Université
franco-allemande
Deutsch-Französische
Hochschule

Introducing Privacy In Current Web Search Engines
by Albin Petit

Zusammenfassung

In den letzten Jahren sind durch den technologischen Fortschritt im Bereich des Sammelns, Speicherns und Verarbeitens von großen Datenmengen mit überschaubaren Kosten schwerwiegende Datenschutzbedenken entstanden. Viele Online-Dienste (z.B. Facebook, Google) besitzen nun die Möglichkeit detaillierte Profile ihrer Nutzer zu erstellen und mit gezielter Werbung anzusprechen. Die Veröffentlichungen von Edward Snowden im Mai 2013 brachten die tiefgreifende globale Überwachung durch die NSA ans Licht die zeigten wie Geheimdienste ebenfalls große Mengen an persönlichen Daten sammeln und nutzen. Tagtäglich werden die Aktivitäten von Nutzern von vielen Seiten ausspioniert, aber die meisten Nutzer sind sich des Ausmaßes der Überwachung nicht bewußt. Der Hauptgrund liegt darin, dass Online-Dienste ihren Nutzern nicht die Möglichkeit geben auf ihre gesammelten Daten zuzugreifen. Darüberhinaus informieren die Nutzungsbedingungen nur ungenügend über die Art der gesammelten Daten und den Zweck der Sammlung.

Datenschutz betrifft viele Bereiche, ist aber besonders wichtig für häufig besuchte Webseiten wie Suchmaschinen (z.B. Google, Bing, Yahoo!). Diese Dienste ermöglichen es Nutzern relevante Inhalte aus dem stetig wachsenden ``Datensee'' im Internet herauszuholen. Die gute Qualität ihrer Ergebnisse beruht auf der Nutzung persönlicher Daten. Als direkte Konsequenz ist es wahrscheinlich, dass Suchmaschinen sensible Informationen über einzelne Nutzer sammeln und speichern (z.B., Interessen, politische und religiöse Orientierung, Gesundheitszustände). In diesem Zusammenhang gewinnt die Entwicklung von Privatsphäre-schützenden Lösungen, die es Nutzern erlauben Suchmaschinen zu befragen, an Wichtigkeit.

In dieser Arbeit, präsentieren wir SimAttack als Angriff gegen existerende Lösungen um eine Suchmaschine in Privatsphäre-erhaltender Weise anzufragen. Dieser Angriff zielt darauf ab die ursprüngliche Nutzeranfrage wiederherzustellen, indem man ungeschützte Nutzeranfragen verwendet die der Angreifer zuvor gesammelt hat. Wir nutzen SimAttack um die Robustheit von drei repräsentativen State-of-the-Art Privatsphäre-erhaltenden Lösungen zu evaluieren. Wir zeigen, dass diese Lösungen die Privatsphäre der Nutzer nicht zufriedenstellend schützt.

Deshalb entwickeln wir PEAS als neuen Mechanismus, um den Schutz in Bezug auf SimAttack zu erhöhen. Dazu verfolgt PEAS zwei Ansätze: verbergen die Identität der Nutzer und maskieren der Suchanfragen. Ersteres wird erreicht durch Verschlüsseln und Versenden der Anfragen durch eine Folge über zwei Knoten, während letzteres Anfragen verschleiert indem sie mit mehreren imitierten Anfragen vermischt. Die Hauptherausforderung unseres Ansatzes ist es realistische, imitierte Anfragen zu erzeugen. Dazu werden Anfragen erzeugt, die von einem anderen Nutzer des Systems geschickt hätten werden können.

Letztlich präsentieren wir ein Verfahren um sensible Anfragen zu identifizieren. Unser Ziel ist es existierende Schutzmechanismen anzupassen um nur sensible Anfragen zu schützen und damit Ressourcen zu schonen (z.B. CPU, RAM). In der Tat braucht eine gewöhnliche Anfrage nach einem Kuchen-Rezept nicht das gleiche Ausmaß an Schutz wie eine Anfrage nach einer HIV-Infektion. Wir entwerfen zwei Module um sensible Anfragen zu identifizieren und kombinieren sie mit real-existierenden Schutzmechanismen. Wir zeigen empirisch, dass diese Modifikation die Leistung existerender Schutzmechanismus dramatisch verbessert.