L'analyse côté client

Qu'est-ce que c'est et pourquoi menace-t-elle la confidentialité et la fiabilité des communications?

Août 2022

Le cryptage est une technologie conçue pour aider les utilisateurs d'Internet à garantir la confidentialité et la sécurité de leurs informations et de leurs communications. Le processus de cryptage brouille les informations, afin que celles-ci ne puissent être lues que par quelqu'un qui dispose de la « clé » pour décrypter ces informations. Le cryptage protège les activités du quotidien telles que la gestion bancaire et les achats en ligne. Il évite également que des données soient volées en cas de piratage de données, et permet de s'assurer que les messages confidentiels le restent. Le cryptage est également crucial pour protéger les communications des forces de l'ordre, du personnel militaire et, de plus en plus, des intervenants en cas d'urgence.

Le cryptage de bout en bout (en anglais, end-to-end ou E2E), où les clés nécessaires pour décrypter une communication cryptée résident uniquement sur les appareils qui communiquent, offre le niveau de sécurité et de confiance le plus élevé. De par sa conception, seul le destinataire prévu détient la clé pour décrypter le message. Le cryptage E2E est un outil crucial pour assurer la sécurité et la confidentialité des communications. L'analyse du message, même si elle a lieu « côté client », détruit le modèle de cryptage E2E et nuit fondamentalement à la confidentialité attendue par les utilisateurs.

Qu'est-ce que l'analyse côté client?

L'analyse côté client (en anglais, client-side scanning ou CSS) est un terme générique faisant référence aux systèmes qui analysent les données du message (ex.: texte, images, vidéos, fichiers) afin de rechercher des correspondances ou des similitudes avec une base de données de contenu répréhensible avant l'envoi du message à son destinataire. Par exemple, votre logiciel anti-virus peut y avoir recours pour trouver et désactiver des logiciels malveillants sur votre ordinateur.

Alors que les principaux fournisseurs de plateformes mettent en œuvre un cryptage E2E et que certains membres des forces de l'ordre demandent à faciliter l'accès au contenu des messages pour mieux identifier et empêcher le partage de données à caractère répréhensible[1], l'analyse côté client

1 https://www.newamerica.org/oti/press-releases/open-letter-law-enforcement-us-uk-and-australia-weak-encryption- puts-billions-internet-users-risk/

pourrait devenir le mécanisme de prédilection pour lutter contre les données à caractère répréhensible partagées sur des services à cryptage E2E sans en compromettre le cryptage.

Cependant, l'analyse côté client compromettrait la confidentialité et la sécurité auxquels les utilisateurs s'attendent et se fient. En empêchant les données du message de rester privées entre l'émetteur et le destinataire, l'analyse côté client détruit le modèle de confiance du cryptage E2E. La complexité que cela vient ajouter limite également la fiabilité du système de communication, et risque d'empêcher des messages authentiques de parvenir à leur destinataire.

L'analyse côté client pour empêcher le partage de données à caractère répréhensible

Lorsqu'elle vise à empêcher le partage de données à caractère répréhensible, l'analyse côté client fait généralement référence à une façon dont un logiciel sur les appareils des utilisateurs (souvent désignés sous le terme de « clients », notamment les smartphones, les tablettes ou les ordinateurs) crée des « empreintes » numériques uniques2 pour les données de l'utilisateur (nommés « hachages »). Le logiciel compare ensuite ces empreintes aux empreintes numériques de données à caractère répréhensible connues, comme des logiciels malveillants (malwares), des images, des vidéos ou des graphiques.3 En cas de correspondance, le logiciel peut empêcher ce fichier d'être envoyé et signaler cette tentative à un tiers, souvent à l'insu de l'utilisateur. Les nouvelles approches d'analyse côté client recherchent également de nouvelles données à caractère répréhensible à l'aide d'algorithmes plus sophistiqués. Ceci est difficile et rend encore plus probable le risque de faux positifs.

Comment fonctionne l'analyse côté client?

Il existe deux méthodes de base d'analyse côté client pour les données à caractère répréhensible sur un service de communication à cryptage E2E. L'une effectue la comparaison des empreintes numériques sur l'appareil de l'utilisateur, et l'autre effectue la comparaison sur un serveur distant (les données restent sur l'appareil).

1. **Comparaison effectuée sur l'appareil de l'utilisateur (mise en correspondance locale des empreintes numériques)**

---

2 Il serait possible de développer un système où les empreintes numériques ne seraient pas aussi uniques, ce qui engendrerait l'utilisation de la même empreinte numérique par différents contenus. Cependant, dans les cas où de faux positifs risquent d'entraîner l'utilisation de ressources importantes (notamment une intervention policière), les concepteurs des systèmes d'analyse côté client chercheront à rendre les empreintes aussi uniques que possible.

3 L'analyse côté client n'est que l'une des manières proposées par les forces de l'ordre ou les agences de sécurité pour obtenir l'accès aux communications cryptées des utilisateurs. Pour plus d'informations, consultez : https://www.internetsociety.org/resources/doc/2018/encryption-brief/

L'application sur l'appareil d'un utilisateur (téléphone, tablette ou ordinateur) dispose d'une base de données complète et à jour d'empreintes numériques fonctionnellement uniques de données surveillées connues. Les données que l'utilisateur est sur le point de crypter et d'envoyer dans un message sont converties en une empreinte numérique avec les mêmes techniques que celles utilisées pour les empreintes numériques de la base de données complète. Si une correspondance est trouvée, ou si un algorithme classe les données comme étant probablement à caractère répréhensible, le message peut ne pas être envoyé et un tiers désigné (tel que les forces de l'ordre, les agences de sécurité nationale ou le fournisseur des services de filtrage) peut être notifié.

2. **Comparaison effectuée sur un serveur à distance**
   Le maintien d'une base de données complète et d'algorithmes sophistiqués, qui effectuent une analyse en temps réel sur l'appareil d'un utilisateur, peut présenter des défis importants. L'autre solution est donc de transférer les empreintes numériques des données d'un utilisateur vers un serveur où sera effectuée la comparaison avec une base de données centrale.

Les problèmes posés par l'analyse côté client des données à caractère répréhensible

Lorsque la comparaison des empreintes numériques est effectuée sur un serveur à distance, cela peut permettre au fournisseur d'accès, ainsi qu'à toute personne avec qui le fournisseur accepte de partager ces informations, de surveiller et de filtrer les données que souhaite envoyer l'utilisateur. Lorsque la comparaison est effectuée sur l'appareil de l'utilisateur et que des tiers sont informés en cas de découverte de données inappropriées, le problème reste le même. Cela va fondamentalement à l'encontre de la raison d'être du cryptage E2E. Les communications à cryptage E2E privées et sécurisées entre deux parties, ou au sein d'un groupe, ont vocation à rester privées. Si des personnes suspectent que leurs données sont analysées, elles peuvent s'auto-censurer, passer à un autre service sans analyse côté client ou utiliser un autre moyen de communication.

Cela crée des vulnérabilités que peuvent exploiter des criminels: l'ajout d'une fonctionnalité d'analyse côté client augmente la surface d'attaque avec la création de moyens supplémentaires d'interférer dans les communications en manipulant la base contenant les données à caractère répréhensible. Des personnes mal intentionnées ayant la possibilité d'ajouter des empreintes numériques dans la base de données et de recevoir des notifications en cas de correspondance avec ces empreintes auraient la possibilité de surveiller les données d'un utilisateur spécifique avant qu'elles ne soient cryptées et envoyées. Cela leur permettrait de surveiller à qui, quand et où certaines données ont été transmises. Ces empreintes pourraient inclure les mots de passe fréquemment utilisés ou d'autres informations, permettant ainsi des attaques telles que le piratage psychologique, l'extorsion ou le chantage. En utilisant les fonctionnalités de blocage d'un système,

des criminels pourraient également décider d'empêcher des utilisateurs d'envoyer des données spécifiques. Cela pourrait servir à cibler des utilisations légitimes, notamment en nuisant aux communications des forces de l'ordre, des intervenants en cas d'urgence ou du personnel de la sécurité nationale.

Cela crée de nouveaux défis techniques et procéduraux: si les comparaisons sont effectuées sur l'appareil de l'utilisateur, tenir à jour la base de données de références complète sur chaque appareil représente déjà de nombreux défis. Parmi ces défis figurent des contraintes procédurales potentielles (ex.: le processus permettant d'ajouter ou de supprimer des empreintes de données dans la base de données ; qui contrôle la base de données ou peut y accéder), la bande passante nécessaire pour transmettre les versions mises à jour de la base de données et la puissance de traitement sur les appareils nécessaire pour effectuer la comparaison en temps réel. Il existe également d'autres considérations à prendre en compte, notamment l'exposition potentielle de la base de données de référence lors de son installation sur l'appareil client, susceptible d'offrir aux criminels des informations sur le système d'analyse. Si les comparaisons sont effectuées sur un serveur central, l'empreinte numérique des données que l'utilisateur cherche à envoyer sera accessible à toute personne contrôlant le serveur central, que ces données soient « à caractère répréhensible » ou non aux yeux de la partie chargée de la surveillance. Cela engendre un nouvel ensemble de problèmes relatifs à la sécurité et à la confidentialité des utilisateurs, avec le risque d'exposer des informations sur leur activité à toute personne ayant accès au serveur.

Risques de dérives et d'utilisation à d'autres fins: les mêmes méthodes mises en œuvre dans l'espoir de lutter contre les pires données (notamment à caractère pédopornographique ou terroriste, les deux exemples les plus souvent cités pour justifier leur utilisation) peuvent également être utilisées pour la surveillance de masse et à des fins répressives. Un article de 2021 sur les risques de l'analyse côté client a révélé qu'un système CSS pourrait être conçu de manière à offrir à une agence la possibilité d'analyser de manière préventive tout type de données sur n'importe quel appareil, à quelque fin que ce soit, sans mandat ni soupçon. De même, les mêmes techniques visant à empêcher la distribution de contenu pédopornographique (en anglais, Child Sexual Abuse Material ou CSAM) peuvent être utilisées pour appliquer des politiques telles que la censure et la suppression de la dissidence politique en empêchant le partage de données légitimes ou en bloquant les communications entre les utilisateurs (tels que les opposants politiques). Il est difficile de restreindre la base de données uniquement aux empreintes d'images, de vidéos ou d'URL relatives à des activités illégales, comme certains le proposent. En créant des empreintes numériques pour davantage de données afin de pouvoir établir des comparaisons avec les empreintes numériques des données de l'utilisateur ou en élargissant la portée d'un algorithme pour classer d'autres types de données de l'utilisateur comme répréhensibles, toute personne responsable de la base de données peut surveiller n'importe quel type de contenu. Un système d'analyse côté client pourrait être étendu pour surveiller les données textuelles des messages envoyés, avec des implications claires et dévastatrices pour la liberté d'expression.

Manque d'efficacité: des systèmes de communication à cryptage E2E existent en dehors de la juridiction de tout gouvernement. Un criminel réellement déterminé pourrait se passer des services dont il sait qu'ils ont recours à une analyse côté client afin d'éviter d'être repéré. Il est techniquement simple pour les criminels d'apporter des modifications à des données à caractère répréhensible, changeant ainsi l'empreinte numérique et évitant la détection par le système d'analyse côté client.

Conclusion

La lutte contre le partage de données à caractère terroriste ou pédopornographique est une cause importante. Cependant, cela ne peut pas être effectué d'une manière qui affaiblirait la sécurité de tous les utilisateurs en modifiant l'infrastructure de communication pour potentiellement surveiller les échanges de tout un chacun. L'analyse côté client réduit la sécurité globale et la confidentialité pour les utilisateurs respectueux de la loi tout en courant le risque de ne pas atteindre son objectif déclaré d'application de la loi. Le cryptage E2E garantit à des milliards d'utilisateurs à travers le monde la possibilité de communiquer de façon sécurisée et confidentielle.[4] Des plateformes majeures poursuivent leurs démarches en vue de son adoption afin d'assurer davantage de fiabilité à leurs plateformes et services.[5] L'analyse côté client des services de communications à cryptage E2E ne constitue pas une solution appropriée pour le filtrage des données à caractère répréhensible, et cela vaut également pour toute méthode qui nuirait au fondement même de la confidentialité et de la fiabilité des communications desquelles nous dépendons.

Références

Internet Society, juin 2018. Encryption Brief.

Matthew Green, décembre 2019. Can end-to-end encrypted systems detect child sexual abuse imagery?

Electronic Frontier Foundation, novembre 2019. Why Adding Client-Side Scanning Breaks End-To-End Encryption.

Centre for Democracy and Technology (CDT), 2021. Content Moderation in Encrypted Systems.

4 https://telegram.org/blog/200-million et https://www.newsweek.com/whatsapp-facebook-passes-two-billion-users-pledges-encryption-support-1486993

5 https://www.facebook.com/notes/2420600258234172/

Hal Abelson, Ross Anderson, Steven M. Bellovin, Josh Benaloh, Matt Blaze, Jon Callas, Whitfield Diffie, Susan Landau, Peter G. Neumann, Ronald L. Rivest, Jeffrey I. Schiller, Bruce Schneier, Vanessa Teague, Carmela Troncoso, octobre 2021. Bugs in our Pockets: The Risks of Client-Side Scanning.