

Sécurité Appliquée

Protection de la vie privée-PVP. TD3

Jean-François COUCHOT
couchot@arobase.femto-st.fr

29 novembre 2021

Sx	Zip Code	Age	Disease	Salary
M	400071	35	bronchitis	10k
M	400182	37	pneumonia	11k
M	400095	39	stomach cancer	12k
F	440672	54	gastritis	12k
F	440123	58	Flu	15k
M	440893	54	bronchitis	16k
M	400022	41	gastric ulcer	16k
M	400135	46	gastritis	17k
F	400182	44	stomach cancer	18k

TABLE 1 – Données médicales sensibles

1 Mise en place de mécanismes sur une base

Dans cette section, on reprend le tableau 1 issu de l'article ¹

Exercice 1.1 (Sensibilités de requêtes). Pour chacune des requêtes suivantes, évaluer sa sensibilité, ou la sensibilité de la fonction d'utilité associée.

- Q_1 : « nombre de patients de moins de 40 ans atteints d'un cancer » ;
- Q_2 : « Répartition des âges (histogramme/camembert) de l'ensemble des patients ».
- Q_3 : « maladie la plus fréquente pour les patients souffrant d'un problème gastrique » ;
- Q_4 : « âge moyen des patients » ;
- Q_5 : « âge moyen des patients arrondi à l'entier le plus proche » ;

Exercice 1.2 (Choisir un mécanisme). Pour chacune des requêtes précédentes, choisir le mécanisme vérifiant la ϵ -DP le plus adapté. Le décrire.

Exercice 1.3 (Mécanisme géométrique pour valeurs entières). On considère le mécanisme suivant qui retourne toujours une valeur entière.

$$\mathcal{M}_G(D) = Q(D) + Gtq. \Pr[G = g] = \frac{1 - \alpha}{1 + \alpha} \alpha^{|g|}$$

avec $\alpha = \exp(-\epsilon/\Delta_Q)$, $g \in \mathbb{Z}$. Ce mécanisme est connu sous le nom de mécanisme géométrique ².

1. $D = (26, 36, 77, 157, 274, 610, 1082, 1517, 1807, 829)$ donne le nombre d'hospitalisés en Île de France ³ pour Covid-19 le 18-11-2020, par classe d'âge d'amplitude 10

1. Elabd, E., Abdulkader, H., & Mubark, A. (2015). L-diversity-based semantic anonymization for data publishing. *IJ Information Technology and Computer Science (IJITCS)*, 10, 1-7.

2. Ghosh, A., Roughgarden, T., & Sundararajan, M. (2012). Universally utility-maximizing privacy mechanisms. *SIAM Journal on Computing*, 41(6), 1673-1693.

3. www.data.gouv.fr/fr/datasets/donnees-hospitalieres-relatives-a-lepidemie-de-covid-19/

2. Soit la requête $Q(D)$ qui retourne le nombre de personnes hospitalisées qui ont moins de 30 ans dans D . Quelle est sa sensibilité ?
3. Remplir le tableau suivant en exploitant le mécanisme géométrique

r	...	134	135	136	137	138	139	140	...
$\Pr[Q = r]$									

2 Preuve de confidentialité différentielle de certains mécanismes

Exercice 2.1. Preuves théoriques. Pour chacun des mécanismes suivants, montrer qu'il vérifie la ϵ -DP.

1. Mécanisme laplacien.
2. Mécanisme exponentiel.
3. Mécanisme géométrique (vu à l'exercice précédent).