

Wirkungsmessung Unfallprävention: Mit Bewegungsdaten Verhaltensänderungen nachweisen

Themenbereich: Data Science

Student: Alexander Becker

Dozent: Andreas Streich

Experte: Philippe Gassmann

Wirtschaftspartner: Suva Luzern

Keywords: Data Science, Mobile Technologies, Externe Datenquellen, Business Intelligence, Wirkungsmessung

1. Kurzfassung der Aufgabenstellung

Die Schweizerische Unfallversicherungsanstalt SUVA ist der grösste Unfallversicherer der Schweiz. In regelmässigen Abständen lanciert die SUVA Sensibilisierungskampagnen, um die Versicherten zu einem risikobewussten Verhalten in Beruf und Freizeit zu animieren. Als konkrete Anwendungsfälle standen Verhaltensmessungen im Sport- und Freizeitbereich im Zentrum. Bisher findet die Wirkungsmessung dieser Kampagnen vorwiegend über Umfragen bei den Versicherten statt. Die Umfragen leiden jedoch unter einer geringen Rücklaufquote. Da diese Umfragen zudem teuer sind und durch die Selbstselektion von Selection Bias betroffen sind, soll ergänzend zu den Umfragen die Nutzbarkeit von Bewegungsdaten evaluiert werden. Alle mobilen Geräte teilen vielerlei Daten mit diversen Stellen wie Herstellern und Dienstleistern. Könnten diese Daten zur Wirkungsmessung genutzt werden, so wäre es der SUVA möglich ihre Wirkungsmessung zu präzisieren und ihre Kampagnen nach deren Wirkung zu steuern. Es soll daher eine Machbarkeitsstudie zur Nutzung dieser Daten durchgeführt werden im speziellen Kontext zur Wirkungsmessung.

2. Vorgehen/Lösungskonzept

Da diese Arbeit explorativ gestaltet wurde, war eine agile Vorgehensweise notwendig. Dabei wurden die Projektschritte und Lösungsvarianten anhand der Ergebnisse der vorherigen Schritte bestimmt.

Zu Beginn der Arbeit galt es, die technische Machbarkeit der Auswertung von Bewegungsdaten zum Zweck der Verhaltensmessung zu prüfen. Dafür sollen aus Bewegungsdaten, welche einen Standort auf einer Zeitachse enthalten, mit Hilfe von Geodaten wie Karten, Rückschlüsse auf das Bewegungsprofil gezogen werden. Wenn dies gelingt, sollte der SUVA ein Überblick über die ihr verfügbaren Optionen geboten werden, um Bewegungsdaten zur Wirkungsmessung einzusetzen. Anhand von Recherchen mittels Interviews und Literatur sollen Erkenntnisse erlangt werden, die dies ermöglichen. Zu durchführbaren Optionen soll dann ein Use-Case definiert werden.

3. Spezielle Herausforderungen

Eine der grössten Herausforderungen war die Arbeit mit den Kartendaten. So wurde zu Beginn versucht eine Schnittstelle zu nutzen. Die Nutzung dieser stellte sich aber nach längerer ausführlicher Untersuchung als nicht zweckmässig heraus. So werden nur die ersten 200 Koordinaten einer Route ausgegeben. Das bedeutet das grössere Netzwerke aus mehreren Strecken mit nur einem Aufruf der Schnittstelle nicht abgerufen werden können. Auch nachdem die Kartendaten als ganzer Datensatz zur Verfügung standen, gestaltete sich die Arbeit mit diesen nicht trivial. Aus Performancegründen ist es notwendig mit Teilmengen des Kartendatensatzes zu arbeiten. Dieser ist in seiner Grundform mehrere Gigabyte gross und beinhaltet viele Daten, die für die Nutzung nicht notwendig sind. Dadurch ist das Berechnen dieser Teilmengen des Datensatzes sehr rechenintensiv und entsprechendes Grundlagenwissen zum Umgang mit georeferenzierten Datensätzen musste erst erlangt werden. Zudem waren das Auslesen und Bearbeiten der XML Daten, welche die Bewegungsdaten darstellen, sehr aufwendig, da diese oft nicht einem einheitlichen Schema folgen, sondern sich von Hersteller zu Hersteller unterscheiden.

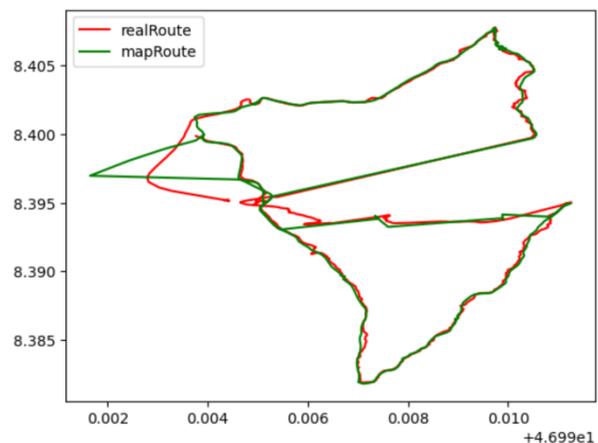
Letztlich war auch das Erlangen der notwendigen Informationen über die Databroker sehr schwierig. Zum einen, weil es hier nur sehr wenig Fachliteratur gibt, zum anderen, weil diese Unternehmen auch nur wenig von sich preisgeben.

4. Konkrete Ergebnisse

Im Rahmen dieser Bachelorarbeit wurden die folgenden Ergebnisse erarbeitet:

- **Use-Case mit Hilfe von Python:** Mit Hilfe eines Pythonscripts wurde die technische Machbarkeit unter Beweis gestellt. Dabei wurden eigene Bewegungsdaten erstellt und Kartendaten genutzt.
- **Evaluation von Datenquellen:** Um die Bewegungsdaten zu erhalten, wurden verschiedene Quellen evaluiert. Dabei wurden Anbieter von Bewegungsdaten sowie weitere Technologien, um Datensammlungen entweder durch die SUVA selbst oder durch Dritte durchzuführen, untersucht.
- **Compliance:** Da durch Bewegungsdaten auch auf die Bewegungsprofile einzelner Personen rückgeschlossen werden kann, wird beleuchtet, welche Massnahmen die SUVA treffen muss, um sich dem Schweizer Datenschutzgesetz (DSG) konform zu verhalten.

Zu Beginn wurde überprüft, ob aus Bewegungsdaten gemessen werden kann, ob jemand von einem offiziellen Wanderweg abweicht. Dazu wurden im Rahmen eines Versuches Bewegungsdaten generiert. Anschliessend sollte anhand dieser Bewegungsdaten eine Abweichung zu einer Idealroute berechnet werden. Um die Idealroute, also die Route, die einem offiziellen Wanderweg entspricht, zu berechnen, wurden die vom Bundesamt für Landestopografie zur Verfügung gestellten Kartendaten aufbereitet. Anschliessend wurden die generierten Daten mit den Idealdaten verglichen und mittels mathematischer Methoden eine Abweichung zu diesen berechnet.



Die real gelaufene Route (rot) im Vergleich mit einem Wanderweg (grün)

Dabei wurde festgestellt, dass die technische Machbarkeit gegeben ist. Durch die erfassten Datenpunkte ist man in der Lage, die Abweichung zu einer Idealroute durch ein Script berechnen zu lassen. Gleichzeitig ist aufgefallen, dass man umfangreiche Fehlerkorrekturmethode implementieren muss, da die Bewegungsdaten mit Fehlern besetzt sein können.

Die anschliessende Evaluation von möglichen Datenquellen hat gezeigt, dass die General Data Protection Regulation (GDPR) und auch das Schweizer Datenschutzgesetz die Gewinnung von Daten sehr schwierig gestaltet. Bei umfangreichen Interviews mit Databrokern ist unisono geäussert worden, dass sich das Erfassen und Verwenden von Bewegungsdaten mittlerweile ebenso sehr schwierig gestaltet. So sind nur Datensätze mit einer geringen Genauigkeit, Auflösung oder Anzahl an mobilen Devices vorhanden. Das Nutzen von Bewegungsdaten erfasst durch Mobilfunkprovider, wie die Swisscom, ist zwar prinzipiell möglich, jedoch ist die Genauigkeit äusserst gering. Auch lassen sich hier keine Einzeldaten darstellen. Diese werden nur aggregiert zur Verfügung gestellt.

Weitere evaluierte Technologien, wie zum Beispiel W-LAN, Bluetooth Low Energy, oder NFC eignen sich je nach Use-Case zum Erfassen von Daten. Eine generelle Erkennung von Human Presence ist ohne weiteres möglich. Da die Hersteller aus Datenschutzgründen das Teilen von Bewegungsdaten immer restriktiver handhaben, wird das genaue Verfolgen von einzelnen Devices jedoch zunehmend verunmöglicht.

Bezüglich des Datenschutzes wurde sich am Schweizer Datenschutzgesetz orientiert. Bei dem Nutzen von Databrokern ist vorab zu prüfen, ob diese den Datenschutzbestimmungen genügen. Zwar ist das DSG hier nicht deckungsgleich mit dem GDPR, aber wenn die Daten dem GDPR genüge tragen, so ist die Wahrscheinlichkeit sehr hoch, dass diese auch dem DSG entsprechen.

Wenn die SUVA selbst Daten erfassen möchte, muss diese sich bewusst sein, dass bereits das Erfassen der MAC-Adresse, wie auch einer IP, bereits das «Verarbeiten» von Personendaten darstellt. Dabei müssen auch hier rechtliche Vorabklärungen getroffen und Massnahmen durchgeführt werden, um diese Daten nicht umkehrbar zu anonymisieren.

5. Ausblick

Die Arbeit hat gezeigt, dass es grundsätzlich möglich ist Bewegungsdaten zu verwenden. Jedoch kommt es hierbei auf den Use-Case an, ob die Technologien und Varianten gewinnbringend eingesetzt werden können.

Der SUVA wurde mit dieser Arbeit eine Art «Handwerkszeug zur Nutzung von Bewegungsdaten» gegeben. Mit diesem ist die SUVA in der Lage ohne grosse Nachforschungsanstrengungen zu evaluieren, ob sich für ihren speziellen Use-Case der Einsatz von Bewegungsdaten lohnt. Da jedoch das Schweizer Datenschutzgesetz im September 2023 in überarbeiteter Form in Kraft tritt, ist hier auch die zukünftige Entwicklung zu beachten. Es ist jedoch davon auszugehen, dass dank immer umfangreicherer Datenschutzmassnahmen seitens der Hersteller und der Gesetzgeber eine Nutzung von genauen Bewegungsdaten immer unwahrscheinlicher wird. Das heisst die Nutzung von Bewegungsdaten wird nicht allumfassend möglich sein, sondern nur für bestimmte Use-Cases.