Université de Paris

Mathématiques discrètes
Projet : Paradoxe des anniversaires

Consignes Le but du projet est de présenter une application dans laquelle les mathématiques discrètes jouent un rôle fondamental.

Le rendu final du projet consistera en un article destiné au grand public au format pdf de 800-1000 mots plus une annexe numérique, qui pourra contenir par exemple une démonstration interactive, une vidéo explicative et/ou des graphiques générés par du code écrit par vous-même ; cette annexe sera rendue sous la forme d'un lien vers un dépôt en ligne. La forme exacte et la technologie utilisée pour l'annexe peut varier et est donc laissé au libre choix des étudiants. L'article et son annexe seront jugés non seulement sur le contenu mais aussi sur la clarté de la présentation, la qualité de rédaction, et la créativité.

Contenu Le sujet détaille quelques points à développer mais ceux-ci sont seulement proposés comme point de départ de votre travail. Vous êtes encouragés à développer d'autres pistes en lien avec les mathématiques discrètes. De même, la bibliographie conseillée est un point de départ. Vous pouvez vous appuyer sur d'autres sources sur lesquelles vous porterez un œil critique et que vous prendrez soin de citer correctement.

Charte de bonne conduite Lisez attentivement la charte de bonne conduite. Portez une attention particulière à citer toutes vos sources, y compris les exemples et les images que vous utiliserez. L'utilisation d'outils d'IA tels que ChatGPT est formellement interdite. L'équipe pédagogique sera très attentive à tous ces aspects lors de la correction.

Calendrier Consultez la page Moodle du cours pour les dates des principales étapes du projet.

Bref descriptif du sujet

Le paradoxe des anniversaire est un résultat célèbre et surprenant de combinatoire qui dit que dans une classe de 23 élèves, il y a à peu près une chance sur deux que deux élèves aient le même anniversaire.

Le paradoxe des anniversaires est un concept clé dans l'étude des fonctions de hashage. Une fonction de hashage h est une fonction qui prend en entrée un fichier F, potentiellement très large, et renvoie une chaîne beaucoup plus courte $h(F)$. La propriété fondamentale que l'on souhaite avoir est que $h(F)$ permette de vérifier, avec très bonne probabilité, qu'un document est authentique et n'a pas été modifié durant un transfert. Cela permet par exemple de télécharger un fichier F, puis d'en calculer le hashé $h(F)$, et de le comparer à la valeur v indiquée sur le site de téléchargement. Si $h(F) = x$, alors il est très probable que le fichier téléchargé soit le bon.

Le paradoxe des anniversaires permet de quantifier la probabilité que des collisions entre fichiers F et F' apparaissent, c'est-à-dire que F et F' soient différents, mais que néanmoins,

$h(F) = h(F')$. Les fonctions de hashage étant au coeur de protocoles cryptographiques, le paradoxe des anniversaires a aussi de très nombreuses applications dans l'étude des système de chiffrement et des protocoles cryptographiques.

Bibliographie conseillée

- `https://fr.wikipedia.org/wiki/Paradoxe_des_anniversaires`
- `https://scienceetonnante.com/2012/05/28/le-paradoxe-des-anniversaires/`
- `http://www.bibmath.net/crypto/index.php?action=affiche&quoi=chasseur/anniversaire`

Pistes de développement

1. Établir proprement la formule donnant la probabilité que deux personnes aient le même anniversaire dans une classe de N élèves.
2. Écrire un programme qui simule des classes de N élèves, tel que l'anniversaire de chaque élève soit un jour pris au hasard dans l'année, afin de vérifier expérimentalement la formule obtenue.
3. Généraliser la formule et la simulation au paradoxe des anniversaires généralisé, pour des années ayant d jours pour $d \neq 365$.
4. Décrire une façon d'exploiter le paradoxe des anniversaires (généralisé) pour créer une attaque sur un protocole utilisant une fonction de hashage dont la sortie est très courte pour vérifier l'authenticité d'un fichier.
5. Explorer une autre application du paradoxe, et appliquer la simulation aux paramètres correspondant à l'application. Par exemple : l'identification de suspects par la technique d'empreinte génétique, ou l'application du paradoxe des anniversaires aux attaques sur le protocole de Wifi.