Schutz der Bewegungsprofile: Geht das?

Prof. Dr. Eric Dubuis
Institutsleiter Research Institute for Security in the Information Society RISIS, BFH

Die Politik denkt wegen häufiger Überlastung der Verkehrsträger (ÖV, Privatverkehr) Mobility Pricing nach. Mobility Pricing heisst, dass aus den Bewegungen eines Nutzers ein Preis für seine Mobilität gebildet wird. Dazu müssen Daten der Art, wer wann wo gewesen ist, erhoben werden. Führt das zwingend zum gläsernen Kunden?

Bewegungsprofile sind Informationen der Nutzer von Verkehrssystemen, die geschützt werden müssen. Es stellt sich die Frage, ob Mobility Pricing mit der Forderung, Bewegungsdaten zu schützen, überhaupt vereinbar ist. Die nachfolgend beschriebenen Ansätze zeigen Lösungen mit keinem, mittlerem und gutem Schutz der Bewegungsprofile.

Kosten aus Bewegungsdaten

Ein Mobility-Pricing-System besteht mindestens aus einem Erfassungsgerät (entweder pro Nutzer oder aber als Terminal im Bus oder in der Bahn) und einem zentralen Server beim Gebührenerheber. Weitere Komponenten wie Smartcard, Smartphone oder Komponenten Dritter kommen fallweise dazu oder verschmelzen mit ihnen (z. B. Smartphone und Terminal). Nutzer der Verkehrsträger erzeugen während einer Bemessungsperiode Bewegungsdaten $x_i = \langle id_i, loc_i, ts_i \rangle$, $1 \le i \le n$, wobei id_i ein Identifikator (je nach Lösung variiert die Bedeutung), loc_i eine Ortsangabe (GPS-Koordinaten oder Identifikator eines Terminals) und ts_i die Zeitangabe der Erhebung i bedeuten. Sei $\mathbf{x} = \langle x_1, x_2, \ldots, x_n \rangle$ die Zusammenfassung von n Erhebungen von Bewegungsdaten. Gefordert ist also eine Funktion, welche aus dem Eingabewert $\mathbf{x}$ die Fahrkosten $K(\mathbf{x})$ unter Berücksichtigung des Verkehrsnetzes, der Tarifzonen und der Tarifzeiten berechnet:

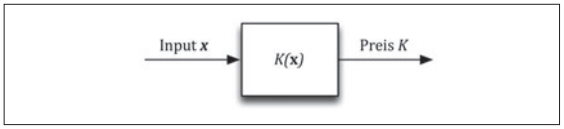

Abbildung 1: Funktion, welche als Input einen Vektor der Bewegungsdaten erhält und daraus den Preis K berechnet.

Betrachten wir nun drei verschiedene Lösungen und überlegen uns, welchen Schutz sie für die Bewegungsprofile der Nutzer bedeuten. Alle verwenden ein Erfassungsgerät (allerdings mit unterschiedlicher Funktionalität), von dem wir annehmen, dass es korrekt funktioniert.

Einfache Lösung

Bei der ersten Lösung erfasst das Erfassungsgerät ab Antritt der Reise periodisch die x_i-Werte und schickt sie an den zentralen Server des Gebührenerhebers.

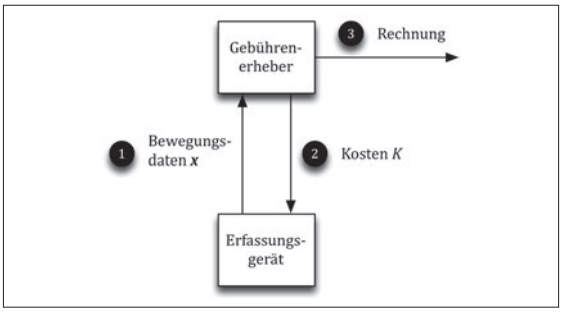

Abbildung 2: Einfache Lösung, bei der die Bewegungsdaten direkt dem Gebührenerheber geschickt werden.

Am Ende der Bemessungsperiode werden die Bewegungsdaten pro Nutzer unter Berücksichtigung des Verkehrsnetzes, der Tarifzonen und der Tarifzeiten ausgewertet, das heisst $K(\mathbf{x})$ berechnet und die errechneten Fahrkosten in Rechnung gestellt. Es ist offensichtlich, dass mit dieser Lösung der Gebührenerheber die Bewegungsprofile sämtlicher Nutzer kennt. Vorteil dieser Lösung ist: $K(\mathbf{x})$ ist einfach zu implementieren.

Schutz durch Pseudonyme

Eine andere Lösung baut erstens auf der Idee auf, dass die Bewegungsdaten x_i nicht mit einem Identifikator versehen werden, welcher die wahre Person identifiziert, sondern mit einem Pseudonym. Zweitens werden die Bewegungsdaten einem Gebührenrechner geschickt, der die fällige Gebühr $K(\mathbf{x})$ berechnet, ohne zu wissen, wem sie gehören. Zudem signiert er diesen Wert, sodass

ihn niemand unbemerkt ändern kann. Das Erfassungsgerät schickt die signierten Fahrkosten $K(x)$ mit der wahren Identität des Nutzers an den Gebührenerheber, der folglich die Rechnung für den Nutzer erhebt.

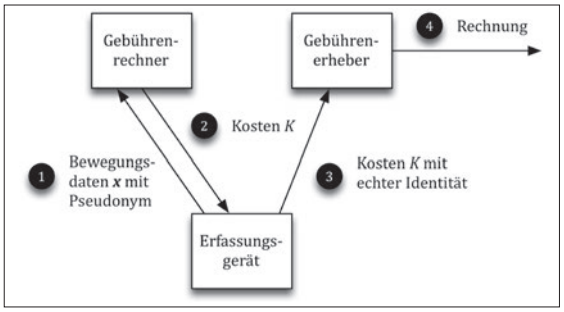

Abbildung 3: Die Bewegungsdaten werden dem Gebührenrechner geschickt, welcher die Kosten K berechnet und dem Erfassungsgerät zurückgibt. Am Ende der Bemessungsperiode wird K mit der echten Identität versehen und dem Gebührenerheber geschickt.

Mit dieser Lösung kennt der Gebührenerheber die Bewegungsprofile seiner Nutzer nicht, solange er nicht mit dem Gebührenrechner kooperiert. Der Gebührenrechner hingegen kennt die Bewegungsprofile mit den Pseudonymen, nicht aber mit den Identitäten der Nutzer. Auf den ersten Blick scheint also der Schutz des Bewegungsprofils eines Nutzers gewährleistet zu sein. Man weiss aber, dass mit Pseudonymen versehene Bewegungsprofile, kombiniert mit anderen Daten, zum Beispiel mit dem Wissen, wer wo wohnt und wo arbeitet, auf die Nutzer zurückgeführt werden können. Bei dieser Lösung muss man also dem Gebührenrechner ähnlich grosses Vertrauen aussprechen wie dem Gebührenerheber bei der ersten Lösung.

Grösstmöglicher Schutz durch Rundenprotokoll

Ein ganz anderer Ansatz wird mit der letzten Lösungsvariante vorgestellt. Als Basis dient die Annahme, dass das Erfassungsgerät leistungsstark ist und komplexe Berechnungen genügend schnell durchführen kann. Zudem besitzt es genügend grosse Speicherkapazität. Moderne Smartphone erfüllen diese Annahme.

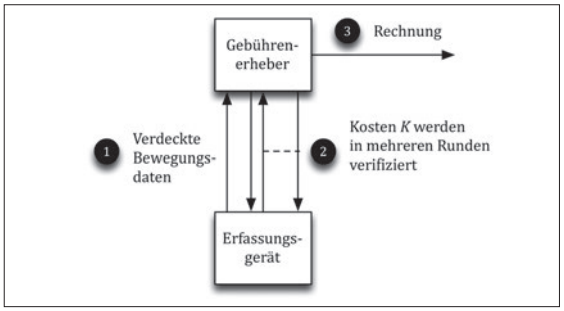

Abbildung 4: Die verdeckten Bewegungsdaten werden laufend an den Gebührenerheber geschickt. Am Ende der Bemessungsperiode berechnet das Erfassungsgerät die Kosten K und teilt sie dem Gebührenerheber mit. Der Gebührenerheber verifiziert die Korrektheit von K mit einem Rundenprotokoll.

Die grundlegende Idee ist, dass das Erfassungsgerät die Bewegungsdaten x_i nach wie vor erstellt und lokal speichert. Eine verdeckte Form der Bewegungsdaten $\bar{x}_i = c(x_i)$ wird laufend oder periodisch, spätestens aber vor Ende der Bemessungsperiode, dem Gebührenerheber geschickt. Die Funktion c ist so beschaffen, dass der eingepackte Wert x_i im Wert $\bar{x}_i$ erst ausgelesen werden kann, wenn das Erfassungsgerät die entsprechende Freigabe erteilt. Die Werte x_i werden somit verbindlich, aber unlesbar beim Gebührenerheber erfasst. Am Ende der Bemessungsperiode werden in einem ersten Schritt die Fahrkosten $K(x)$ durch das Erfassungsgerät bestimmt und dem Gebührenerheber mitgeteilt. In einem zweiten Schritt beweist das Erfassungsgerät dem Gebührenerheber in einem Rundenprotokoll (siehe Kasten), dass es die Fahrkosten richtig berechnet hat. Ist das der Fall, so stellt der Gebührenerheber die Rechnung dem Kunden.

Bei den vorgestellten Lösungsansätzen müssen noch Massnahmen für den Fall ergriffen werden, dass der Nutzer das Erfassungsgerät gar nicht verwendet. Aus Platzgründen wird darauf nicht näher eingegangen.

Fazit

Die Lösungsansätze zwei und drei zeigen, dass man mit geeigneten organisatorischen und kryptografischen Massnahmen den Schutz der Bewegungsprofile gewährleisten kann. Allerdings gilt: Mehr Schutz bedingt komplexere Lösungen.

Kontakt
– eric.dubuis@bfh.ch

Infos
– risis.bfh.ch

Rundenprotokoll

In einer festen Anzahl von Runden versucht der Gebührenerheber zu eruieren, ob die vom Erfassungsgerät berechneten Fahrkosten korrekt sind. Dazu stellt er dem Gerät Fragen, die das Gerät beantwortet, ohne dass es die aufgezeichneten Bewegungen bekannt gibt. Sind die Antworten richtig, dann hat das Gerät die Fahrkosten mit hoher Wahrscheinlichkeit korrekt berechnet. Ist hingegen eine Antwort falsch, so weiss der Gebührenerheber, dass das Erfassungsgerät betrogen hat.