Sécurité différentielle dans les Bases de Données et Complexité

Richard Lassaigne
IMJ/Logique mathématique
CNRS-Université Paris Diderot

Quelques acteurs majeurs

Cynthia Dwork Moritz Hardt

Christos Papadimitriou

- Gouvernement, entreprises, centres de recherche collectent des informations **personnelles** et les **analysent**
- **Réseaux sociaux** : Facebook, LinkedIn
- YouTube et Amazone utilisent les enregistrements d'achâts et de vidéos
- Les courriers dans **Gmail** sont utilisés pour des publicités ciblées

Mesures conventionnelles :

- contrôler l'**accès** à l'information
- contrôler le **flôt** d'information
- contrôler l'**utilisation** de l'information

Approches classiques de la communication privée :

- **anonymisation** (suppression des identifiants, k-anonymisation)
- **assainissement** (communication d'un échantillon)

Ces approches ne garantissent pas vraiment le **respect** de la vie privée

- Parsimonie des données : avec grande probabilité,
  jamais deux profils ne sont **similaires** à plus de α
  Exemple : Internet Movie Database $\alpha = 50\%$

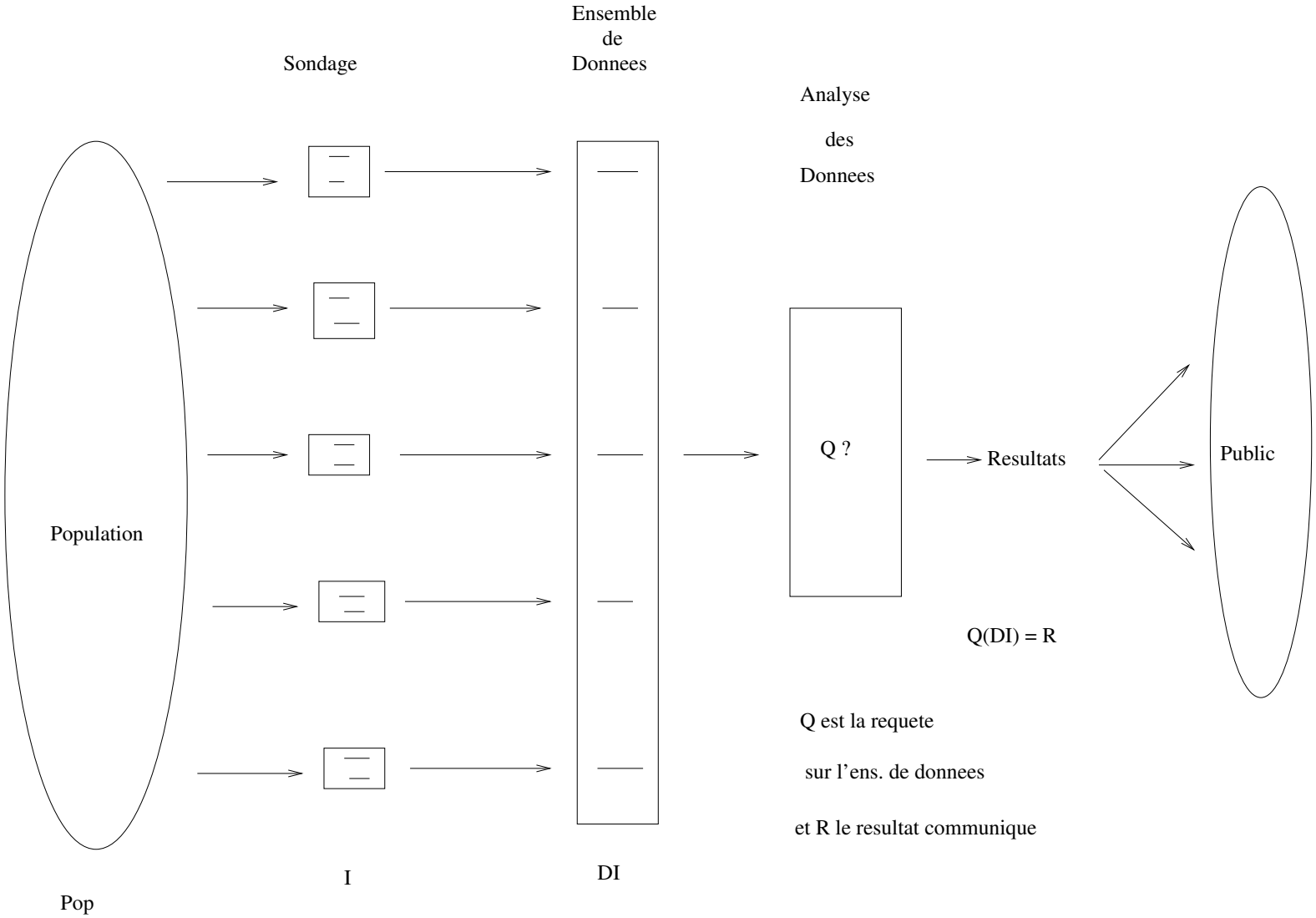
- Si le profil peut être aparié à 50% de **similarité**
  à un profil dans IMDB, alors l'adversaire connaît
  avec grande probabilité la vraie identité du profil

- Le papier suivant propose un
  algorithme probabiliste efficace
  pour casser l' **anonymisation** dans de telles BD

  A. Narayanan et V. Shmatikov :
  Robust de-anonymization of large sparse datasets
  Proc. 29th IEEE Symposium on Security and Privacy, 2008

- Aimez-vous écouter Carla Bruni ?
- **Combien** d'albums de Carla Bruni possédez-vous ?
- Quel est votre sexe ?
- Quel est votre âge ?

Si votre goût pour la musique est une **information sensible**, quel **mécanisme** de réponse à cette requête pourrait vous inciter à répondre ? **Anonymisation** ?


Ensemble
de
Donnees
Sondage
Analyse
des
Donnees
Population
Q ?
Resultats
Public
Q(DI) = R
Q est la requete
sur l'ens. de donnees
et R le resultat communique
Pop
I
DI

- Une **réponse personnelle** n'a pas d'impact sur le résultat communiqué :

$$Q(D_{I-i}) = Q(D_I)$$

- Un **attaquant** consultant le résultat publié ne peut apprendre (**avec grande probabilité**) aucune nouvelle information personnelle :

$$Prob[\text{secret}(i) \mid R] = Prob[\text{secret}(i)]$$

- Si les **réponses individuelles** n'ont aucun impact sur le résultat communiqué, alors le résultat pourrait n'avoir aucune utilité. Par induction :
  $Q(D_{I-i}) = Q(D_I) \Longrightarrow Q(D_I) = Q(D_\emptyset)$

- Si le résultat montre qu'il existe une **forte tendance** dans la population, alors, **avec grande probabilité**, la propriété correspondante est vraie, pour un enregistrement donné : $Prob[\text{secret}(i) \mid \text{secret(Pop)}] > Prob[\text{secret}(i)]$

- Si un **attaquant** connaît une **fonction** dépendant de faits généraux et fournissant des renseignements personnels, alors en communiquant juste ces faits généraux, on permet à l'attaquant de déduire une **information spécifique**.

La personne répondant à une enquête
se sentira plus en **sécurité** si elle sait que

La **probabilité** que le **résultat communiqué**
soit R est presque la même, **indépendamment**
du fait qu'elle ait répondu ou non

Respect **différentiel** de la vie privée

Sensibilité d'une fonction

Mécanisme de **Laplace**

Divergence max, divergence en moyenne

Mécanisme **exponentiel**

Complexité

[Dwork, Mc Sherry, Nisssim et Smith, 2006]

Deux ensembles de données D, D' sont **voisins** s'ils ne diffèrent que par un seul élément

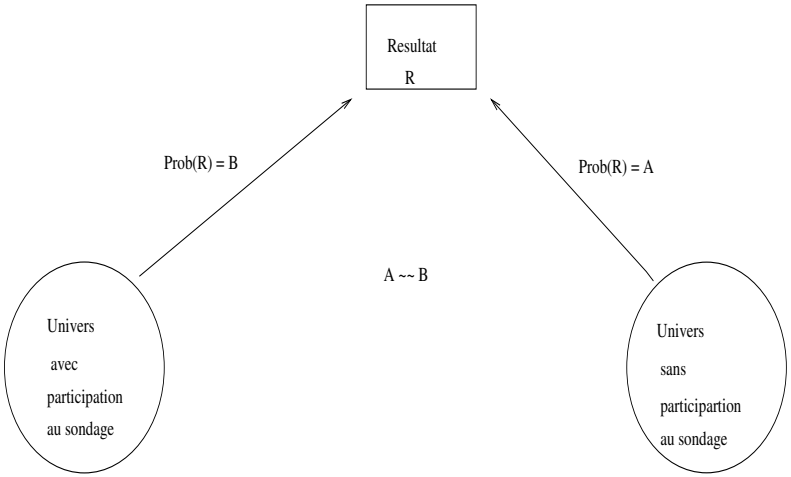
Définition :
Un algorithme probabiliste M est ε-**différentiellement privé** si pour tous ensembles de données D, D' **voisins** et tout évènement S

$$Prob[M(D) \in S] \leq \exp(\varepsilon).Prob[M(D') \in S]$$

La **probabilité** que le résultat communiqué soit R est presque la même, indépendamment du fait que l'on fournit ou non une **information personnelle**

$\frac{Prob[R|\text{univers}=D_I]}{Prob[R|\text{univers}=D_{I\pm i}]} \leq \exp(\varepsilon)$ pour tous I, i, R

Etant donné un **résultat** R, comment un **attaquant** peut-il deviner de quel **univers** il provient ?



- Les résultats publiés donnent une **indication minimale** sur le fait qu'un individu donné ait contribué à la BD
- Si les individus fournissent seulement des informations sur eux-mêmes, alors il y a **protection relative** du caractère d'identification de l'information
- Résistance aux **attaques par lien** avec d'autres BD's (résistance à la puissance d'une information **auxiliaire**)
- Résistance aux **attaques par composition** de requêtes

Soit f une fonction définie sur D, à valeurs réelles
Exemple : Requêtes **statistiques** ou de **comptage**

$$\Delta(f) = max_{x,x' \text{voisins}} \mid f(x) - f(x') \mid$$

- Les ens. de données **voisins** diffèrent par au plus un enregistrement
- Les requêtes de **comptage** ont une sensibilité de 1
- La **sensibilité** reflète combien les données d'un individu peuvent affecter la sortie

- Combien de personnes sondées sont des femmes ?

$$\Delta(f) = 1$$

- Au total, combien d'albums différents de Carla Bruni ont été achetés par des personnes sondées ?

$$\Delta(f) = 3$$

(puisqu'elle a enregistré 3 albums différents ?)

- Requête du type :

"**Combien** d'enregistrements dans la BD
satisfont la **propriété** P"
(P prédicat **booléen** sur l'univers D)

- Algorithme :
  Calculer la **vraie réponse** à la requête $q(D)$
  Ajouter un **bruit aléatoire** suivant une distribution
  satisfaisant

$$\forall z, z' \mid z - z' \mid \leq 1 \Rightarrow Prob[z] \leq \exp(\varepsilon).Prob[z']$$

Distribution de **Laplace**

Paramètre b

$$p(z) = \frac{1}{2b} \exp(-\frac{\mid z \mid}{b})$$

Variance $2b^2$

Accroître b applatit la courbe

$b = \frac{1}{\varepsilon}$ La **densité** en z est proportionnelle à $\exp(-\varepsilon \mid z \mid)$

Sensibilité de f :

$$\Delta(f) = max_{x,x' \text{voisins}} \mid f(x) - f(x') \mid$$

Théorème (DMNS06) :
Pour une requête $f : D \mapsto \mathbb{R}$, le mécanisme qui ajoute un bruit engendré suivant la **distribution** $Lap(b)$ avec $b = \frac{\Delta(f)}{\varepsilon}$ satisfait la propriété de respect ε-**différentiel** de la vie privée.

Remarques :
- Le **bruit** dépend de f et de ε, non de la BD
- Une **sensibilité** plus petite correspond à une distorsion moindre

$$\frac{Prob[f(x) + Lap(\frac{\Delta(f)}{\varepsilon}) = y]}{Prob[f(x') + Lap(\frac{\Delta(f)}{\varepsilon}) = y]} =$$

$$\exp(-\frac{\mid y - f(x) \mid}{\Delta(f)}\varepsilon) / \exp(-\frac{\mid y - f(x') \mid}{\Delta(f)}\varepsilon) =$$

$$\exp(\frac{\varepsilon}{\Delta(f)}(\mid y - f(x') \mid - \mid y - f(x) \mid))$$

$$\leq \exp(\frac{\varepsilon}{\Delta(f)}(\mid f(x) \mid - \mid f(x') \mid)) \leq \exp(\varepsilon)$$

Combien de personnes dans la BD sont des femmes ?

- **Sensibilité** $= 1$
- Il suffit d'ajouter un **bruit** distribué suivant $Lap(\frac{1}{\varepsilon})$

  Problème : Requête de comptage à **valeurs multiples**

$$f : D \longrightarrow \mathbb{R}^d$$

$$\Delta(f) = max_{x,x' \text{voisins}} \mid\mid f(x) - f(x') \mid\mid_1$$

$$\Delta(f) = max_{x,x' \text{voisins}} \sum_{i=1}^{d} \mid f(x)_i - f(x')_i \mid$$

Théorème (DMNS06) :
Pour une requête $f : D \longrightarrow \mathbb{R}^d$, le mécanisme qui ajoute un bruit engendré suivant la **distribution** $[Lap(\frac{\Delta(f)}{\varepsilon})]^d$ satisfait la propriété de respect ε-**différentiel** de la vie privée.

- Suite de requêtes $f_1, f_2, \ldots, f_m$
- Mécanisme avec **distribution** du bruit $Lap(\sum_{i=1}^{m} \frac{\Delta(f_i)}{\varepsilon})$
- Conservation de la propriété de respect ε-**différentiel**
- Même si les requêtes sont "**adaptatives**"

Paramètre de **divergence**

Propriété (DP) :

Pour tous D, D' voisins, pour tout $S \subseteq \text{domaine}(M)$

$$\frac{Prob[M(D) \in S]}{Prob[M(D') \in S]} \leq \exp(\varepsilon)$$

Le paramètre ε mesure la **perte** :

$$ln(\frac{Prob[M(D) \in S]}{Prob[M(D') \in S]}) \leq \varepsilon$$

Du **pire** des cas au cas **moyen**

- **Divergence max** :

$$D_\infty(Y \mid\mid Z) = max_{S \subseteq \text{supp}(Y)} ln(\frac{Prob[Y \in S]}{Prob[Z \in S]})$$

- **Divergence** KL (**en moyenne**) :

$$D(Y \mid\mid Z) = E_{y \sim Y} ln(\frac{Prob[Y = y]}{Prob[Z = y]})$$

- Un lemme utile fournit une borne sur la **divergence en moyenne** ([DRV10]) :

si $D_\infty(Y \mid\mid Z) \leq \varepsilon$ alors

$D(Y \mid\mid Z) \leq 2\varepsilon^2$ (lorsque $\varepsilon < 1$)

Lorsque l'addition de **bruit** n'a pas de sens

- Exemples :
  Requêtes à valeurs dans les **chaînes de caractères**,
  les arbres, les stratégies,...
  Requêtes de choix d'un meilleur **objet** dans un ensemble,
  non nécessairement continu, d'objets à valeurs réelles

- On suppose l'existence d'une **fonction d'utilité**
  en (D, y) qui mesure la **qualité** d'une réponse y
  sur l'ensemble de données D

- Exemple : D ensemble de points étiquetés dans $\mathbb{R}^d$
  y vecteur de $\mathbb{R}^d$ décrivant un **hyperplan**
  de dimension $d-1$ cherchant à classer les points
  u fonction donnant le nombre de points
  correctement classés

- Fournit y avec **probabilité** proportionnelle à $\exp(\frac{u(D,y)\varepsilon}{\Delta(u)})$

- $\Delta(u)$ **sensibilité** de la fonction d'utilité u bornant pour tous ensembles de données D, D' **adjacents** la différence $\mid u(D,y) - u(D',y) \mid$

  Assure la propriété (DP) de **respect différentiel** de la vie privée

- Problème de la génération de données synthétiques (GDS) :
  Etant donné un ens. de données D, produire un ens.
  dont les statistiques reflètent fidèlement celles de D
  et qui préserve la propriété de respect différentiel

- Résultat : sous des hypothèses cryptographiques standard,
  il existe des instances du problème GDS qui n'ont pas
  d'implémentation en temps polynomial

- Conséquence : existence de cas où
  le mécanisme exponentiel n'a pas
  d'implémentation efficace

- Fondement **mathématique** pour l'analyse de données **privées**
- Mécanisme **différentiel** pour les **requêtes de comptage** et les **requêtes statistiques**
- Préservation du caractère **différentiel** par **composition**
- **Extensions** à d'autres types de requêtes
- Problème de l'**efficacité** des extensions (mécanisme **exponentiel**)

• [AW89] N.R. Adam and J. Wortmann. *Security-control methods for statistical databases : A comparative study*. ACM Computing Surveys, 21, p.515-556, 1989.

• [CMNS06] C. Dwork, Frank McSherry, K.Nissim and A. Smith. *Calibrating noise to sensitivity in private data analysis*. 3rd Theory of Cryptography Conference, 2006.

• [DNRRV09] C. Dwork, M. Naor, O. Reingold, G. Rothblum and S. Vadhan. *On the complexity of differentially data relase*. International ACM Symposium on Theory of Computing, p.381-390, 2009.

• [D11] C. Dwork. *A Firm Foundation for Private Data Analysis*. Communications of the ACM 54(1), 2011.

- [HT09] M. Hardt and K. Talwar. *On the geometry of differential privacy.* arXiv :0907.3754v2, 2009.

- [KPR00] J. Kleinberg, C. Papadimitriou and P. Raghavan. *Auditing boolean attributes*. Proc. 19th ACM Symposium on Principles of Database Systems, p. 86-91, 2000.

- [MT07] Frank McSherry and K. Talwar. *Mecanism design via differential privacy.* Proc. 48th Annual Symposium on Foundations of computer Science, 2007.

- [NS08] A. Narayanan et V. Shmatikov. *Robust de-anonymization of large sparse datasets* Proc. 29th IEEE Symposium on Security and Privacy, 2008

Merci pour votre attention