

Herausforderung

Profile im Internet haben in der heutigen Zeit einen wichtigen Stellenwert eingenommen um eine Identität online zu repräsentieren. Des Weiteren hat die allgemeine Beteiligung an sozialen Netzwerken in den letzten Jahren drastisch zugenommen. Dies führt dazu, dass soziale Netzwerke und deren Profile zur Zielscheibe von Identitätsdieben geworden sind. Dieser Beitrag diskutiert die allgemeine Problematik persönliche Daten in sozialen Netzwerken zu veröffentlichen und versucht, mathematische Funktionen zu definieren, die von Angreifern genutzt werden könnten um raffinierten Identitätsdiebstahl zu betreiben. Es werden Hinweise auf Gegenmaßnahmen seitens der Anwender und Betreiber entsprechender Plattformen gegeben und eine Formel zur Messung der Kritikalität einer Identität im Internet wird diskutiert.

Methoden

Datenaggregation

Während der Aggregationsphase sammelt der Angreifer zunächst im großen Stil identitätsbezogene Profilinformationen und -relationen aus mehreren, unterschiedlichen sozialen Netzwerken. Mit Hilfe von *Crawling* ist es dem Angreifer möglich, viele Nutzerdaten automatisiert zu beziehen. Dies hat unter anderem den Grund, dass viele Turing-Tests in Form von so genannten Captchas unzureichend gegen heutige Angriffstechnologien gewappnet sind. Durch *Cracking* von Webanwendungen hat der Angreifer die Möglichkeit illegal an einen großen Pool Nutzerdaten zu gelangen. Zusätzlich kann durch den *Kauf von Nutzerdaten* der Vorrat an Nutzerdaten weiter aufgestockt werden.

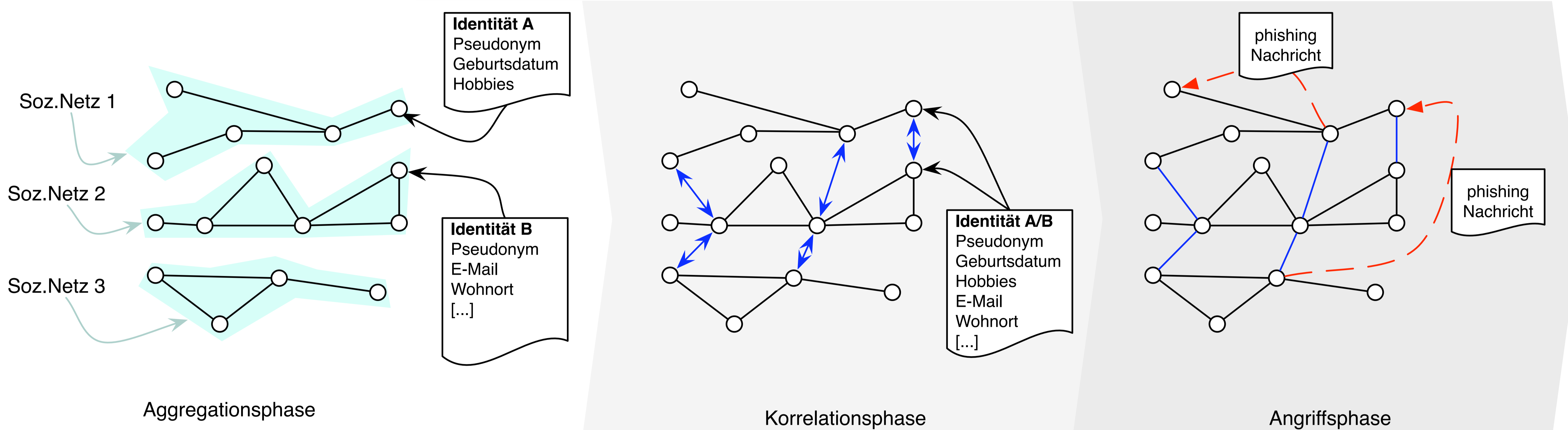
Datenkorrelation

Korrelationen -die der Angreifer herstellen kann- ergeben sich aus Übereinstimmungen oder Ähnlichkeiten verschiede-

ner Profilattribute und gegenseitig verweisende Relationen. Ist ist davon auszugehen, dass ein Internetnutzer ist meist gleich auf mehreren Web-2.0-Anwendungen aktiv ist: Auf einer Plattform speichert er dabei seine persönlichen Bilder, auf einer nächsten betreibt er sein persönliches Weblog und auf einer dritten pflegt er schließlich die Kontakte zu Freunden, Bekannten und Arbeitskollegen. Durch Ähnlichkeiten dieser verschiedenen Profile kann man Schlüsse über die Originalperson ziehen.

Angriffsszenarien

In der abschließenden dritten Phase werden die Daten analysiert und die entsprechenden Opfer ausgewählt. Der Angreifer macht sich dabei die so genannten Trusted Relation zwischen zwei Identitäten zu Nutze, indem er sich in betrügerischer Weise für eine Identität ausgibt und anschließend die zweite im Namen dieser kontaktiert.



Gegenmaßnahmen

Um Gegenmaßnahmen zu entwickeln, sollte zunächst der Kreislauf eines Angriffs näher betrachtet werden. In der Angriffsphase sind zwei grundlegende Schwächen der Opfer zu identifizieren: die Unfähigkeit eine Vertrauenswürdigkeit einer Nachricht einzuschätzen und die Reaktion auf eine nicht vertrauenswürdige Nachricht. Für dieses Verhaltensschema gab es vielerlei Vorschläge für Gegenmaßnahmen von der Etablierung eines Sicherheitsbewusstseins der Nutzer bis hin zu technischen Gegenmaßnahmen wie S/MIME und PGP. Bezüglich der Korrelationsphase können folgenden Gegenmaßnahmen die Auswirkung eines Angriffs vereiteln bzw. mindern: Die Nutzung von Pseudonymen in Kombination mit der Eingrenzung von persönlichen Attributen vermindert die Chance, dass der Angreifer eine Korrelation zwischen den Profilen herstellen kann. Der Aggregationsphase kann zum einen technisch mit Anti-Crawler Techniken begegnet werden. Zum anderen müssen die Nutzer über die möglichen Auswirkungen maßloser, zuordnungsfähiger, personenbezogener Daten besser informiert sein. Eine somit motivierte Eingrenzung von persönlichen Attributen, welche auch in der Korrelationsphase schützt, macht es dem Angreifer somit unmöglich personenbezogene Attribu-

te in der Sophisticated-Phishing-Nachricht weiterzuverwenden. Die Problematik der Umsetzung solcher Gegenmaßnahmen liegt aber nicht nur auf Seite des Nutzers. Auch Anbieter großer sozialer Netzwerke müssen auf Sicherheitsrisiken noch stärker hinweisen als bisher geschehen. Da die Betreiber solcher Netzwerke aber von privaten Datensätzen profitieren, haben sie zunächst wenig Interesse daran, dem Nutzer die Veröffentlichung zu erschweren bzw. zu verbieten.

Wir schlagen daher folgende Bewertung für personenbezogenen Daten vor: Personenbezogene Daten, wie den Namen oder das Geburtsdatum, definieren wir als $(\alpha_i \in \{0, 1\})$ mit dem jeweiligen Gewichtungsfaktor $(0 < g_{\alpha_i} \leq 1)$. Die Anzahl der Relationen bezeichnen wir als (ρ) und den Netz-abhängigen Gewichtungsfaktor als (g_{ρ_i}) . Aus diesen Definitionen lässt sich eine Formel zur Berechnung der Profilkritikalität κ ableiten:

$$\kappa = \frac{\sum g_{\alpha_i} \alpha_i + g_{\rho_i} \rho}{\nu}$$

wobei die Gewichtungsfaktoren $(g_{\alpha_i}, g_{\rho_i})$ und die Normierung (ν) entsprechen einer Skalierung, beispielsweise null bis zehn angepasst werden müssen.

Fazit & Ausblick

Für persönliche Daten gibt es vielerlei Missbrauchsmöglichkeiten, zunächst einmal die gewollte und/oder ungewollte unerwünschte Veröffentlichung durch private und/oder gar staatliche Stellen zu den verschiedensten Zwecken. Darüber hinaus lassen sich spätestens mit Hilfe der Korrelation von Profildaten auch rasterfahndungsähnliche Methoden umsetzen. Und nicht zuletzt birgt die Veröffentlichung von Daten auch ein großes ökonomisches Potenzial, dass beim Nutzer in Form von (unerwünschter) Werbung wieder ankommt. Angesichts der vorher geschilderten Szenarien ist dies wohl eher ein harmloses Übel. Eine sichere Separierung der Identitäten bleibt bisher eine noch ungelöste Aufgabe für die einzelnen Nutzer (Verwendung von Pseudonymen, allgemeine Datensparsamkeit, ...), für die Technik (Anonymisierungsdienste, ...) und für die Gesetzgebung (BDSG, europäische Datenschutzgesetzgebung, ...). Dabei scheint vor allem dringend notwendig, den Nutzer bei der Justierung und Einhaltung des von ihm gewünschten Datenschutzniveaus zu unterstützen. Konzepte, die versuchen, dieses Niveau anhand der Anzahl und Qualität der veröffentlichten Daten zu ermitteln, können ein erster Schritt zu einem besseren Verständnis sein und damit insgesamt zu einem höheren Datenschutzniveau beitragen.