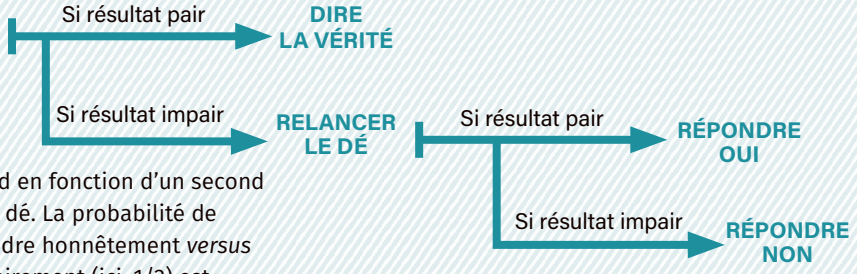
UN SONDAGE SENSIBLE

Question
Avez-vous déjà pris de la drogue ?

Lancer de dé



Comment répondre à un sondage sensible en dissimulant sa réponse de façon efficace et résistante aux attaques ? Grâce au mécanisme de réponse randomisé proposé par Stanley Warner en 1965. Demandons à 1000 individus s'ils ont déjà pris de la drogue. Selon le résultat d'un jet de dé (pair ou impair), ils disent la vérité ou bien répondent au hasard en fonction d'un second jet de dé. La probabilité de répondre honnêtement *versus* aléatoirement (ici, 1/2) est fonction du paramètre de vie privée ε. Résultat : 500 interrogés répondent au hasard, et parmi eux, 250 prétendent avoir pris de la drogue, sans lien avec la vérité. Si, sur les 1000 interrogés, 450 individus répondent « oui », seuls 200 font partie de la moitié (500 individus) qui dit la vérité. Sur l'ensemble, en supposant que le groupe des interrogés ayant répondu au hasard suit la même distribution que le groupe ayant répondu honnêtement, c'est donc à peu près le double (400) qui a déjà pris de la drogue.

Ce développement récent d'outils a accompagné la mise en place de la confidentialité différentielle au sein d'applications et de services de la vie quotidienne. Cependant, son utilisation par les acteurs du numérique est pour l'instant concentrée sur la collecte de statistiques ou bien la publication d'informations de nature statistique.

ET DANS LA VRAIE VIE ?

En 2016, lors de sa conférence annuelle WWDC, Apple introduisait l'utilisation de la confidentialité différentielle pour, entre autres, la collecte de statistiques sur l'utilisation des émojis dans les messages privés. Depuis, l'idée, auparavant cantonnée au milieu académique, s'est diffusée dans les médias qui, même spécialisés, touchent un plus grand public. Avant cela, en 2014, Google avait mis en place un outil nommé Rappor afin d'analyser l'utilisation de son navigateur Chrome. Grâce à cet algorithme, fondé sur un mécanisme proposé en 1965 par le statisticien Stanley Warner, un individu peut répondre à un sondage sensible en conservant la possibilité de nier de façon plausible sa réponse, tout en offrant des garanties sur la quantité de bruit une fois l'ensemble des réponses agrégées. Le mécanisme en jeu est le suivant : étant donné une question sensible dont la réponse est binaire (« oui » ou « non »), l'individu choisit aléatoirement, par exemple à l'aide d'un jet de dé, de répondre la vérité (avec une certaine probabilité qui dépend du bruit, c'est-à-dire au paramètre ε), ou de répondre « oui » ou « non » aléatoirement à l'aide d'un second jet de dé *(voir l'encadré ci-dessus)*. Les résultats individuels sont agrégés dans une structure probabiliste qui donne ensuite accès à des statistiques globales correctes tout en respectant la confidentialité différentielle.

Plus récemment, des jeux de données d'importance ont été anonymisés par des algorithmes satisfaisant la confidentialité différentielle. En 2020, le bureau du recensement de la population aux États-Unis (Census) l'a adoptée comme modèle de vie privée pour la publication de ses statistiques. En plus des nombreuses valeurs agrégées publiées, le Census distribue des données synthétiques générées aléatoirement à partir des distributions anonymisées. La confidentialité différentielle a aussi été utilisée pour publier le

© Eleven/shutterstock.com (dé) ; Pour la Science

Search trends symptoms dataset, un jeu de données anonymisées sur les recherches sur Google liées aux symptômes associés au Covid-19.

La recherche académique est aujourd'hui très active et de nombreux chantiers sont ouverts. Certaines pistes de recherche sont particulièrement prometteuses, notamment pour répondre à des questions importantes. D'abord, comment combiner équité et vie privée ? En effet, si la confidentialité différentielle est aujourd'hui une référence pour la protection de la vie privée lors de la publication de données, son impact sur l'équité n'est pas anodin. Son adoption par le Census a par exemple posé clairement ce problème. En effet, le bruit ajouté aux requêtes pour anonymiser les données du recensement peut impacter différemment certains groupes vis-à-vis d'autres. C'est notamment le cas des statistiques concernant les langues minoritaires, utilisées afin de décider dans quelles langues le matériel de vote doit être distribué. Le taux d'erreur, c'est-à-dire l'impact du bruit ajouté, varie entre les juridictions en fonction du nombre de locuteurs minoritaires recensés.

DANS LA TÊTE DE L'ATTAQUANT

Autre problématique, comment compléter les techniques de protection formelles comme la confidentialité différentielle avec des approches empiriques d'attaque ? L'objectif est d'évaluer l'efficacité d'une anonymisation au-delà des garanties théoriques du modèle. La tâche est cruciale, car un modèle de vie privée ne capture pas tout. Il est ici nécessaire de se mettre à la place d'un attaquant qui observerait des données anonymisées et tenterait d'en extraire une information sensible. De quelles connaissances auxiliaires dispose-t-il (partagées sur les réseaux sociaux, obtenues de façon illicite…) ? Avec quelles méthodes va-t-il exploiter les données anonymisées et les connaissances auxiliaires (réidentification, reconstruction ou inférence d'appartenance) ? Le niveau de protection assuré par une méthode d'anonymisation peut être estimé en comparant le succès d'attaques sur les données brutes et sur les données anonymisées. Cette approche empirique, jamais exhaustive, de gestion du risque est complémentaire des approches formelles, car elle guide les choix complexes de modèle de vie privée, d'algorithme d'anonymisation et de paramétrage (par exemple, le budget).

Enfin, dernier exemple : comment combiner chiffrement et perturbations aléatoires liées à la confidentialité différentielle, par exemple dans un contexte de calcul sécurisé (donc hors du domaine d'application initial de la confidentialité différentielle, c'est-à-dire l'anonymisation) ? Les techniques de protection habituelles sont parfois trop coûteuses ou trop rigides quand les données sont massives ou distribuées. En autorisant un dévoilement contrôlé d'informations au cours du calcul, les techniques de perturbation aléatoire satisfaisant la confidentialité différentielle peuvent compléter les techniques cryptographiques en accélérant les performances tout en conservant des garanties formelles. Le prix à payer est la perte en précision des calculs due aux perturbations aléatoires.

En pleine ébullition, les travaux sur la confidentialité différentielle et son amélioration nous promettent un avenir où notre vie privée sera garantie de façon plus sûre qu'elle ne l'était jusqu'à présent. De la sorte, nous aurons moins de réticences à confier nos données. À l'heure de l'open data, c'est une bonne nouvelle.

— Les auteurs —

> Tristan Allard
est maître de conférences à l'université de Rennes 1 et rattaché au laboratoire Irisa.

> Louis Béziaud
est doctorant en informatique à l'université Rennes 1 et à l'université du Québec à Montréal.

> Sébastien Gambs
est titulaire de la chaire de recherche du Canada en analyse respectueuse de la vie privée et éthique des données massives, et professeur au département d'informatique de l'université du Québec à Montréal.

— À lire —

S. Wagh *et al.*, DP-cryptography: Marrying differential privacy and cryptography in emerging applications, *Commun. ACM*, vol. 64 (2), pp. 84-93. 2021.

C. Sahin *et al.*, A differentially private index for range query processing in clouds, 2018 IEEE 34th international conference on data engineering (ICDE), pp. 857-868, 2018.

C. Dwork et A. Roth, The algorithmic foundations of differential privacy, *Found. Trends Theor. Comput. Sci.*, vol. 9, pp. 211-407, 2014.

M. Alaggan *et al.*, BLIP: Non-interactive differentially-private similarity computation on bloom filters, *SSS*, 2012.