Die Mobilität sicher verstehen

Anonymitätstest | Die Plattform Mobility Insights von Swisscom ist ein Online-Tool, das anonyme, aus Mobilfunkdaten erhobene Statistiken liefert. Aber ist die Anonymität der mit ihr erfassten Personen durch die verwendete Anonymisierungsmethode auch ausreichend geschützt? Das Security and Privacy Engineering Lab der ETH Lausanne ging mit einem Audit dieser Frage nach.

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

Im Jahr 2015 initiierte Swisscom konkrete Smart-City-Projekte. Gleichzeitig wollte die Gemeinde Pully durch ihre Beobachtungsstelle für Mobilität eine mindestens zehnjährige Studie über ihre Hauptstrasse, die einen grossen Wandel erfahren sollte, durchführen. Der damalige Verantwortliche bei Swisscom erkannte sofort das Potenzial von Mobilfunkdaten, die auf einem begrenzten Gebiet an 365 Tagen des Jahres rund um die Uhr Mobilitätsanalysen ermöglichten.

Auf dieser Basis entstand in Zusammenarbeit mit der Gemeinde Pully die erste Version der Mobility Insights Plattform (MIP) von Swisscom. Da der Schutz der Privatsphäre der Abonnenten ein hochsensibles Thema ist, beschloss der Betreiber 2018, als die Plattform für eine neue Version überarbeitet wurde, das Security and Privacy Engineering Lab der EPFL mit der Prüfung der Plattform zu beauftragen. Dieser Artikel stellt die erzielten Ergebnisse vor.

Die Plattform Mobility Insights

Die Plattform wurde entwickelt, um aus dem Netz des Mobilfunkbetreibers stammende anonyme Ereignisse auszuwerten und Statistiken über die erfassten Bewegungen zu erstellen. Sie ermöglicht die visuelle Darstellung der aggregierten «Fahrten» in kundenspezifisch ausgewählten Gebieten.

Eine Fahrt ist definiert als eine Bewegung zwischen einem Ausgangs- und einem Ankunftsort, verbunden mit einem Zeitraum und einem Transportmittel. Die Fahrten sind in ankommende Fahrten (die an anderer Stelle beginnen und im Beobachtungsgebiet enden), abgehende Fahrten (die im Beobachtungsgebiet beginnen und ausserhalb dieses Gebiets enden) und lokale Fahrten (die innerhalb des Beobachtungsgebiets verlaufen) unterteilt. Die Fahrten werden über einen Zeitraum von einer Woche oder über einen längeren Zeitraum hinweg pro Stunde oder pro Tag angegeben.

Datenschutz auf Plattformebene

Um die Privatsphäre der Abonnenten zu schützen, wurde die Plattform so konzipiert, dass die Anzahl der Personen in einem Beobachtungsgebiet nur dann bekannt gegeben wird, wenn dort mehr als k Personen gleichzeitig erkannt werden. Aktuell ist k auf 20 festgelegt: Die Plattform nennt die Anzahl der Personen in einem Gebiet also nur dann, wenn mindestens 21 Personen erkannt werden. Wird diese Schwelle nicht erreicht, liefert die Plattform für dieses

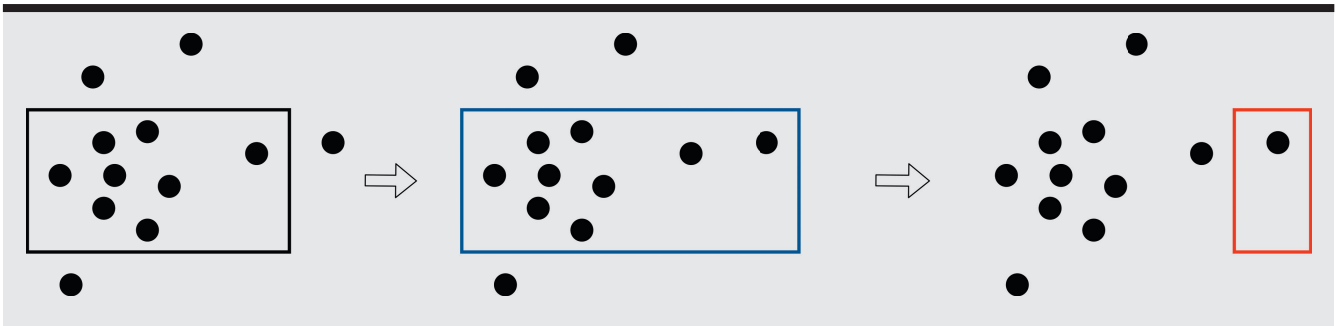

Bild 1 Ablauf eines Erweiterungsangriffs.

Gebiet kein Ergebnis. Diese Praxis folgt dem sogenannten k-Anonymitätsfilterprinzip.

Diese Berechnungen basieren nicht direkt auf der Anzahl der Swisscom-Abonnenten. Da dieser Betreiber über einen Marktanteil von rund 60 % verfügt, schätzt die Plattform die Anzahl Personen in einer Region, indem sie jeden Abonnenten mit der Marktdurchdringung von Swisscom in der Stadt seines Wohnsitzes in Verbindung bringt; das von der Plattform angezeigte Ergebnis entspricht der Summe der in der interessierenden Region ermittelten Abonnentenzahl, gewichtet mit den jeweils zugehörigen Marktdurchdringungswerten. Diese gewichtete Summe wird dann auf die nächste ganze Zahl gerundet, bevor der Anonymitätsfilter angewendet wird.

Risiko einer Datenschutzverletzung

Um im Rahmen einer proaktiven Logik Angriffsszenarien vorwegzunehmen und das Risiko einer Datenschutzverletzung zu bewerten, hat das Security and Privacy Engineering Lab der ETH Lausanne folgende Frage untersucht: Besteht die Gefahr, dass das Datenschutzrecht die zu einem gegebenen Zeitpunkt durch die Plattform bereitgestellten Daten als «personenbezogene Daten» einstuft? Zur Beantwortung dieser Frage wurde entschieden, die Studie gemäss den in der «Stellungnahme zu Anonymisierungstechniken» der Gruppe für den Schutz von Personen bei der Verarbeitung personenbezogener Daten der EU [1] genannten Kriterien durchzuführen. Die Arbeitsgruppe führt in ihrer Stellungnahme drei Risiken auf, die zu nicht anonymen Daten führen können:

- das Herausgreifen, d. h. die Möglichkeit, in einem Datenbestand einige oder alle Datensätze zu isolieren, die die Identifizierung einer Person ermöglichen;
- die Verknüpfbarkeit, d. h. die Fähigkeit, (mindestens) zwei Datensätze, die dieselbe Person oder Personengruppe betreffen, zu verknüpfen (in derselben Datenbank oder in zwei verschiedenen Datenbanken);
- die Inferenz, das heisst, die Möglichkeit, den Wert eines Merkmals mit einer signifikanten Wahrscheinlichkeit von den Werten einer Reihe anderer Merkmale abzuleiten.

Der durch Swisscom gewählte Datenschutzmechanismus zielt darauf ab, nie Daten offenzulegen, die nicht mindestens k Personen zugeordnet werden können. Die Analyse der Studie konzentriert sich somit auf die Möglichkeit, dass ein Angreifer in der Lage sein könnte, Benutzer anhand der durch die Plattform gelieferten Daten voneinander zu unterscheiden. Der allererste Weg der Deanonymisierung ist das Herausgreifen: Sind die Daten eines Benutzers erst einmal isoliert, kann der Angreifer durch Verarbeitung dieser Daten versuchen, personenbezogene Informationen dieses Benutzers zu erhalten.

Der Erweiterungsangriff

Die in dieser Studie verwendete Strategie zur Isolierung von Benutzern wurde als «Erweiterungsangriff» bezeichnet. Bei dieser Art von Angriff identifiziert der Gegner zunächst eine Region, für die die Plattform einen Wert zurückgibt (d. h. es hat dort mindestens 21 Personen). Dann dehnt der Gegner diese Region aus, bis die Zahl um 1 erhöht wird. Nun weiss der Angreifer, dass sich im Erweiterungsgebiet nur ein einziger Benutzer befindet. **Bild 1** illustriert diese Art Angriff unter der Annahme $k = 7$.

Im linken Teil von **Bild 1** befinden sich acht Personen im ausgewählten Gebiet. Diese Anzahl wird auch von der Plattform ausgegeben. In der Mitte liefert die Plattform entsprechend die Anzahl neun. Daraus kann der Angreifer schliessen, dass sich in dem im rechten Teil von **Bild 1** ausgewählten Gebiet nur eine Person aufhält. Der Datenschutzmechanismus funktioniert hier also nicht. Ausgehend von dieser Vorgehensweise ist Folgendes möglich:

- Auf der Karte wird ein beliebiges Zielgebiet ausgewählt. Die Plattform gestattet es den Benutzern, die Beobachtungsgebiete durch Zeichnen beliebiger Polygone festzulegen (wobei der Angreifer beispielsweise ein grosses Gebiet innerhalb einer Stadt wählen und dieses dann um ein kleines Gebiet ausserhalb des Stadtzentrums erweitern kann);
- Das Zielgebiet wird verkleinert (z. B. durch Halbierung), wobei immer kleinere Erweiterungsgebiete bis hin zu einer beliebig kleinen Grösse ausprobiert werden (und der Gegner zum Beispiel ein Haus oder ein Gebäude ins Visier nehmen kann).

Der Angreifer kann mittels dieser gezielten Erweiterungen die Adresse eines Benutzers ermitteln, indem er beobachtet, wo er seine Nächte verbringt. Mit diesen Informationen ist er dann potenziell in der Lage, die Identität dieses Benutzers herauszufinden.

Erweiterungsangriff auf der MIP

Die Forscher der ETH Lausanne haben untersucht, ob diese Art des Angriffs auf der Mobility Insights Plattform möglich ist. Sie stellten fest, dass zwei Abfragen ausreichen, um Regionen zu finden, in denen es nur einen Benutzer gibt, und sie konnten beweisen, dass der Angreifer diesen Schluss in beliebigen Regionen ziehen kann.

Bilder: EPFL

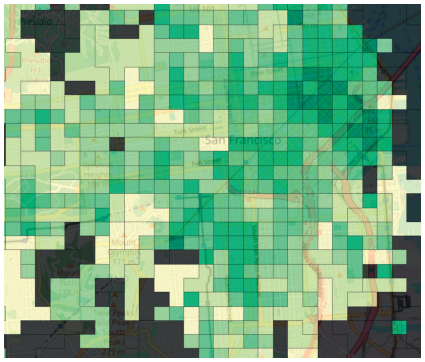

Bild 2 Auswahl vordefinierter Gebiete (Beispiel entsprechend der sogenannten «Zusammenlegungsstrategie»).

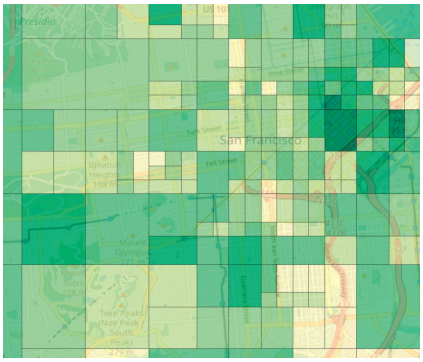

Bild 3 Teilungsstrategie: beispielhaftes Ergebnis nach Anwendung der «Aufteilungsstrategie».

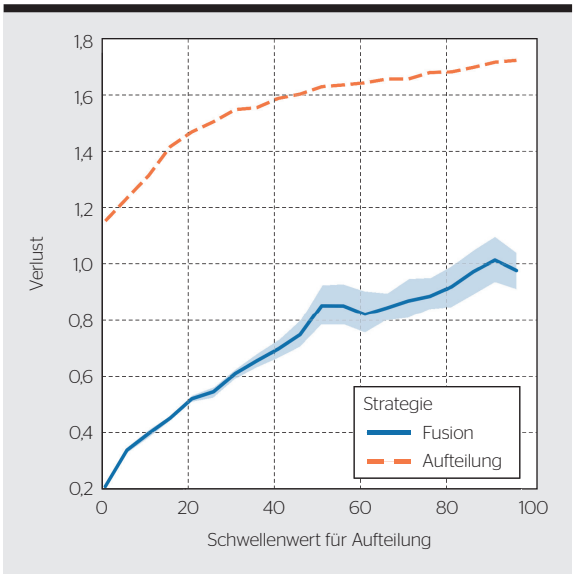

Bild 4 Vergleich der jeweiligen Nutzbarkeitsverluste beider Teilungsstrategien in Abhängigkeit vom für die Teilung herangezogenen Schwellenwert (Berechnungszeit für die Zusammenlegungsstrategie: 45,6 s; Berechnungszeit für die Aufteilungsstrategie: 4,1 s).

Daraufhin wurde ein Szenario untersucht, bei dem der Angreifer jemanden ins Visier nimmt, dessen Wohnadresse er kennt. Dazu wurde ein Erweiterungsangriff auf die Wohnadresse eines Mitarbeiters des Labors durchgeführt: So konnte an einem Morgen im März 2018 eine seiner Bewegungen innerhalb von Lausanne ermittelt werden.

Konkret wurden die möglichen Kombinationen von Abonnenten und Marktanteilen untersucht, die den angegebenen Ergebnissen entsprechen könnten (auch nach dem Expansionsangriff). Dann wurde ein kombinatorischer Ausdruck abgeleitet, der die Eintrittswahrscheinlichkeit der verschiedenen Ausgabewerte auf der Basis einer realen Anzahl von Swisscom-Abonnenten bestimmt. Dabei zeigte sich, dass bei Werten unter k (20) ein Angreifer die tatsächliche Zahl der Abonnenten mit hoher Wahrscheinlichkeit vorhersagen konnte. Folglich bietet dieser Mechanismus keinen zusätzlichen Schutz der Privatsphäre der Abonnenten.

Möglichkeiten, Angriffe zu vermeiden

Eine vielversprechende Lösung, um bei gleicher Nutzbarkeit und Relevanz der Ergebnisse einen besseren Schutz der Kundendaten zu erreichen, wäre eine Einschränkung der Abfragemöglichkeiten auf der Plattform. Zum Beispiel, indem Abfragen nur in vordefinierten, quadratischen Regionen zugelassen sind **(Bild 2)**.

Dieser Ansatz lässt sich wie folgt begründen: Halten sich in einer Region nur wenige Personen auf (z. B. in den Bergen), werden die Nutzer der Plattform weniger an Abfragen interessiert sein, die nur kleine Gebiete betreffen, da sich dort mit hoher Wahrscheinlichkeit nur wenige Personen befinden. Die Plattform kann in diesen Regionen also gröbere Unterteilungen vornehmen, die den meisten Abfragen entsprechen. An üblicherweise stark frequentierten Orten kann die Plattform feinere Unterteilungen anbieten, da auch diese dort mit hoher Wahrscheinlichkeit einen Wert zurückliefern.

Um jegliches Datenschutzproblem auszuschliessen, wäre es sinnvoll, die Unterteilungen auf der Grundlage aggregierter historischer Daten vorzunehmen. Da sich die Modelle wahrscheinlich selten ändern werden, können dieselben Unterteilungen lange Zeit verwendet werden.

Jede Teilungsstrategie verfolgt ein doppeltes Ziel: Einerseits muss sie den Datenschutz gewährleisten: Sie muss verhindern, dass Hacker den Aufenthaltsort einzelner Personen (oder kleiner Personengruppen) ermitteln können. Somit muss sie Erweiterungsangriffe abwehren können. Andererseits muss die Nutzbarkeit der Plattform erhalten bleiben: Die Plattform muss ihren Kunden weiterhin gestatten, präzise Informationen über ihre Beobachtungsgebiete zu erhalten.

Definition der Gebiete

Zur Erstellung der Unterteilungen werden zwei Strategien angeboten. Mit diesen beiden Methoden, Zusammenlegung und Aufteilung, lässt sich die Nutzbarkeit gegen das Risiko einer Neuauthentifizierung abwägen. Ausgangspunkt beider Strategien ist ein Schwellenwert für die Mindestanzahl Personen, die ein einziges Gebiet im Durchschnitt enthalten muss.

Die **Zusammenlegungsstrategie** geht von einer feinen Unterteilung aus. Solange es innerhalb dieser Unterteilung Gebiete gibt, in denen sich weniger Personen als durch den Schwellenwert definiert aufhalten, wird eines dieser Gebiete zufällig ausgewählt und mit einem benachbarten Gebiet zusammengelegt (auch zufällig bestimmt). Die Anzahl der für das neue Gebiet angezeigten Personen entspricht der Summe der Personen in den beiden zusammengelegten Gebieten. Der Prozess ist abgeschlossen, wenn in jedem Gebiet die Zahl der Menschen über der Schwelle liegt. **Bild 2** zeigt das Ergebnis dieser Strategie.

Es wurde beschlossen, die zu verschmelzende Region und ihren Nachbarn nach dem Zufallsprinzip auszuwählen, aber es könnten auch andere

Teilungsstrategie	0	*k*/2	*k*
Zusammenlegung	0,793	0,789	1,121
Aufteilung	1,50	1,496	1,558

Tabelle Nutzbarkeitsverlust für die zwei Strategien, wenn der Schwellenwert *k* in einer Region nicht erreicht wird. Im Idealfall wäre der Verlust = 0.

Strategien in Betracht gezogen werden, wie z. B. die Auswahl der Regionen, in denen sich die wenigsten Personen aufhalten; ihre Umsetzung bedeutet jedoch zusätzlichen Aufwand und kostspieligere Berechnungen und dürfte zu einem sehr ähnlichen Ergebnis führen.

Die **Aufteilungsstrategie** beginnt im Gegensatz dazu mit einem Rechteck, das die gesamte Karte als ein einziges Gebiet abdeckt. Dann werden schrittweise jeweils alle Gebiete in vier Teilgebiete unterteilt, und es wird die Anzahl der Personen in den jeweiligen neuen Gebieten ermittelt. Halten sich darin ausreichend viele Personen auf (d.h. mehr Personen als durch den Schwellenwert vorgegeben), wird die Aufteilung übernommen. Andernfalls wird die vorherige Aufteilung wieder hergestellt und die verbleibenden Gebiete werden weiter geteilt. Die Aufteilung in vier Teilgebiete (statt in zwei) wurde aus praktischen Gründen gewählt: Sie erleichtert die Umsetzung und führt zu Gebieten, deren Formen einfacher zu handhaben sind. **Bild 3** zeigt das Ergebnis dieser Strategie.

Nach Anwendung einer der genannten Strategien kann Swisscom die Gebiete den Kundenbedürfnissen entsprechend manuell anpassen. So können beispielsweise personalisierte Gebiete geschaffen werden, indem erzeugte Gebiete zusammengelegt oder aufgeteilt werden. Zu beachten ist jedoch, dass einmal abgeschlossene Unterteilungen langfristig bestehen bleiben können (beispielsweise während eines Jahres).

Welche Strategie wählen?

Die Wahl der Teilungsstrategie erfolgt auf Grundlage der Nutzbarkeit, der Berechnungskosten und schliesslich dem zu erreichenden Datenschutz. Datenschutzüberlegungen werden weiter unten erörtert. Im Hinblick auf Nutzungs- und Rechenkosten wurde eine Reihe vorläufiger Experimente auf der Grundlage eines öffentlichen Datensatzes mit 313 289 Positionsbestimmungen in San Francisco durchgeführt.

Der Nutzbarkeitsverlust der Plattform wurde anhand der Verteilungsdistanz zwischen der in einer Reihe von Zonen angegebenen Anzahl von Personen vor und nach der Teilung gemessen (siehe untenstehendes Beispiel einer Teilungskombination). Konkret wurde als Einstiegspunkt für den Zusammenlegungsvorgang ein feines Raster herangezogen (Gebiete mit einigen Häuserblocks) und der Wert der einzelnen Zellen wurde mit dem des Gebiets verglichen, das sie nach dem Unterteilungsvorgang enthielt. **Bild 4** zeigt, dass die Zusammenlegungsstrategie auf Kosten eines höheren Berechnungsaufwands für das Raster zu einer besseren Nutzbarkeit führt.

Von der Plattform angezeigte Ergebnisse

Nachdem die Gebiete festgelegt sind, kann sie der Kunde einzeln abfragen und erhält von der Plattform die Anzahl der entsprechenden Personen. In einigen Fällen und für bestimmte Zeitfenster kann es jedoch vorkommen, dass die im vorherigen Schritt erstellten Gebiete nicht genügend Personen enthalten, um sicher genutzt werden zu können. Dies liegt daran, dass dort manchmal, beispielsweise nachts oder zur Urlaubszeit, weniger Personen unterwegs sind. Die Plattform liefert daher nur dann einen Wert zurück, wenn im vordefinierten Gebiet mehr als *k* Benutzer (also mehr als 20) gezählt werden. Es ist jedoch wichtig, festzulegen, was die Plattform anzeigen muss, wenn die Anzahl der Personen in einem Gebiet kleiner als *k* ist.

Einerseits sollten die Ergebnisse der Plattform aus Datenschutzsicht unabhängig von der tatsächlichen Anzahl der Personen im Gebiet sein. Andererseits darf dies die Benutzerfreundlichkeit der Plattform nicht reduzieren.

Hier einige Möglichkeiten, die sich unter diesen Umständen anbieten: Die Plattform kann einen von *k* abhängigen Wert anzeigen (beispielsweise den Schwellenwert oder die Hälfte dieses Werts) oder sie kann «0» (also «niemand») oder einen auf Grundlage der Anzahl Personen in benachbarten Gebieten interpolierten Wert anzeigen. Die Tabelle vergleicht für die genannten Fälle den Nutzbarkeitsverlust, ausgedrückt als mittlere Differenz zwischen den Summen der für die einzelnen Gebiete angezeigten Zahlen vor und

Klarstellungen

Anmerkung von Swisscom

Die Arbeit des Teams von Carmela Troncoso hat eine Möglichkeit aufgezeigt, die von unserer Plattform genutzten Datenschutzvorkehrungen zu umgehen. Genau das ist das Ziel dieser Art von Übungen, die regelmässig wiederholt werden müssen. Die Bedingungen, unter denen diese Ergebnisse erzielt wurden, müssen verdeutlicht werden, da sie nicht der «normalen» Nutzung der Plattform entsprechen. Tatsächlich wurde dem EPFL-Team eine dedizierte Version zur Verfügung gestellt, die genau die gleichen Daten wie die Produktionsplattform verwendete, aber die Definition einer unbegrenzten Anzahl von Zonen und den direkten Zugriff auf die Front-End-API erlaubte. Nach Ansicht des EPFL-Teams war es nicht so sehr die Anzahl der Zonen (theoretisch reichen zwei Zonen aus, um die Erweiterung erfolgreich anzugreifen), sondern die Möglichkeit, direkt auf die API zuzugreifen, die die Angriffszyklen erheblich verkürzte und zu dem erzielten Ergebnis führte. In der Produktion und im Anschluss an die Studie implementierte das Team sofort provisorische Lösungen zur Vermeidung des Angriffs, während es an einem automatischen Prozess arbeitete, der dieses Problem endgültig lösen würde. Wir wissen, dass wir beim Datenschutz keinen Raum für Fehler haben und verbessern das Produkt Tag für Tag weiter.

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

Bild 5 Kombination von Gebieten.

nach der Unterteilung (siehe das folgende Beispiel für den Fall der Kombination von Gebieten). Selbst in diesem Fall führt die Zusammenlegung zu einer besseren Nutzbarkeit als die Aufteilung. Zudem scheint es hinsichtlich der Nutzbarkeit die beste Wahl zu sein, den Wert $k/2$ zurückzuliefern.

Greifen wir einen Fall mit einer Kombination mehrerer Gebiete heraus: Fragt ein Kunde mehr als ein Gebiet ab, muss das Ergebnis der Plattform für die einzelnen Gebiete von der Kombination unabhängig sein. Nehmen wir beispielsweise an, dass ein Kunde auf der Karte in **Bild 5**, die Gebiete A (40 Personen), B (15 Personen) und C (45 Personen) auswählt. Auch wenn die Gesamtzahl der Personen im Gesamtgebiet (A, B, C) grösser ist als k, muss die Plattform zunächst die Anzahl Personen in B entsprechend einer der oben genannten Strategien ersetzen und dann die Gesamtzahl für die drei Gebiete berechnen. Ist als Strategie beispielsweise gewählt, die Hälfte des Schwellenwerts zu liefern, zeigt die Plattform Folgendes an: $40 + k/2 + 45 = 95$ (anstatt $40+15+45=100$). Dies ist ein kritisches Element, um Angriffe ähnlich dem Erweiterungsangriff abzuwehren.

Analyse des Datenschutzes

Die Verwendung eines auf einem Raster basierenden Systems verhindert grundsätzlich Erweiterungsangriffe. Tatsächlich können keine Gebiete frei gewählt werden. Somit kann die Auswahl auch nicht beispielsweise um ein Haus erweitert werden. Es sind jedoch Angriffe denkbar, bei denen verschiedene Kombinationen vordefinierter Gebiete über die Plattform abgefragt werden und aus den Ergebnissen Informationen über individuelle Fahrten abgeleitet werden können. Die Untersuchung, inwieweit derartige Angriffe möglich sind, stellt einen weiteren Forschungsgegenstand dar. Zu beachten ist jedoch, dass die in diesem Artikel empfohlene Abgrenzungsstrategie bereits verhindert, gezielt einzelne Gebäude oder Beobachtungsgebiete zu betrachten, selbst wenn der Angreifer Bewegungen isolieren könnte.

Referenz

[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

Autoren

Giovanni Cherubin ist Postdoktorand am Security and Privacy Engineering Lab der ETH Lausanne.
→ EPFL, 1015 Lausanne
→ giovanni.cherubin@epfl.ch

Bogdan Kulynych ist Doktorand am Security and Privacy Engineering Lab.
→ bogdan.kulynych@epfl.ch

Marion LeTilly ist Studentin und verfasst ihre Masterarbeit am Security and Privacy Engineering Lab.
→ marion.letilly@epfl.ch

Carmela Troncoso ist Assistenzprofessorin (Tenure-Track) am Security and Privacy Engineering Lab.
→ carmela.troncoso@epfl.ch

La version française de cet article est parue dans le Bulletin 8/2020.