

Praktikumsbericht

KDDI R&D Laboratories, Japan
Juni-August 2016

Charlotte Burmeister

October 21, 2016

1 EINLEITUNG

Mein Praktikum habe ich bei KDDI R& D Laboratories in Japan absolviert. KDDI ist ein japanisches Telekommunikationsunternehmen mit Hauptsitz in Tokio. Ich habe in der Research and Development Abteilung gearbeitet, die ihren Sitz in Saitama hat, was circa eine Stunde außerhalb von Tokio liegt. Das Praktikum habe ich über den IAESTE bekommen. Der IAESTE ist eine Studentenorganisation die Praktia im Ausland vermittelt. Die Organisation gibt es weltweit und es werden Praktikumsstellen zwischen den Ländern ausgetauscht. Man bewirbt sich zuerst mit Motivationsschreiben und Lebenslauf beim IAESTE, dann gibt es eine Börse bei der man eine Stelle zugeteilt bekommt. Damit hat man die Praktikumsstelle schon relativ sicher, muss sich aber nochmal beim Arbeitgeber bewerben. Aber solange man keinen völligen Unsinn schreibt ist einem die Stelle sicher heißt es. Für mich war klar, dass ich mein Praktikum in Asien absolvieren möchte, da ich schon zuvor ein Jahr in Vietnam verbracht habe und mich die Gegend sehr interessiert. Nach Japan wollte ich schon länger und die Arbeitsstelle klang spannend, sodass ich mich sehr gefreut habe, als ich die Stelle bekommen habe.

2 TÄTIGKEIT

Mein Praktikum habe ich in der Privacy Abteilung von KDDI R&D Labs. gemacht. Dort habe ich zu Algorithmen zur Sicherung der Privatsphäre in Datenbanken untersucht. Hierzu eine kleine Erklärung: Durch die zunehmende Digitalisierung der Gesellschaft nimmt auch der Umfang an gesammelten Daten zu. Im Zuge dessen werden auch intime Details, wie z.B. Krankheitsdaten gespeichert. Diese Datenbanken werden aber ausgewertet um Statistiken zu erstellen und Erkenntnisse zu gewinnen. Standardmäßig werden eindeutig identifizierende

Attribute wie Name, Adresse, Krankenversicherungsnummer, etc. vor der Publikation entfernt. Dies allein sichert aber nicht die Privatsphäre, denn durch verschiedenen Attacken gibt es die Möglichkeit Individuen zu identifizieren. Eine Option ist es z.B. durch sogenanntes "Matching" verschiedenen Datenbanken zu verbinden und durch Attribute die in beiden Datenbanken vorkommen, einzelne Personen rauszufinden. Dies ist möglich wenn eine Attributkombination nur einmal vorkommt.

2.1 LÖSUNGSANSÄTZE UND MEINE ARBEIT

Um das zu verhindern, gibt es verschiedene Ansätze. Gleichzeitig aber sollen die Daten weiterhin verwertbar bleiben. Ziel jedes neuen Algorithmus ist es also ein optimales Verhältnis zwischen Sicherheit und Verwendbarkeit zu finden. Ich habe mich mit k-Anonymität und differential privacy auseinandergesetzt. Nachdem ich mich zuerst in k-Anonymität eingearbeitet habe und dort verschiedene Algorithmen und deren Effizienz verglichen habe, meinte mein Mentor, dass das nicht mehr so aktuell ist und ich mir auch noch differential privacy angucken soll. Nachdem ich mich auch in das Thema reingelesen habe, hat er mir ein Paper gegeben in dem k-Anonymität und differential privacy kombiniert wurde um die jeweiligen Nachteile auszugleichen und die Vorteile zu kombinieren. Zuerst aber eine kurze Erklärung zu den zwei unterschiedlichen Ansätzen:

2.1.1 K-ANONYMITÄT

K-Anonymität kann wie folgt definiert werden:

K-ANONYMITÄT Eine Tabelle erfüllt k-Anonymität wenn jede Kombination von Quasi-Identifizierern mindestens k-mal auftaucht.

Dafür brauchen wir noch die Definition der Quasi-Identifizierer.

QUASI-IDENTIFIZIERER Ein Quasi-Identifizierer (QI) ist ein Set von Attributen die wenn sie mit externen Informationen verlinkt werden, Detail über Individuen preisgeben.

Im Beispiel in der Einleitung ist der QI die PLZ, das Geburtsdatum und das Geschlecht, aber auch Name und Adresse, da diese Attribute auf jeden Fall die Identität preisgeben.

Um k-Anonymität zu erreichen gibt es viele Algorithmen. In dem Paper wurde der MDAV Algorithmus benutzt. Dieser funktioniert folgendermaßen:

Der Vorteil von k-Anonymität ist, dass sich eine Datenbank schnell und mit geringen Verlusten verschlüsseln lässt, d.h. die Daten können später gut ausgewertet werden. Der Nachteil ist die geringe Sicherheit, k-Anonymität war eine der ersten Methoden zur Verschlüsselung von Datenbanken und mittlerweile sind viele Attacken bekannt.

2.1.2 DIFFERENTIAL PRIVACY

Dieser Anonymisierungsmechanismus ist für ein interaktives Szenario entworfen und sitzt zwischen dem Anwender und der Datenbank. Interaktiv heißt, dass der Anwender eine Frage an die Datenbank stellt und dann eine Antwort von der Datenbank erhält. Die Antwort wird

dann durch den Algorithmus verändert, sodass die Sicherheit gewährleistet ist. Differential privacy leistet eine probabilistische Garantie, dass der Einfluss eines Individuums auf eine Antwort limitiert ist, d.h. es soll nicht erkennbar sein, ob ein Individuum in der Antwort enthalten ist oder nicht.

Um differential privacy zu erreichen, wird oft Geräuschaddition verwendet. Bei der Geräuschaddition wird eine Zufallsvariable zur (numerischen) Antwort hinzugefügt. Im Paper wurden Zufallsvariablen aus der Laplacedistribution verwendet.

Der Vorteil von differential privacy ist, dass es eine hohe Garantie für die Sicherheit gibt. Der Nachteil ist die geringe Verwendbarkeit der Daten, denn ja nach Beschaffenheit der Daten, kann die Addition von Geräuschen zu starker Verzerrung führen.

2.1.3 KOMBINATION DER BEIDEN ALGORITHMEN

Die Idee der Autoren des Papers von J. Soria-Comas et al., „Enhancing Data Utility in Differential Privacy via Microaggregation-based k-Anonymity“ ist, die beiden Methoden zu kombinieren und so die Nachteile auszugleichen. Dafür soll der Datensatz zuerst mithilfe von Mikroaggregation k-anonymisiert werden und dann wird zu jedem Eintrag in der Datenbank Laplacegeräusch addiert (eine Abwandlung des interaktiven Szenarios). Die vorherige Mikroaggregation des Datensatzes soll die Geräusche reduzieren die später hinzugefügt werden und damit die Verwertbarkeit der Daten erhöhen. Meine Aufgabe war es die These aus dem Artikel nachzuprüfen. Dafür habe ich den beschriebenen Algorithmus in Octave nachprogrammiert und zudem den MDAV „k-Anonymity“ Algorithmus und einen Algorithmus für „differential privacy“ der Laplace Zufallszahlen benutzt programmiert. Dann habe ich die drei Algorithmen anhand von drei Kriterien verglichen: das Offenlegungsrisiko, der Informationsverlust und eine Metrik die beides kombiniert. Zum Testen habe ich den öffentlich verfügbaren Datensatz "Adult" vom UCI verwendet. Leider konnte ich die Ergebnisse der Forscher nicht bestätigen. Bei mir hat der Kombi-Algorithmus keine Verbesserung zu den einzelnen Algorithmen gezeigt.

3 UMFELD

3.0.1 ARBEIT

Unsere Abteilung bestand aus acht Mitarbeiteren von denen zwei aus Indien und einer aus China kam. Ich war die einzige Praktikantin dort. Wir saßen in einem Großraumbüro wo die einzelnen Bereiche durch Regale abgetrennt waren. Unseren Bereich haben wir mit noch einer Abteilung geteilt. Am Anfang wurde mir ein Mentor zugeteilt, der mir meine Aufgaben zugeteilt hat und mir bei Problemen geholfen hat. Mein Mentor war sehr nett und stand für Fragen immer zur Verfügung. Allerdings gestaltete sich die Kommunikation manchmal schwierig, da sein Englisch nicht ganz so gut war. So war es schwer fachliche Fragen zu stellen und ich war auf mich allein gestellt eine Antwort im Internet zu finden. Das hat den Arbeitsprozess zum Teil sehr verlangsamt, da ich für kleine Fragen lange brauchte um eine verständliche Erklärung zu finden. Allgemein ist zwei Monate eine kurze Zeit für ein Praktikum, insbesondere wenn man sich in ein komplett neues Themenfeld einarbeiten muss. Ich denke die Aufgabe die

ich gestellt bekommen habe war für zwei Monate angemessen, aber mit mehr Unterstützung hätte ich sie noch gründlicher bearbeiten können.

Mit den anderen Mitarbeiter hatte ich nicht viel Kontakt, da ich meine Aufgaben unabhängig von deren Aufgaben erledigen konnte. Leider waren sie auch nicht so kommunikativ. Meine Mittagspause habe ich eigentlich immer mit meinem Mentor verbracht. Allerdings waren sie nicht nur mir gegenüber verschlossen, sondern haben auch untereinander nicht viel gemacht. Einige waren nur an manchen Tagen im Büro, da sie auch in der Hauptzentrale in Tokio gearbeitet haben oder unterwegs auf Konferenzen waren. Andere haben ihr Mittagessen am Schreibtisch gegessen und gar keine richtige Pause gemacht. Ich hätte mir eine bisschen lockere und offenere Arbeitsatmosphäre gewünscht und gerne mehr über meine Kollegen erfahren. Trotzdem waren alle nett und hilfbereit wenn man im Gespräch war. Ich habe mich auch nie unwohl oder ausgegrenzt gefühlt, da sich alle so zueinander verhalten haben. Etwas schade fand ich, dass ich erst am Ende erfahren habe, dass es noch einen deutschen Praktikanten bei KDDI R& D gibt. Ich hätte es interessant gefunden sich auszutauschen und fände es gut wenn KDDI dort Kontakte herstellen würde. Nichtsdestotrotz KDDI als Arbeitgeber würde ich insgesamt als gut bewerten. Meine Arbeitszeiten wurden eingehalten, ich habe mein Geld pünktlich bekommen, bei Problemen hatte ich einen Ansprechpartner und die Kommunikation im Vorfeld verlief reibungslos. Auch allgemein als Arbeitgeber in Japan ist KDDI glaub ich ganz gut. Dazu muss man im Vorfeld sagen, dass die Menschen in Japan sehr viel arbeiten. Überstunden gehören zum guten Ton, kaum Urlaub ist normal und Samstags ins Büro auch. Von meinen Kollgen weiß ich, dass sie "nur" einmal die Woche Überstunden machen und auch nicht Samstag arbeiten. Es gab extra jeden Mittwoch durchsagen, dass alle Mitarbeiter pünktlich gehen sollen (was nicht immer eingehalten wurde). Außerdem wurde es verboten im Büro zu übernachten(!). Anscheinend ist es noch in einigen Firmen gang und gäbe, dass die Leute solange arbeiten, dass es sich nicht mehr lohnt nach Hause zu gehen.

3.0.2 PRIVAT

Ein Ausgleich zu der eher ruhigen Atmosphäre auf Arbeit war der Kontakt zu japanischen Studenten und anderen Praktikanten. Dadurch dass es auch ein Lokalkomitee des IAESTE in Tokio gibt, findet man gleich Anschluss. Das Komitee besteht aus motivierten japanischen Studenten die viele Veranstaltungen organisieren auf denen man schnell die anderen Praktikanten kennenlernt. Unter anderem gab es japanische Kulturabende, Sightseeing, einen Kochabend und wir haben zusammen den Berg Fuji, Fuji-San auf Japanisch, bestiegen. Das war sehr anstrengend, aber auch sehr schön und definitiv lohnenswert. Sehr beliebt in Japan ist "All-You-Can-Drink" und natürlich Karaoke. Oft auch beides zusammen. Bei "All-You-Can-Drink" zahlt man einen festen Betrag und kann dann für 2 oder 3 Stunden soviel trinken wie man will, oft gibt es noch kleine Snacks dazu. Da Alkohol relativ teuer in Japan ist (wie so vieles) lohnt sich das meistens schon nach zwei Getränken. Meine Wohnung lag leider außerhalb Tokios und ich brauchte so eine Stunde rein in die Stadt. Zudem ist Bahnfahren relativ teuer in Tokio, eine Fahrt hat mich ca. 7€ gekostet. Trotzdem bin ich sehr oft in die Stadt gefahren, da es in meinem Vorort nicht wirklich was zu unternehmen gab und auch keine anderen Praktikanten dort gewohnt haben.

3.1 LEBEN IN JAPAN

Tokio ist eine aufregende Stadt in der man wahrscheinlich ein Jahr wohnen kann und jeden Tag etwas Neues entdecken kann. Es gibt tausende Restaurants, Spielhallen, Bars und allerlei verrückte Sachen. Leider ist Japan ziemlich teuer. Wie oben erwähnt ist der Transport in Tokio teuer. Alternativ kann man mit dem Rad fahren, wenn man das Glück hat zentral zu wohnen, allerdings gibt es nicht nur ein Zentrum in Tokio. Essen gehen fand ich im Verhältnis zu den Lebensmittelpreisen günstig. Eine Ramensuppe (Nudelsuppe mit Fleisch) bekommt man ab 6€ und wird davon ordentlich satt. Die Lebensmittel sind dagegen unfassbar teuer, ein(!) Apfel kostet so 2-3 €, eine Gurke 1-2€ und auch Reis ist nicht so billig wie man denken würde. Fisch ist dafür relativ günstig und sehr lecker! Kulinarisch hat Japan einiges zu bieten, allerdings wird es schwer wenn man kein Fleisch oder Fisch ist, denn selten gibt es vegetarische Alternativen. Mit Englisch kommt man durch, aber man darf nicht erwarten, dass jeder Englisch spricht. Auch junge Leute können oft nur schlecht Englisch. In größeren Städten ist auf den Bahnhöfen meist alles auch auf Englisch ausgeschildert, in kleinern Orten oft nur auf Japanisch. Dafür sind dort die Bahnhöfe kleiner und irgendwie findet man sich dann schon zurecht, da alles auch sehr strukturiert und organisiert ist. Dagegen ist es nicht so einfach sich auf den zwar auf Englisch ausgeschilderten aber sehr sehr großen Bahnhöfen in Tokio zurechtzufinden. Die Bahnhöfe zählen zu den größten der Welt und in Shinjuku (der belebtesten Station der Welt) steigen täglich etwas 3,6 Millionen Menschen um. Eine Eigenschaft beschreibt Japan sehr gut für mich: Japan ist "convenient", also bequem. Japaner verstehen es sich das Leben (in einigen Aspekten) möglichst angenehm zu machen. Da sind die Toiletten, die es in jedem, wirklich jedem, Bahnhof und Supermarkt kostenlos zur Nutzung gibt. Dann gibt es an jeder Ecke einen "Convenience Store", ein 24h-Laden der ziemlich gut ausgestattet ist. An jeder Ecke stehen auch Getränkeautomaten. Dann die japanischen Schnellzüge, die sehr schnell sind und sehr komfortabel, riesige Beinfreiheit (nicht nur für kleine Menschen wie mich) und breite Sitze.

4 FAZIT

Japan ist ein spannendes Land und Tokio eine noch spannendere Stadt. Es gibt sehr viel zu sehen und zu erleben. Die zwei Monate haben nicht ansatzweise ausgereicht um alles zu sehen. Definitiv möchte ich nocheinmal nach Japan reisen und mehr sehen. Mein Praktikum war fachlich interessant und ich habe viel gelernt. Es war eine interessante Erfahrung in einem japanischen Büro zu arbeiten. Ein Praktikum über den IAESTE zu machen kann ich jeden nur empfehlen, es ist eine tolle Möglichkeit Praktikum, Auslandserfahrung und Urlaub zu verbinden!