Figure: Shutterstock.com/NTL studio

Localisation par le réseau mobile

Quel anonymat ? | La plateforme Mobility Insights de Swisscom est un outil en ligne présentant des statistiques de groupes anonymes basées sur les données mobiles des clients de l'opérateur. Le processus de *k*-anonymisation utilisé protège-t-il efficacement l'anonymat de ces derniers ? Afin de pouvoir le garantir, un audit a été réalisé par le Security and Privacy Engineering Lab de l'EPFL.

GIOVANNI CHERUBIN, BOGDAN KULYNYCH, MARION LE TILLY, CARMELA TRONCOSO

En 2015, alors que Swisscom cherchait à mener des projets concrets dans le domaine de la smart city, la ville de Pully souhaitait, par le biais de son observatoire de la mobilité, réaliser une étude d'au moins 10 ans sur sa rue principale qui allait subir des transformations importantes. Chez Swisscom, le responsable de l'époque a tout de suite décelé le potentiel des données mobiles qui permettaient d'analyser la mobilité sur un territoire donné 24 h/24, 7 j/7 et 365 j/an.

C'est ainsi que la première version de la plateforme Swisscom Mobility Insights (MIP) a été développée en co-création avec la ville de Pully. Le thème de la protection de la vie privée des utilisateurs étant particulièrement sensible, l'opérateur a décidé en 2018, à l'occasion de la refonte de la plateforme dans une nouvelle version, de mandater le Security and Privacy Engineering Lab de l'EPFL pour auditer cette dernière. Cet article présente les résultats obtenus.

La plateforme Mobility Insights

La plateforme MIP a été développée afin de traiter les événements anonymes provenant du réseau de l'opérateur et d'établir des statistiques sur les déplacements effectués. Elle permet à ses clients de visualiser le nombre de trajets agrégés dans les régions qu'ils ont sélectionnées.

Un trajet est défini comme un chemin, entre un point de départ et un point d'arrivée, associé à un temps et à un mode de transport. Ils sont regroupés en trajets entrants (pour ceux qui commencent ailleurs et se terminent dans la région d'intérêt), sortants (lorsqu'ils partent de la région d'intérêt et aboutissent en dehors de celle-ci) et locaux (lorsque les points de départ et d'arrivée se trouvent dans la région d'intérêt). Ils sont donnés par heure et par jour, sur une période d'une semaine ou plus.

Protection de la vie privée au niveau de la plateforme

Afin de protéger la vie privée des abonnés, la plateforme a été conçue de manière à ne révéler le nombre de personnes dans une région d'intérêt que si plus de *k* personnes y sont détectées simultanément. Actuellement, *k* est fixé à 20 : la plateforme n'indique donc le nombre de personnes dans une région que si elle en détecte au moins 21. Si ce seuil n'est pas atteint, la plateforme ne retourne aucun résultat pour cette zone. Cette pratique suit le principe dit du filtre d'anonymat *k*.

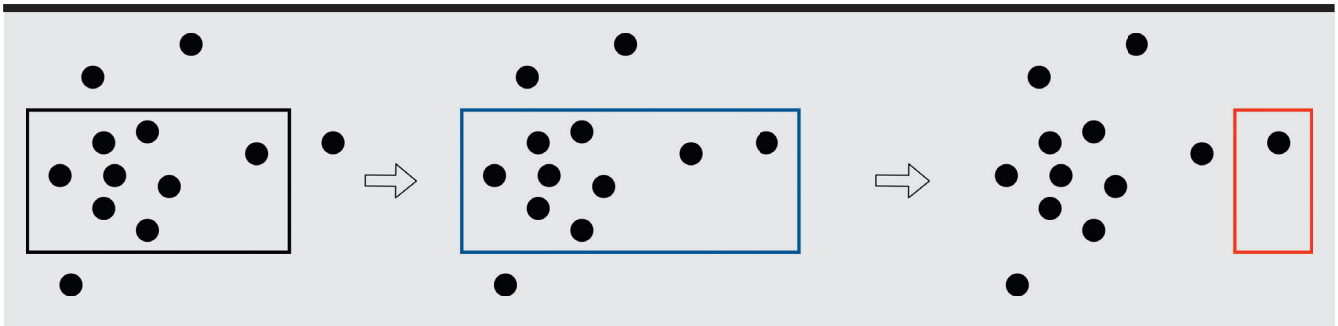

Figure 1 Déroulement de l'attaque de l'extension.

À noter aussi que ces calculs ne sont pas effectués directement à partir du nombre d'abonnés de Swisscom. Cet opérateur bénéficiant d'une part de marché d'environ 60 %, la plateforme estime le nombre de personnes dans une région en associant chaque abonné à la pénétration du marché de Swisscom dans la ville où il réside ; le résultat indiqué par la plateforme correspond à la somme du nombre d'abonnés détectés dans la région d'intérêt, pondérée par les valeurs respectives de pénétration du marché qui leur sont liées. Cette somme pondérée est ensuite ramenée au nombre entier le plus proche avant l'application du filtre d'anonymat *k*.

Risques d'atteinte à la vie privée

Dans une logique proactive, afin d'imaginer des scénarios d'attaques et d'évaluer le risque d'atteinte à la vie privée, le Security and Privacy Engineering Lab de l'EPFL a étudié la question suivante : existe-t-il un risque que les données fournies par la plateforme puissent, à un moment donné, être considérées comme des « données personnelles » selon la législation sur la protection des données ? Afin de pouvoir répondre à cette question, il a été décidé de réaliser cette étude en suivant les critères fournis dans l'« Avis sur les techniques d'anonymisation » du groupe de travail de l'UE sur la protection des personnes à l'égard du traitement des données à caractère personnel [1]. Dans ce rapport, le groupe de travail expose en détail trois risques qui rendraient les données non anonymes :

- le ciblage, qui correspond à la possibilité d'isoler certains ou tous les enregistrements associés à une personne dans l'ensemble de données ;
- le couplage, qui correspond à la possibilité de relier (au moins) deux enregistrements concernant la même personne ou un groupe de personnes concernées, et ce, soit dans la même base de données, soit dans deux bases de données différentes ;
- et l'inférence, qui consiste en la possibilité de déduire avec une probabilité significative la valeur d'un attribut à partir des valeurs d'un ensemble d'autres attributs.

Le mécanisme de protection de la vie privée choisi par Swisscom a pour but de ne jamais révéler de données pouvant être associées à moins de *k* personnes. Au cours de cette étude, l'analyse a donc été axée sur la possibilité qu'un adversaire soit en mesure de distinguer des utilisateurs à partir des données produites par la plateforme. Le ciblage constitue le tout premier pas vers la désanonymisation : une fois les données d'un utilisateur isolées, l'adversaire peut commencer à effectuer des opérations sur ces données afin d'essayer de découvrir des informations privées sur cet utilisateur.

L'attaque d'extension

La stratégie utilisée au cours de cette étude pour isoler les utilisateurs a été dénommée « attaque d'extension ». Dans ce type d'attaque, l'adversaire identifie d'abord une région pour laquelle la plateforme renvoie une valeur (c'est-à-dire qu'il s'y trouve au moins 21 individus). Ensuite, l'adversaire continue d'étendre cette région jusqu'à ce que le nombre soit augmenté de 1. À ce moment, l'adversaire sait qu'il n'y a qu'un seul utilisateur dans la région étendue. Cette attaque est illustrée dans la **figure 1,** en supposant que $k = 7$.

Dans la partie gauche de la **figure 1,** la région sélectionnée indique 8 individus, valeur également retournée par la plateforme. De manière identique, au milieu, la plateforme indique 9 individus. À partir de là, l'adversaire peut déduire que la région sélectionnée dans la partie droite de la **figure 1** ne contient qu'un seul individu, ce qui constitue une violation du mécanisme de protection de la vie privée. En suivant cette démarche, il est possible de :

- sélectionner n'importe quelle région cible sur la carte, la plateforme permettant aux utilisateurs de dessiner des polygones arbitraires comme régions d'intérêt (par exemple, l'adversaire peut sélectionner une grande zone à l'intérieur d'une ville et ensuite s'étendre à une petite région à l'extérieur du centre urbain) ;
- réduire la région cible, par exemple par la méthode de la bissection, en essayant des régions d'extension plus petites jusqu'à une taille arbitraire (l'adversaire peut notamment cibler une maison ou un bâtiment).

Par le biais de l'attaque d'extension ciblée, l'adversaire peut apprendre où un utilisateur vit en regardant où il passe ses nuits. Avec cette information, il est alors potentiellement capable de récupérer l'identité de cet utilisateur.

L'attaque d'extension sur la plateforme MIP

Les chercheurs de l'EPFL ont testé la faisabilité de cette attaque sur la plateforme MIP. Ils ont pu constater que deux requêtes suffisaient à un adversaire pour trouver des régions dans lesquelles il n'y avait qu'un seul utilisateur, et ils ont pu prouver que celui-ci pouvait effectuer cette inférence dans des régions arbitraires.

Un scénario dans lequel l'adversaire cible un utilisateur dont il connaît l'adresse de domicile a ensuite été envisagé. Pour ce faire, l'attaque d'extension a été déployée sur l'adresse du domicile d'un des collaborateurs du

Figures : EPFL

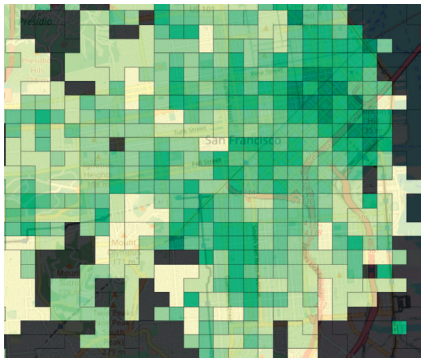

Figure 2 Sélection de partitions prédéfinies : exemple issu de la stratégie dite « de fusion ».

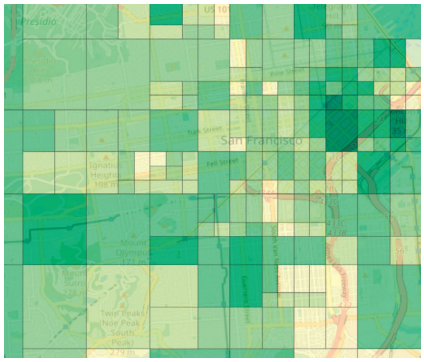

Figure 3 Stratégie de partitionnement : exemple de résultat obtenu après l'utilisation de la stratégie dite « de scission ».

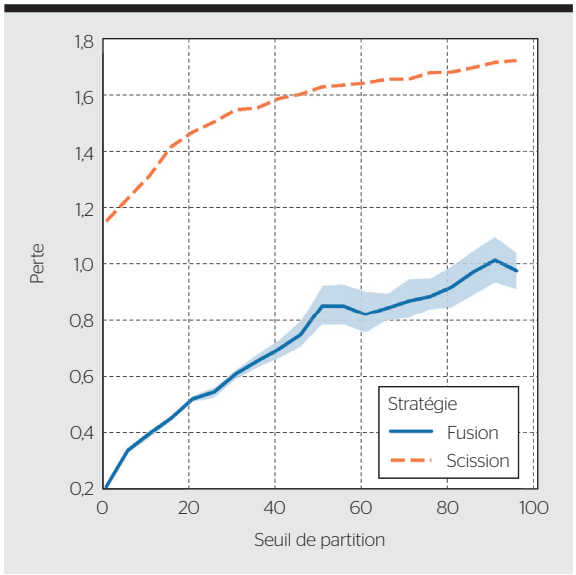

Figure 4 Comparaison des pertes d'utilité respectives des deux stratégies de partition en fonction du seuil de partitionnement (temps de calcul nécessaire pour la stratégie de fusion : 45,6 s ; temps de calcul nécessaire pour la stratégie de scission : 4,1 s).

laboratoire : il a ainsi été possible d'identifier l'un de ses déplacements à l'intérieur de la ville de Lausanne, un matin de mars 2018.

Ensuite, étant donné que les chiffres produits par la plateforme ne proviennent pas directement des abonnés mais qu'ils sont adaptés pour refléter la population suisse, il a été évalué dans quelle mesure il était possible de déduire le nombre effectif d'abonnés de Swisscom à partir des résultats indiqués par la plateforme. Concrètement, les combinaisons possibles d'abonnés et de parts de marché pouvant correspondre aux résultats indiqués ont été explorées (également après l'attaque d'extension). Puis, une expression combinatoire a été dérivée, qui détermine la probabilité d'occurrence des différentes valeurs de sortie à partir d'un nombre réel d'abonnés de Swisscom. Ceci a permis de constater que pour des valeurs inférieures à k (20), un adversaire pouvait prédire avec une forte probabilité le nombre réel d'abonnés. Par conséquent, ce mécanisme n'offre pas de protection supplémentaire de la vie privée des clients.

Vers la prévention de l'attaque

Une piste de solution prometteuse en vue d'améliorer la confidentialité des clients, tout en préservant l'usage ou la pertinence des résultats, consiste à limiter la façon dont les utilisateurs peuvent interroger la plateforme. Par exemple, en n'autorisant les requêtes que dans des régions prédéfinies sous forme de carrés (**figure 2**).

La justification de cette approche est la suivante : si une région a peu d'utilisateurs (par exemple si elle est située en montagne), les clients de la plateforme seront moins intéressés par des requêtes concernant de petites zones, car celles-ci ne contiendront, avec une forte probabilité, que très peu de personnes. Ainsi, la plateforme peut utiliser des partitions plus importantes dans ces régions, de sorte à répondre à la plupart des requêtes. Dans les endroits habituellement très fréquentés, la plateforme peut se permettre des partitions plus fines, car il est très probable qu'elle retournera une valeur.

Pour éviter tout problème de confidentialité, il serait judicieux de construire les partitions sur la base de données historiques agrégées. Comme il est peu probable que les modèles changent très fréquemment, le même partitionnement peut être utilisé pendant une longue période.

Le but d'une stratégie de partitionnement est double. D'une part, elle doit garantir la vie privée : elle doit interdire à un hacker de déterminer la localisation d'un individu (ou d'un petit nombre de personnes) et donc être apte à contrer les attaques d'extension. D'autre part, du point de vue de l'utilité, elle doit permettre aux clients de la plateforme de trouver des informations précises sur leurs zones d'intérêt.

Création des partitions

Pour construire les partitions, deux stratégies sont proposées. Ces deux méthodes, la fusion et la scission, permettent de faire varier l'utilité en fonction du risque de réauthentification. Les deux stratégies prennent comme entrée un seuil qui définit le nombre minimum de personnes qu'une seule partition devrait contenir en moyenne.

La **stratégie de fusion** part d'une petite partition. Ensuite, tant qu'il existe dans cette partition des régions comptant un nombre de personnes inférieur au seuil, l'une d'entre elles est choisie au hasard et fusionnée avec l'une des régions voisines (également déterminée de manière aléatoire). Le nombre de personnes indiqué pour la nouvelle région correspond à la somme des nombres des deux régions fusionnées. Ce processus se termine lorsque chaque région compte un nombre de personnes supérieur au seuil fixé. La **figure 2** illustre le résultat de cette stratégie.

Il a été décidé de choisir au hasard la région à fusionner et sa voisine, mais d'autres stratégies pourraient être envisagées, comme la sélection des régions ayant le plus petit nombre de personnes ;

Stratégie de partitionnement	0	*k*/2	*k*
Fusion	0,793	0,789	1,121
Scission	1,50	1,496	1,558

Tableau Perte d'utilité, pour les stratégies de fusion et de scission, en fonction du choix du type d'indication quand le seuil *k* n'est pas atteint dans une région.

toutefois, leur mise en œuvre représente des efforts supplémentaires et des calculs plus coûteux, et elle devrait mener à un résultat très similaire.

La **stratégie de scission** commence, quant à elle, par considérer un rectangle couvrant toute la carte comme une partition unique. À chaque étape et pour chaque partition existante, la partition est divisée en 4 et le nombre de personnes se trouvant à l'intérieur de chaque partie est compté : si ces dernières contiennent toutes suffisamment de personnes (c'est-à-dire un nombre supérieur au seuil fixé), ce partitionnement est accepté. Dans le cas contraire, la stratégie retourne à la partition précédente et passe à la division d'autres partitions. Le partage à 4 voies (par opposition à un partage à 2 voies) est simplement dicté par la commodité : il permet une implémentation plus simple et produit des partitions dont les formes sont plus faciles à gérer. La **figure 3** illustre le résultat de cette stratégie de partitionnement.

Après avoir appliqué l'une de ces stratégies, Swisscom peut modifier manuellement les partitions en fonction des besoins de ses clients. Par exemple, il est possible de choisir de créer une partition personnalisée en fusionnant ou en divisant les partitions générées. À noter cependant qu'une fois le partitionnement finalisé, il ne doit plus être modifié pendant une longue période (par exemple un an).

Quelle stratégie utiliser ?

Le choix de la stratégie de partitionnement devrait dépendre de l'utilité, des coûts de calcul et, enfin, de la protection de la vie privée. Les questions relatives à la protection de la vie privée seront abordées plus bas. En ce qui concerne l'utilité et les coûts de calcul, une série d'expériences préliminaires, basées sur un ensemble de données publiques contenant 313289 localisations de personnes à San Francisco, ont été menées.

La perte d'utilité de la plateforme a été mesurée selon la distance de distribution entre le nombre de personnes indiqué dans un ensemble de zones avant et après le partitionnement (voir l'exemple de combinaison de partitions donné plus bas). Concrètement, une petite grille a été considérée comme entrée pour la procédure de fusion (les régions comprennent quelques pâtés de maisons) et la valeur de chaque cellule a été comparée à la région qui la contient après le partitionnement. La **figure 4** montre que la stratégie de fusion fournit une meilleure utilité, au prix d'un calcul plus lourd pour établir la grille.

Résultats indiqués par la plateforme

Une fois les partitions construites, un client peut interroger une ou plusieurs partitions et la plateforme lui retournera le nombre de personnes correspondant. Dans certains cas et pour des tranches horaires particulières, les partitions créées à l'étape précédente peuvent toutefois ne pas contenir suffisamment de personnes pour être utilisées en toute sécurité. Ceci est dû au fait qu'à certains moments, par exemple la nuit ou pendant les périodes de vacances, moins de personnes se trouvent à cet endroit. Ce n'est que lorsqu'une région prédéfinie compte plus que *k* utilisateurs (soit plus de 20), que la valeur est retournée par la plateforme. Or, il est important d'établir ce que la plateforme doit indiquer si le nombre de personnes dans une partition est inférieur à *k*.

D'une part, du point de vue de la protection de la vie privée, les résultats de la plateforme devraient être indépendants du nombre réel de personnes dans la partition. D'autre part, du point de vue de l'utilité, ils ne devraient pas avoir d'impact sur l'expérience des clients de la plateforme.

Voici certains choix naturels dans ces circonstances : la plateforme peut indiquer une valeur qui dépend de *k* (par exemple la valeur du seuil ou la moitié de cette valeur), « 0 » (c'est-à-dire, personne) ou une valeur interpolée à partir du nombre de personnes se trouvant dans les partitions voisines. Le **tableau** compare, pour 0, *k*/2 et *k*, la perte d'utilité calculée comme étant la différence moyenne entre les sommes des chiffres indiqués pour chaque zone avant et après le partitionnement (voir l'exemple

Quelques précisions

Note de Swisscom

Le travail de l'équipe de Carmela Troncoso a permis de mettre en lumière une façon de contourner la protection de confidentialité offerte par notre plateforme. C'est précisément l'objectif de ce type d'exercice qui doit être répété régulièrement. Les conditions dans lesquelles ces résultats ont été obtenus doivent être clarifiées, car elles ne répondent pas à une utilisation dite « normale » de la plateforme. En effet, une version dédiée a été mise à disposition de l'équipe de l'EPFL ; celle-ci utilisait exactement les mêmes données que la plateforme de production, mais permettait de définir un nombre illimité de zones et également d'accéder directement à l'API du front-end. D'après l'équipe de l'EPFL, ce n'est pas tant le nombre de zones (en théorie, deux zones suffisent pour réussir l'attaque de l'extension), mais la possibilité d'accéder à l'API directement qui a permis de raccourcir considérablement les cycles d'attaques et d'aboutir au résultat obtenu. En production et suite à l'étude, l'équipe a immédiatement mis en place des solutions temporaires afin d'éviter l'attaque tout en travaillant à un processus automatique qui réglerait définitivement ce problème. Nous savons qu'en matière de protection des données, nous n'avons pas droit à l'erreur et continuons jour après jour à améliorer le produit.

Yann Steimer
Product Manager Mobility Insights Platform
insights.info@swisscom.com

Figure 5 Combinaison de partitions.

donné ci-dessous). Même dans ce cas, la fusion donne une meilleure utilité que la scission. En outre, le choix de retourner la valeur $k/2$ semble être la meilleure option en termes d'utilité.

Prenons un cas de combinaison de partitions: si un client interroge plus d'une partition, le résultat de la plateforme pour chaque partition doit être choisi indépendamment de la somme des partitions. Par exemple, supposons qu'un client choisisse, dans la carte de la **figure 5,** les partitions A (40 personnes), B (15 personnes) et C (45 personnes). Même si le nombre total de personnes dans la partition (A, B, C) est supérieur à k, la plateforme doit d'abord remplacer le nombre de personnes dans B selon l'une des stratégies décrites ci-dessus, puis calculer la somme des trois partitions. Par exemple, si la stratégie choisie consiste à utiliser la moitié du seuil, la plateforme indiquera $40 + k/2 + 45 = 95$ (au lieu de $40 + 15 + 45 = 100$). Cet élément est crucial pour éviter des attaques similaires à l'attaque d'extension.

Analyse de confidentialité

L'utilisation d'un système à base de grilles évite en principe l'attaque d'extension. En effet, il n'est pas possible de sélectionner des régions librement et donc, d'étendre la sélection sur une maison, par exemple. Cependant, il peut y avoir des attaques qui interrogent la plateforme sur différentes combinaisons de régions prédéfinies et qui obtiennent des informations sur des trajets individuels. L'évaluation de la mesure dans laquelle ces attaques sont possibles constitue un prochain sujet de recherche. À noter toutefois que même si l'adversaire pouvait isoler les déplacements, la stratégie de cloisonnement proposée dans cet article l'empêche déjà de cibler des bâtiments ou des régions d'intérêt.

Référence

[1] Article 29 of Directive 95/46/EC Data Protection Working Party, «Opinion 05/2014 on Anonymisation Techniques», adopted on 10 April 2014. ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf

Auteurs

Giovanni Cherubin est postdoc au sein du Security and Privacy Engineering Lab de l'EPFL.
→ EPFL, 1015 Lausanne
→ giovanni.cherubin@epfl.ch

Bogdan Kulynych est doctorant au sein du Security and Privacy Engineering Lab de l'EPFL.
→ bogdan.kulynych@epfl.ch

Marion LeTilly est étudiante et effectue son Master au sein du Security and Privacy Engineering Lab de l'EPFL.
→ marion.letilly@epfl.ch

Carmela Troncoso est professeure assistante Tenure Track au sein du Security and Privacy Engineering Lab de l'EPFL.
→ carmela.troncoso@epfl.ch

Die deutsche Version dieses Beitrags wird im Bulletin 9/2020 erscheinen.