S} p_{q^*,s} \log(p_{q^*,s}) \geqslant \log(l)$$

where $p_{q^*,s} = \dfrac{n_{(q^*,s)}}{\sum_{s' \in S} n_{(q^*,s')}}$ is the fraction of tuples with sensitive attribute s

Entropy *l*-Diversity

Zip Code	Age	Nationality	Condition
1305*	≤40	*	Heart Disease
1305*	≤40	*	Viral Infection
1305*	≤40	*	Cancer
1305*	≤40	*	Cancer
1485*	>40	*	Cancer
1485*	>40	*	Heart Disease
1485*	>40	*	Viral Infection
1485*	>40	*	Viral Infection
1306*	≤40	*	Heart Disease
1306*	≤40	*	Viral Infection
1306*	≤40	*	Cancer
1306*	≤40	*	Heart Disease

l-Diversity

- The entropy of the entire table limits the entropy of each block (90% of the patients in a certain hospital have a heart problem) - Recursive (c, l)-Diversity
- It might be possible to disclose some sensitve values (healthy) - Positive Disclosure-Recursive (c, l)-Diversity
- A privacy breach might occur if attacker knows an object does not have a certain sensitive attribute (99.9% are not infected with a virus) - NPD-Recursive (c_1, c_2, l)-Diversity

Recursive (c, l)-Diversity

Let s_1, s_2, ... , s_m be the sensitive values appearing in a q^*-block and r_1, r_2, ... , r_m be the number each sensitive value appears. WLOG $r_i \geq r_{i+1}$.

The q^*-block is recursive (c, l)-diverse if

$$r_1 < c\,(r_l + r_{l+1} + \ldots + r_m)$$

A table is recursive (c, l)-diverse if every block is recursive (c, l)-diverse.

Positive Disclosure-Recursive (c, l)-Diversity

Let Y denote the set of sensitive attributes for which positive disclosure is allowed, y the minimal value for which s_y is not in Y. A q^*-block is Positive Disclosure-Recursive (c, l)-Diverse if

$$y \leq l\text{-}1 \text{ and } r_y < c\,(r_l + r_{l+1} + \ldots + r_m)$$

or

$$y > l\text{-}1 \text{ and } r_y < c\,(r_{l-1} + \ldots + r_{y-1} + r_{y+1} + \ldots + r_m)$$

Negative/Positive Disclosure-Recursive (c_1, c_2, l)-Diversity

Let W denote the set of sensitive attributes for which negative disclosure is not allowed, A table is Negative/Positive Disclosure-Recursive (c_1, c_2, l)-Diverse if it is Positive Disclosure-Recursive (c_1, l)-diverse and each s in W occurs in at least c_2 percent of the tuples in every q^*-block.

Multiple Sensitive Attributes Considerations

Zip Code	Age	Salary	Condition
4760*	2*	10K	Gastric Ulcer
4760*	2*	4K	Gastritis
4760*	2*	10K	Stomach Cancer
4790*	2*	6K	Gastritis
4790*	≥40	11K	Flu
4790*	≥40	8K	Bronchitis
4790*	≥40	7K	Bronchitis
4790*	≥40	11K	Pneumonia

t-Closeness: Privacy Beyond *k*-Anonymity and *l*-Diversity

Ninghui Li
Tiancheng Li
Suresh Venkatasubramanian

Is *l*-Diversity Enough?

Zip Code	Age	Salary	Condition
476**	2*	3K	Gastric Ulcer
476**	2*	4K	Gastritis
476**	2*	5K	Stomach Cancer
4790*	≥40	6K	Gastritis
4790*	≥40	11K	Flu
4790*	≥40	8K	Bronchitis
476**	3*	7K	Bronchitis
476**	3*	9K	Pneumonia
476**	3*	10K	Stomach Cancer

Is *l*-Diversity Enough?

- l-diversity does not protect against learning some semantic category of the sensitive value

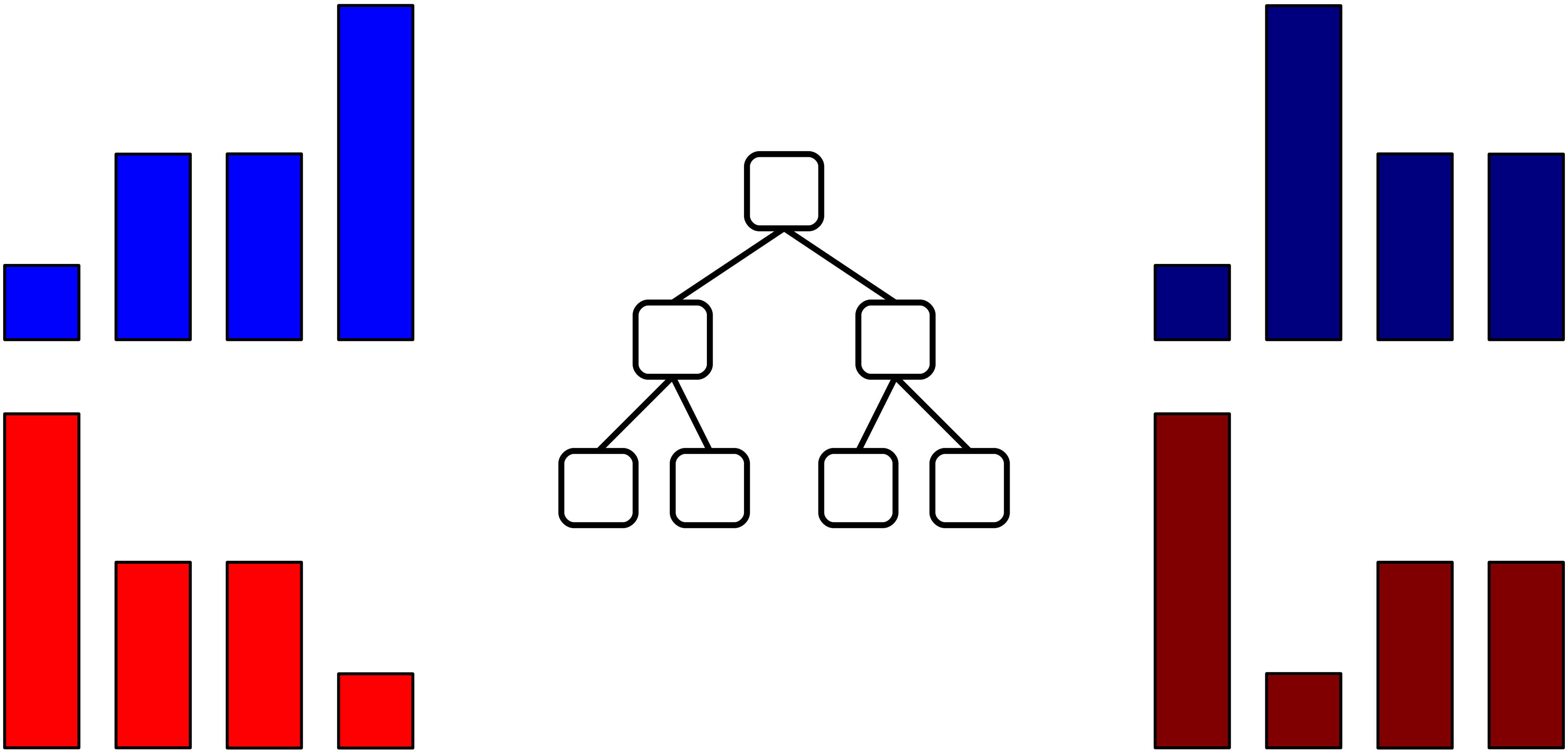
t-Closeness

- The attacker must learn the distribution of the sensitive values in the published database
- If nothing more can be learned – no privacy breach (unless the database will not be released)
- *t*-Closeness: the distribution of the sensitive values of every q^*-block is t close (w.r.t. some distance) to the distribution of the sensitive values of the entire database

Earth Mover Distance

- In the paper, the earth mover distance is used to capture semantic closeness
- How much and how far a “mass” of a distribution needs to be moved to be equal to another distribution

Earth Mover Distance



t-Closeness

Zip Code	Age	Salary	Condition
4767*	<40	3K	Gastric Ulcer
4767*	<40	5K	Stomach Cancer
4767*	<40	9K	Pneumonia
4790*	≥40	6K	Gastritis
4790*	≥40	11K	Flu
4790*	≥40	8K	Bronchitis
4760*	<40	4K	Gastritis
4760*	<40	7K	Bronchitis
4760*	<40	10K	Stomach Cancer

0.167-closeness w.r.t. Salary

0.278-closeness w.r.t. Disease

t-Closeness

- t-closeness limits the amount of useful information that can be derived from the database
- t-closeness only captures a certain semantic difference, an attacker might be interested in a completely different semantic categories

Information Disclosure Under Realistic Assumptions: Privacy Versus Optimality

Lei Zhang
Sushil Jajodia
Alexander Brodsky

Attack on an Algorithm

Private Database

Name	Age	Gender	Condition
Alan	Old	Male	Heart Disease
Bob	Old	Male	Viral Infection
Clark	Middle	Male	Cancer
Diana	Middle	Female	Cancer
Ellen	Young	Female	Flu
Fen	Young	Female	Ulcer

(Age, *)

Name	Age	Gender	Condition
Alan	Old	*	Heart Disease
Bob	Old	*	Viral Infection
Clark	Middle	*	Cancer
Diana	Middle	*	Cancer
Ellen	Young	*	Flu
Fen	Young	*	Ulcer

(*, Gender)

Name	Age	Gender	Condition
Alan	*	Male	Heart Disease
Bob	*	Male	Viral Infection
Clark	*	Male	Cancer
Diana	*	Female	Cancer
Ellen	*	Female	Flu
Fen	*	Female	Ulcer

(*, *)

Name	Age	Gender	Condition
Alan	*	*	Heart Disease
Bob	*	*	Viral Infection
Clark	*	*	Cancer
Diana	*	*	Cancer
Ellen	*	*	Flu
Fen	*	*	Ulcer

Attacker Point of View

"Public" Data

Name	Age	Gender
Alan	Old	Male
Bob	Old	Male
Clark	Middle	Male
Diana	Middle	Female
Ellen	Young	Female
Fen	Young	Female

Released Database

Age	Gender	Condition
*	Male	Heart Disease
*	Male	Viral Infection
*	Male	Cancer
*	Female	Cancer
*	Female	Flu
*	Female	Ulcer

Not 2-Diverse

Name	Age	Gender	Condition
Alan	Old	Male	?
Bob	Old	Male	?
Clark	Middle	Male	?
Diana	Middle	Female	?
Ellen	Young	Female	?
Fen	Young	Female	?

Name	Age	Gender	Condition
Alan	Old	*	?
Bob	Old	*	?
Clark	Middle	*	?
Diana	Middle	*	?
Ellen	Young	*	?
Fen	Young	*	?

Conclusions

- All privacy preserving schemes presented assume knowledge of attacker auxiliary information (quasi identifiers, attributes that can be dismissed) or the data the attacker is interested in (semantic catgory)
- As shown by the algorithm attack, all schemes presented implicitly assume the method the attacker will use